

Algebra I

A Basic Course in Abstract Algebra



Rajendra Kumar Sharma
Sudesh Kumari Shah
Asha Gauri Shankar

Algebra-I

A Basic Course in Abstract Algebra

Rajendra Kumar Sharma

Indian Institute of Technology Delhi

Sudesh Kumari Shah

Sri Venkateswara College, University of Delhi

Asha Gauri Shankar

Lakshmibai College, University of Delhi

PEARSON

Delhi • Chennai • Chandigarh

Copyright © 2011 Dorling Kindersley (India) Pvt. Ltd.

Licenseses of Pearson Education in South Asia

No part of this eBook may be used or reproduced in any manner whatsoever without the publisher's prior written consent.

This eBook may or may not include all assets that were part of the print version. The publisher reserves the right to remove any material present in this eBook at any time.

ISBN 9788131760864

eISBN 9788131797624

Head Office: A-8(A), Sector 62, Knowledge Boulevard, 7th Floor, NOIDA 201 309, India

Registered Office: 11 Local Shopping Centre, Panchsheel Park, New Delhi 110 017, India

*To our
past, present and future
students*

This page is intentionally left blank.

Contents

<i>Preface</i>	<i>XIII</i>
<i>About the Authors</i>	<i>XV</i>
UNIT - 1	1
1 Sets and Relations	2
1.1 Sets	2
1.2 Exercise	6
1.3 Algebra of Sets	7
1.4 Exercise	18
1.5 Binary Relation	20
1.6 Exercise	33
1.7 Supplementary Exercises	37
1.8 Answers to Exercises	40
2 Binary Operations	49
2.1 Definition and Examples	49
2.2 Exercise	54
2.3 Introduction to Groups	55
2.4 Symmetries	58
2.5 Exercise	63
2.6 Solved Problems	65
2.7 Supplementary Exercises	68
2.8 Answers to Exercises	72

3	Functions	77
3.1	Definition and Representation	77
3.2	Images and Inverse Images	80
3.3	Types of Functions	83
3.4	Real Valued Functions	86
3.5	Some Functions on the Set of Real Numbers	88
3.6	Exercise	95
3.7	Inverse of a Function	97
3.8	Composition of Functions	100
3.9	Solved Problems	105
3.10	Exercise	109
3.11	Cardinality of a Set	110
3.12	Countable Sets	112
3.13	Exercise	121
3.14	Solved Problems	122
3.15	Supplementary Exercise	125
3.16	Answers to Exercises	127
4	Number System	134
4.1	Number Systems	134
4.2	Division Algorithm	141
4.3	Exercise	148
4.4	Greatest Common Divisor	150
4.5	Least Common Multiple	159
4.6	Exercise	162
4.7	Congruence Relation	164
4.8	Exercise	174
4.9	Supplementary Problems	177
4.10	Answers to Exercises	178
	UNIT - 2	183
5	Group: Definition and Examples	184
5.1	Definition of Group	184
5.2	Exercise	187
5.3	Groups of Numbers	187
5.4	Exercise	189
5.5	Groups of Residues	190

5.6	Exercise	193	
5.7	Groups of Matrices	194	
5.8	Exercise	197	
5.9	Groups of Functions	197	
5.10	Exercise	199	
5.11	Group of Subsets of a Set	199	
5.12	Exercise	200	
5.13	Groups of Symmetries	200	
5.14	Supplementary Exercise	204	
5.15	Answers to Exercises	207	
6	Group: Properties and Characterization		211
6.1	Properties of Groups	211	
6.2	Solved Problems	215	
6.3	Exercise	219	
6.4	Characterization of Groups	220	
6.5	Solved Problems	223	
6.6	Exercise	227	
6.7	Supplementary Exercises	228	
6.8	Answers to Exercises	229	
7	Subgroups		231
7.1	Criteria for Subgroups	231	
7.2	Solved Problems	235	
7.3	Exercise	237	
7.4	Centralizers, Normalizers and Centre	238	
7.5	Exercise	244	
7.6	Order of an Element	245	
7.7	Solved Problems	249	
7.8	Exercise	251	
7.9	Cyclic Subgroups	253	
7.10	Solved Problems	255	
7.11	Exercise	257	
7.12	Lattice of Subgroups	257	
7.13	Exercise	262	
7.14	Supplementary Exercises	263	
7.15	Answers to Exercises	265	

8	Cyclic Groups	271
8.1	Definition and Examples	271
8.2	Description of Cyclic Groups	273
8.3	Exercise	276
8.4	Generators of a Cyclic Group	276
8.5	Exercise	278
8.6	Subgroups of Cyclic Groups	280
8.7	Subgroups of Infinite Cyclic Groups	280
8.8	Subgroups of Finite Cyclic Groups	281
8.9	Number of Generators	283
8.10	Exercise	286
8.11	Solved Problems	287
8.12	Supplementary Exercise	290
8.13	Answers to Exercises	292
	UNIT - 3	297
9	Rings	298
9.1	Ring	298
9.2	Examples of Ring	299
9.3	Constructing New Rings	303
9.4	Special Elements of a Ring	304
9.5	Solved Problems	305
9.6	Exercise	309
9.7	Subrings	312
9.8	Exercise	316
9.9	Integral Domains and Fields	318
9.10	Examples	319
9.11	Exercise	326
9.12	Solved Problems	327
9.13	Supplementary Exercises	328
9.14	Answers to Exercise	332
	UNIT - 4	337
10	System of Linear Equations	338
10.1	Matrix Notation	340

10.2	Solving a Linear System	341
10.3	Elementary Row Operations (ERO)	342
10.4	Solved Problems	344
10.5	Exercise	353
10.6	Row Reduction and Echelon Forms	357
10.7	Exercise	371
10.8	Vector Equations	373
10.9	Vectors in \mathbb{R}^2	374
10.10	Geometric Descriptions of \mathbb{R}^2	374
10.11	Vectors in \mathbb{R}^n	377
10.12	Exercise	388
10.13	Solutions of Linear Systems	393
10.14	Parametric Description of Solution Sets	397
10.15	Existence and Uniqueness of Solutions	399
10.16	Homogenous System	403
10.17	Exercise	416
10.18	Solution Sets of Linear Systems	422
10.19	Exercise	430
10.20	Answers to Exercises	430

11 Matrices

441

11.1	Matrix of Numbers	441
11.2	Operations on Matrices	443
11.3	Partitioning of Matrices	451
11.4	Special Types of Matrices	455
11.5	Exercise	460
11.6	Inverse of a Matrix	463
11.7	Adjoint of a Matrix	464
11.8	Negative Integral Powers of a Non-singular Matrix	466
11.9	Inverse of Partitioned Matrices	467
11.10	Solved Problems	470
11.11	Exercise	473
11.12	Orthogonal and Unitary Matrices	476
11.13	Length Preserving Mapping	478
11.14	Exercise	481
11.15	Eigenvalues and Eigenvectors	483
11.16	Cayley Hamilton Theorem and Its Applications	490
11.17	Solved Problems	493
11.18	Exercise	500

11.19	Supplementary Exercises	501
11.20	Answers to Exercises	504
12	Matrices and Linear Transformations	509
12.1	Introduction to Linear Transformations	509
12.2	Exercise	513
12.3	Matrix Transformations	513
12.4	Surjective and Injective Matrix Transformations	514
12.5	Exercise	521
12.6	Linear Transformation	524
12.7	Exercise	528
12.8	The Matrix of a Linear Transformation	530
12.9	Exercises	532
12.10	Geometric Transformations of \mathbb{R}^2 and \mathbb{R}^3	534
12.11	Exercises	552
12.12	Supplementary Problems	553
12.13	Supplementary Exercise	555
12.14	Answers to Exercises	559
UNIT - 5		563
13	Vector Space	564
13.1	Definition and Examples	564
13.2	Exercise	573
13.3	Subspaces	574
13.4	Exercise	581
13.5	Linear Span of a Subset	583
13.6	Column Space	586
13.7	Exercise	591
13.8	Solved Problems	593
13.9	Exercise	595
13.10	Answers to Exercises	599
14	Basis and Dimension	601
14.1	Linearly Dependent Sets	601
14.2	Solved Problems	607
14.3	Exercise	613
14.4	Basis of Vector Space	615

14.5	Coordinates Relative to an Ordered Basis	617
14.6	Exercise	625
14.7	Dimension	627
14.8	Rank of a Matrix	634
14.9	Exercise	640
14.10	Solved Problems	644
14.11	Supplementary Exercises	645
14.12	Answers to Exercises	649
15	Linear Transformation	653
15.1	Definitions and Examples	653
15.2	Exercise	662
15.3	Range and Kernel	664
15.4	Exercise	673
15.5	Answers to Exercises	676
16	Change of Basis	678
16.1	Coordinate Mapping	678
16.2	Change of Basis	679
16.3	Procedure to Compute Transition Matrix $P_{\mathcal{B} \leftarrow \mathcal{B}}$ from Basis \mathcal{B}_1 to Basis \mathcal{B}_2	683
16.4	Exercise	687
16.5	Matrix of a Linear Transformation	691
16.6	Working Rule to Obtain $[T]_{\mathcal{B}_1 \mathcal{B}_2}$	693
16.7	Exercise	702
16.8	Supplementary Exercises	705
16.9	Answers to Exercises	711
17	Eigenvectors and Eigenvalues	719
17.1	Eigenvectors and Eigenspace	719
17.2	Solved Problems	723
17.3	Exercise	724
17.4	Characteristic Equation	726
17.5	Exercise	735
17.6	Diagonalization	737
17.7	Exercise	742
17.8	Supplementary Exercises	743
17.9	Answers to Exercises	745

18 Markov Process	748
18.1 Exercise	756
18.2 Answers to Exercises	758
<i>Index</i>	<i>761</i>

Preface

There are two schools of thought: one, a particular topic be chosen and taught at an advanced level; the other, topics be chosen and first taught at the introductory level and then at the advanced level. Each has its own advantages and disadvantages. In the first case, study becomes very focused and a significant level of the course can be achieved. But at the same time it leaves inter-related topics introduced together. Even the choice of examples becomes very limited. Thus, for the first timers a course on algebra requires to go through several books. Connecting topics becomes a difficult task. Different assumptions and notations too pose a big problem.

In the second approach, however, related topics are studied simultaneously. They are collected at one place. Though an advanced level coverage may not be taught at the first stage, a reasonably good introduction and sound background can be prepared. For a student coming to study a course on algebra for the first time, this book presents all the basic materials in one place and gives an opportunity to begin understanding the topics in a most easy and comfortable way. The presentation of the text is lucid. A large number of examples are used to explain the concepts. This also prepares the students to attempt exercises themselves. The book contains a large number of exercises, together with answers of varying difficulty. These help students build confidence. Difficult exercises follow simple ones. Graphics are also introduced. We have tried to make a complete textbook for a first course in Algebra. That is why it is *Algebra – I*. Two more volumes, *Algebra – II* and *Algebra – III*, will take the students to higher and sufficiently advanced levels, as is expected from a three-year undergraduate degree programme of any university or institute.

The entire text of *Algebra – I* is divided into six different units formed of related topics. Each unit is divided into chapters and each chapter into sections. All theorems, lemmas and examples are continuously numbered by a three-digit number. That is, $x.y.z$ means that the result is in Chapter x , Section y and within the section its serial number is z . At the end of each section, an exercise set is given. It is also numbered as a section. The answers to the exercises in a chapter are given at the end of the chapter. Each chapter begins with learning objectives and is concluded by a summary of the topics covered. An attempt has been made to make this book a complete resource book of a first course in Algebra.

We are grateful to a lot of people. It is not possible to include the names of all of them here. We thank a large large number of students who have helped us in bringing out this book in its present form.

Bhavya Chauhan, Sweta Mishra, Neha Makhijani and Parvesh Lathwal are some of them. We remain indebted to all those who have helped us in any manner in bringing out this book.

RAJENDRA KUMAR SHARMA
SUDESH KUMARI SHAH
ASHA GAURI SHANKAR

About the Authors

Rajendra Kumar Sharma is professor and head of the Department of Mathematics at the Indian Institute of Technology, Delhi. Earlier, he was in the faculty of Mathematics at IIT, Kharagpur. He has been teaching undergraduate and postgraduate classes for more than 20 years, and guided eight Ph.D. theses and more than 30 M.Tech. projects. Sharma has published more than 45 research papers in international journals. He has participated in several conferences, including the coveted International Congress of Mathematicians (ICM), 1994, in Zurich, Switzerland. Sharma has travelled widely and delivered talks at several places. He was a postdoctoral fellow in France and Germany for three years. Several students are working with him on sponsored projects. His main areas of research are Algebra and Cryptography.

Sudesh Kumari Shah is associate professor in the Department of Mathematics at Venketashwara College, University of Delhi. She has an experience of teaching undergraduate and postgraduate classes for more than 35 years at the college and in the north and south campuses of the University of Delhi. Shah has published research papers in several international journals. She is the co-author of *Mathematics for Life Sciences*. Her research interest is Algebra.

Asha Gauri Shankar is associate professor in the Department of Mathematics at Lakshmi Bai College, University of Delhi. She has been teaching undergraduate and postgraduate classes for more than 35 years at the college. She also has the experience of teaching in the north and south campuses of the University of Delhi; at Vivekananda College, University of Delhi; at Imperial College of Science, Technology and Medicine, London; and at the Institute of Advanced Studies, Meerut. She holds two Ph.D. degrees, one from Imperial College of Science, Technology and Medicine, University of London; and the other from Chaudhary Charan Singh University, Meerut. Shankar has been awarded the Bharat Excellence Award and a gold medal by Friendship Forum of India (FFI), which was conferred by Dr G.V.G. Krishnamurthy, former Election Commissioner of India, on 29 September 2010. Earlier, she was nominated by the Vice Chancellor, University of Delhi, to be honoured as ‘Teacher of Excellence’, which was conferred by Dr A. P. J. Abdul Kalam, former President of India, on 7 September 2009. Shankar is also a recipient of Shiksha Ratan Puraskar awarded by India International Federation Society and of Mahila Shree Award given by FFI. She has published several research papers in international journals. She has to her credit a research-level book *Numerical Integration of Differential Equations*. Her research interests are Mathematics Education and Numerical Analysis.

This page is intentionally left blank.

UNIT - 1

Chapter 1

Sets and Relations

In our daily life we come across words such as set, collection, group, clear etc... In this chapter we begin by giving a mathematical definition of the word "set" and proceed to study various types of relations on them. To make concepts easily comprehensible, lots of examples and diagrams, called Venn diagrams have been given. So, first thing first.

1.1 Sets

Definition 1.1. (*Set*): A set is a well defined collection of objects.

The adjective 'well defined' means that given an object, it should be possible to decide whether it belongs to the collection or not. There should not be any ambiguity. The objects that belong to a set are called its members or elements.

Example 1.1. The following are examples of sets:

- (i) Factors of 120.
- (ii) Roots of the equation $x^2 - 3x + 2 = 0$.
- (iii) Letters of the word Boole.
- (iv) The rivers of India originating in the Himalayas.
- (v) The students of Sri Venkateswara College taking admission in 2009.

Example 1.2. The following are not sets:

- (i) The collection of all intelligent teachers of Delhi University.
- (ii) The collection of all fat ladies in Shalimar Bagh, Delhi.
- (iii) The collection of all rich people in Delhi.
- (iv) The collection of all hardworking students of Lakshmibai College.
- (v) The collection of 9 natural numbers.

This is because the adjectives intelligent, fat, rich and hardworking are not well defined. These terms are all relative. The collection (v) is not a set as it is not well defined. The natural number 7 may or may not belong to the collection. If the statement is modified as 'the collection of first nine natural numbers', then it is a set and 7 is a member of the set.

Notation: Sets are usually denoted by capital letters and their members by lower-case letters. The statement “ a is an element of A ” is written as $a \in A$ and is read as “ a belongs to A ”, where \in is the Greek letter epsilon.

Example 1.3. $3 \in \mathbb{Z}$, $5 \in \mathbb{N}$, $-5 \notin \mathbb{N}$.

Definition 1.2. (Universal Set): The set from which we pick the elements to test whether the properties under consideration are satisfied or not, is called the Universal Set. This set may change depending on the context.

For instance, if we consider the set of all students obtaining more than 85% marks in mathematics, it is not clear from where do we have to pick these students. They can be picked from any one of the following sets:

- (i) All students of Sardar Patel School, Delhi who have appeared for class VIII.
- (ii) All students of Delhi University who appeared in the annual examination of 2009.
- (iii) All students of Lakshmibai College who appeared in first year examination of 2009.

Depending upon which of the sets (i), (ii) or (iii) we choose, our set defined above will change.

Sometimes the universal set is not mentioned, then it is understood from the context. We shall denote it by U .

Example 1.4. Let $A = \{x \in \mathbb{R} \mid -4 \leq x \leq 4\}$.

$$B = \{x \in \mathbb{Z} \mid -4 \leq x \leq 4\}.$$

$$C = \{x \in \mathbb{N} \mid -4 \leq x \leq 4\}.$$

In the above examples the universal sets are \mathbb{R} , \mathbb{Z} and \mathbb{N} respectively. Thus the sets A , B , C are different though the condition is the same in all the three cases, A is the interval $[-4, 4]$, $B = \{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$ and $C = \{1, 2, 3, 4\}$.

Description of a set: There are two ways of describing a set:

- (i) Tabular method or Listing method or Roster method.
- (ii) Set builder method or Property method or Rule method.

In the **Roster method**, all the elements are listed. Two elements are separated by a comma and the entire set of elements is enclosed by curly brackets (or braces). The elements should not be repeated, i.e., no element should be written more than once. Moreover, the order in which the elements are written is immaterial.

Example 1.5. 1. $S = \{1, 2, 3, 4, 5, 6\}$
is the set of the first 6 natural numbers.

2. The set of the letters of word ‘mathematics’ is $T = \{m, a, t, h, e, i, c, s\}$. Note that though the letter m occurs twice in the word, it is written only once when writing the set. In fact, if we write the letters in alphabetic order, then S can be written as

$$\{a, c, e, h, i, m, s, t\}.$$

3. $E = \{2, 4, 6, 8, \dots\}$ is the set of even natural numbers.

In the **Set builder method** the elements are described by means of a property which is possessed by all the elements.

Example 1.6. 1. In the preceding example, the set S can be written as

$$S = \{n \in \mathbb{N} \mid n \leq 6\}.$$

2. If A is the set of all alphabets, then the set T can be written as
 $T = \{x \in A \mid x \text{ is a letter of the word 'mathematics'}\}.$

3. The set E can be written as
 $E = \{x \in \mathbb{N} \mid x \text{ is an even natural number}\}.$

We can describe it as

$$E = \{x \in \mathbb{N} \mid x \text{ is divisible by } 2\}.$$

Now, we give some sets which are written in Roster form as well as Set builder form.

Example 1.7.

S. No.	Roster Form	Set Builder Form
1	$\{-3, -2, -1, 0, 1, 2\}$	$\{x \in \mathbb{Z} \mid -3 \leq x \leq 2\}$
2	$\{-2, -1, 0, 1, 2\}$	$\{x \in \mathbb{Z} \mid x^2 \leq 5\}$
3	$\{1, 2, 3\}$	$\{x \in \mathbb{N} \mid x^3 \leq 50\}$
4	$\{B, o, l, e\}$	$\{x \mid x \text{ is a letter of the word Boole}\}$

At times the Roster method is not good, for example in the set $\{\text{cat, dog, rabbit, ...}\}$ it is not clear what are the other elements of the set. But when the same set is written in set builder form, namely, $\{x \mid x \text{ is a mammal}\}$, it is clear which elements have to be included in the set.

Definition 1.3. (Empty set): A set which does not have any element is called the empty set (or null set or void set).

It is denoted by $\{\}$ or ϕ . The latter symbol is read as phi.

Example 1.8. 1. The set of all alive persons in India born before 1800

2. $\{x \in \mathbb{Z} \mid x^2 < 0\}$

3. $\{x \in \mathbb{Z} \mid x > 2 \text{ and } x < 1\}$

are all null sets.

Definition 1.4. (Singleton): A set consisting of exactly one element is called a singleton. It is written as $\{a\}$.

Example 1.9. 1. $\{\phi\}$ is a singleton whose only element is the null set ϕ .

2. $\{x \in \mathbb{Z} \mid x \geq 2 \text{ and } x \leq 2\} = \{2\}$.

Definition 1.5. (Equality of Sets): Two sets A and B are equal if and only if they have the same elements.

We write $A = B$. If two sets A and B are not equal, we write $A \neq B$.

Example 1.10. 1. If $A = \{2, 3, 5, 7\}$, $B = \{x \in \mathbb{N} \mid x \text{ is a prime number and } x < 8\}$, then $A = B$.

2. $A =$ letters of the word wolf, $B =$ letters of the word flow.

Then $A = \{w, o, l, f\}$, $B = \{f, l, o, w\}$, so that $A \neq B$.

3. If $A = \phi$, $B = \{\phi\}$, then $A \neq B$.

Definition 1.6. (Finite Set): A set is said to be finite if it is either empty or it is in one-to-one correspondence with $\{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$. The number of elements in a finite set A is denoted by $o(A)$.

A set which is not finite is infinite.

Example 1.11. 1. The set A of the months in a year is a finite set with $o(A)=12$.

2. $B = \{x \in \mathbb{Z} \mid x \text{ is divisible by } 2\}$, B is an infinite set.

3. The set of all natural numbers, integers, rationals, and reals are infinite sets.

4. The set of all persons living in India on Sep 1, 2009 is a finite set.

Definition 1.7. (Subset): If A and B are two sets such that every element of A is an element of B , then A is called a subset of B , we write $A \subseteq B$. If A is a subset of B and $A \neq B$, then we say that A is a proper subset of B and we write $A \subset B$.

When $A \subseteq B$, we may also say that A is contained in B . We can write this as $B \supseteq A$ and we say that B is a superset of A or B contains A .

Example 1.12. 1. Let $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 3, 4, 5\}$.

Then $A \subseteq B$. Since $5 \in B$ and $5 \notin A$, $\therefore A$ is a proper subset of B .

2. If A is the set of letters in the word 'algebra' and B is the set of letters in the word 'real',

then $A = \{a, b, e, g, l, r\}$, $B = \{a, e, l, r\}$. Clearly $B \subset A$.

Theorem 1.1. For any set A ,

(i) $\phi \subseteq A$

(ii) $A \subseteq A$

Proof:

(i) Suppose on the contrary $\phi \not\subseteq A$. Then there exists $x \in \phi$ such that $x \notin A$. This is absurd as ϕ does not contain any element. \therefore Our assumption is wrong, so that, $\phi \subseteq A$.

(ii) Since every element of A is an element of A , therefore $A \subseteq A$. □

Theorem 1.2. If A and B are two sets, then $A = B$, if and only if $A \subseteq B$ and $B \subseteq A$.

Proof: Left to the reader. □

The above theorem gives a practical way to prove the equality of two sets.

Definition 1.8. (Power Set): The set of all subsets of a given set A is called the power set of A . It is denoted by $\mathcal{P}(A)$.

In symbols, $\mathcal{P}(A) = \{B \mid B \subseteq A\}$.

Example 1.13. 1. If $A = \{1\}$, then $\mathcal{P}(A) = \{\{1\}, \phi\}$

$o(\mathcal{P}(A)) = 2$

2. If $A = \{x, y\}$, then

$\mathcal{P}(A) = \{\phi, \{x\}, \{y\}, \{x, y\}\}$

$o(\mathcal{P}(A)) = 4 = 2^2$.

3. If $A = \{p, q, r\}$, then

$\mathcal{P}(A) = \{\phi, \{p\}, \{q\}, \{r\}, \{p, q\}, \{p, r\}, \{q, r\}, \{p, q, r\}\}$

$o(\mathcal{P}(A)) = 8 = 2^3$.

Can you guess the number of elements in $\mathcal{P}(A)$, when number of element in A is given. It is interesting to note that the power set of the empty set is not empty. In fact, $\mathcal{P}(\phi) = \{\phi\}$.

Problem 1.1. *If A is a finite set containing n element, then $\mathcal{P}(A)$ has 2^n elements.*

Solution: Let $o(A) = n$.

For $0 \leq r \leq n$, the number of subsets of A containing r elements is ${}^n C_r$. Thus total number of subsets of A

$$\begin{aligned} &= {}^n C_0 + {}^n C_1 + \dots + {}^n C_n \\ &= 2^n. \end{aligned}$$

Hence $o(\mathcal{P}(A)) = 2^n$.

It follows from the above result that the power set of an infinite set is infinite.

1.2 Exercise

- Write the following sets in Roster Form:
 - $\{x \mid x \text{ is a natural number, } x = x^2\}$
 - $\{x \in \mathbb{N} \mid x \text{ is divisible by } 5\}$
 - $\{x \in \mathbb{Z} \mid x^4 - 64 = 0\}$
 - $\{a \in \mathbb{Z} \mid -1 \leq |a| \leq 5\}$
 - $\{x \in \mathbb{Z} \mid -6 \leq x \leq 8\}$
- Write the following sets in Set Builder Form:
 - $A = \{3, 6, 9, 12, 15, 18\}$
 - $B = \{1, 2\}$
 - $C = \{2, 5, 10, 17, 26, \dots\}$
 - $D = \{-3, -2, -1, 0, 1, 2, 3\}$
 - $E = \{1, -1, -i, i\}$, where $i^2 = -1$
- List 3 elements of the following sets:
 - $\{p \mid p \text{ is a four letter word ending with ice}\}$
 - $\{x + y\sqrt{7} \mid x, y \text{ are rationals}\}$
 - $\{x + y \mid x, y \in \mathbb{R}, x^2 + y^2 = 4\}$
 - $\{x + y \mid x, y \in \mathbb{Z}, x^2 + y^2 = 25\}$
 - $\{x \in \mathbb{Q} \mid (x^2 - 1)(x^2 - 2)(x^3 + 3x^2 + 2x) = 0\}$
 - $\{\frac{x}{y} \mid x \in \{0, 1, 2\}, y \in \{-1, 1\}\}$
- Let $P = \{2, 4, 6, 8, 10\}$, write a subset Q of P such that
 - $\{2, 10\} \subseteq Q$
 - $\phi \subseteq Q$
 - $\{4, 6\} \subset Q$
 - $Q \subseteq \{4, 10\}$
 - $Q \subset \{6, 8\}$
 - $Q \not\subseteq \{2, 4, 6\}$
 - $\{2, 4\} \subseteq Q \subseteq \{2, 4, 8\}$

5. Let $A = \{p, q, r, s, t\}$ write a subset of A such that
- (i) q belongs to the set A
 - (ii) It contains 3 elements
 - (iii) It contains s and t
 - (iv) It does not contain r or s
 - (v) It contains none of p, q, r, s or t .
6. Write the power set of A and tell the number of element in it, where
- (i) $A = \phi$
 - (ii) $A = \{\phi\}$
 - (iii) $A = \{w, x, y, z\}$
7. Let $A = \{p, q, r\}$. Indicate whether the following are true or false, with justification.
- (i) $\phi \in A$
 - (ii) $\phi \subseteq A$
 - (iii) $p \subseteq A$
 - (iv) $p \in A$
 - (v) $A \in A$
 - (vi) $A \subseteq A$
 - (vii) $A \in \mathcal{P}(A)$
 - (viii) $\phi \in \mathcal{P}(A)$
 - (ix) $\{q, r\} \subseteq A$
 - (x) $\{q, r\} \in A$
 - (xi) $\{q, r\} \subset \mathcal{P}(A)$
 - (xii) $\{q, r\} \in \mathcal{P}(A)$

1.3 Algebra of Sets

We now discuss some ways in which two or more sets can be combined to give a new set.

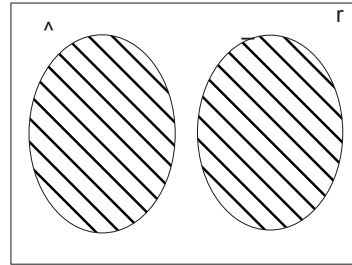
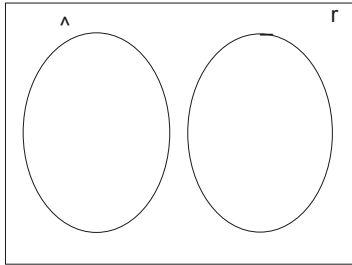
Definition 1.9. (Union of two sets): Let A and B be two sets. The set of all elements which belong to A or B or both is called the union of A and B . It is denoted by $A \cup B$.

Symbolically $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$.

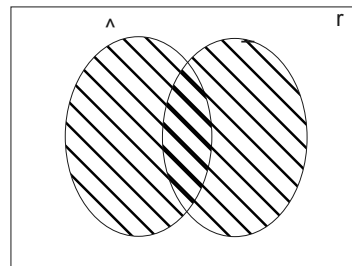
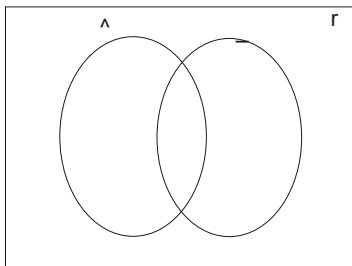
Example 1.14. If $A = \{1, 2, 3, 4\}$, $B = \{3, 5, 7, 11\}$, then $A \cup B = \{1, 2, 3, 4, 5, 7, 11\}$.

$A \cup \phi = A$, $B \cup \{\phi\} = \{3, 5, 7, 11, \phi\}$.

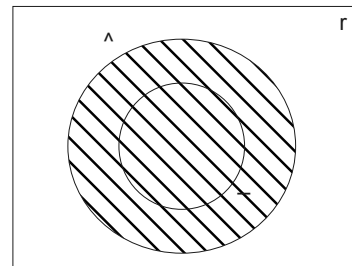
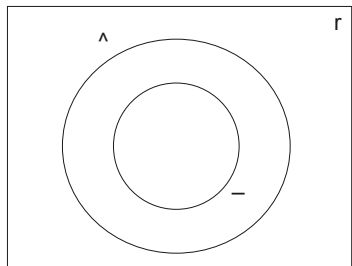
Using Venn diagram, the shaded region represents $A \cup B$ in different cases.



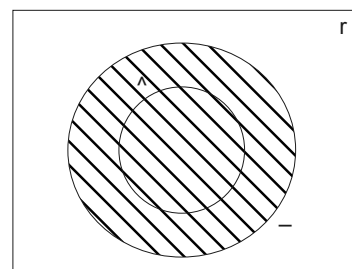
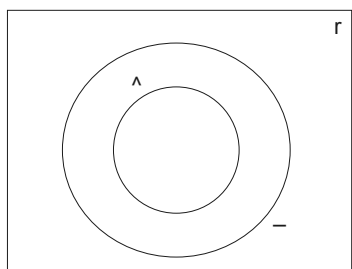
$$A \cup B$$



$$A \cup B$$



$$A \cup B = A$$

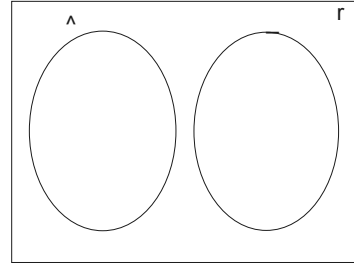
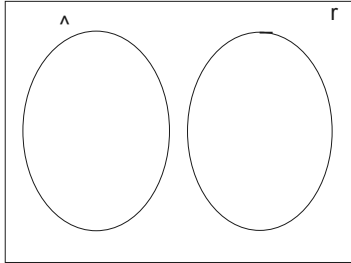


$$A \cup B = B$$

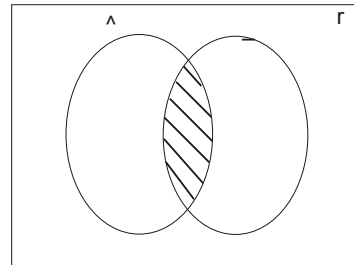
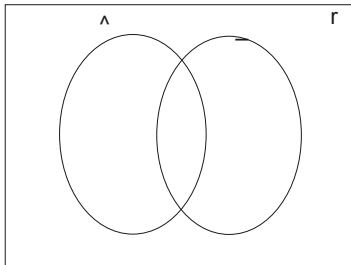
Definition 1.10. (Intersection of two sets): Let A and B be two sets. The intersection of A and B is the set of all those elements which are in A as well as in B . It is denoted by $A \cap B$.
Symbolically $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$.

Example 1.15. If $A = \{a, b, c, d, e\}$, $B = \{a, e, i, o, u\}$, then $A \cap B = \{a, e\}$, $A \cap \phi = \phi$.

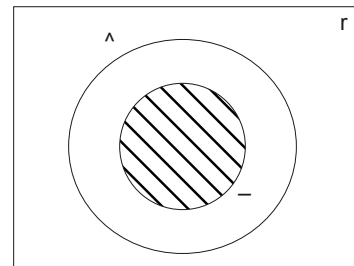
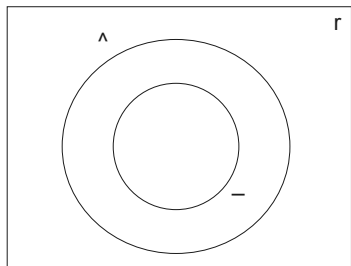
Using Venn diagram, the shaded region represents $A \cap B$ in different cases.



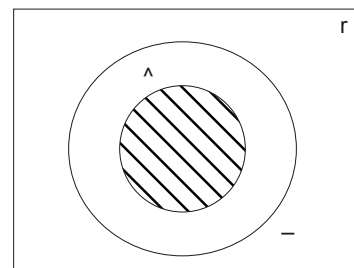
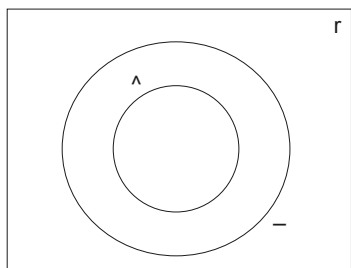
$$A \cap B = \phi$$



$$A \cap B$$



$$A \cap B = B$$



$$A \cap B = A$$

Definition 1.11. (Disjoint sets): Two sets A and B are said to be disjoint if $A \cap B = \phi$.

Example 1.16. A = set of all vowels, B = set of all consonants.
Then $A \cap B = \phi$.

The following results, though simple to prove, are important in set theory.

Theorem 1.3. *If A , B and C are three subsets of the Universal set U , then*

- $A \subseteq A \cup B$, $B \subseteq A \cup B$
- $A \cup B = B \cup A$
- $A \cup (B \cup C) = (A \cup B) \cup C$
- $A \cup A = A$
- $A \cup U = U$
- $A \cup \phi = A$
- $A \cap B \subseteq A$, $A \cap B \subseteq B$
- $A \cap B = B \cap A$
- $A \cap (B \cap C) = (A \cap B) \cap C$
- $A \cap A = A$
- $A \cap U = A$
- $A \cap \phi = \phi$

Proof: Left to the reader. □

Since the union of three sets is associative, therefore there is no need to use parentheses, we can simply write $A \cup B \cup C$. Similarly if we have n sets A_1, A_2, \dots, A_n , we can write

$$A_1 \cup A_2 \cup \dots \cup A_n \text{ as } \bigcup_{i=1}^n A_i.$$

Similarly, if Λ is some index set and for each $\lambda \in \Lambda$, there is defined a set A_λ , then union of all these sets A_λ is written as $\bigcup_{\lambda \in \Lambda} A_\lambda$. Similar notation holds for intersection. Note that if union and intersection are used in the same expression then it is essential to use parentheses. That is $A \cap B \cup C$ is not well defined, as the two sets $(A \cap B) \cup C$ and $A \cap (B \cup C)$ are different. This is shown by following example.

Example 1.17. *Let $A = \{a, b, c, d, e\}$, $B = \{a, e, i\}$, $C = \{b, d, e, f, g\}$.*

Then $A \cap B = \{a, e\}$, $(A \cap B) \cup C = \{a, b, d, e, f, g\}$.

$B \cup C = \{a, b, d, e, f, g, i\}$, $A \cap (B \cup C) = \{a, b, d, e\}$.

Hence $(A \cap B) \cup C \neq A \cap (B \cup C)$.

In fact union (intersection) is distributive over intersection (union).

Theorem 1.4. *If A , B , C are any three sets, then*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

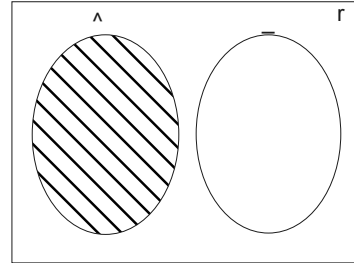
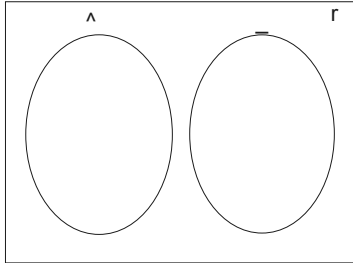
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Proof: Try yourself. □

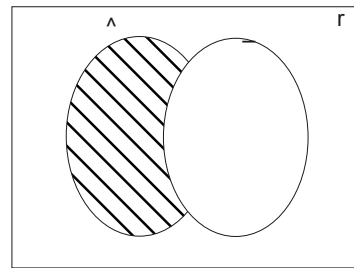
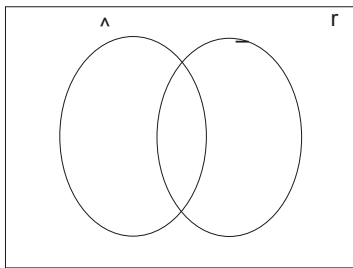
Definition 1.12. (Difference of two sets): *Let A and B be two sets. The difference of A and B in that order is the set of all elements of A which do not belong to B . It is denoted by $A \sim B$ or $A \setminus B$ or $A - B$. It is also called the complement of B in A .*

Symbolically $A \setminus B = \{x \in A \mid x \notin B\}$.

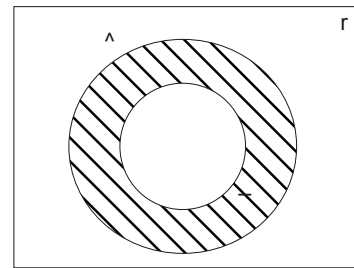
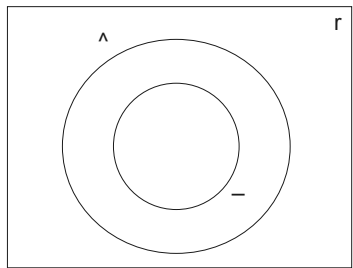
Using Venn diagram, $A \setminus B$ is represented by the shaded region.



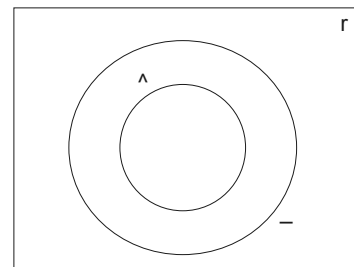
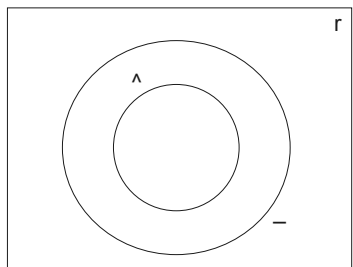
$$A \setminus B = A$$



$$A \setminus B$$



$$A \setminus B$$



$$A \setminus B = \phi$$

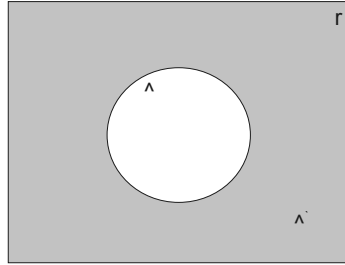
From the definition it is obvious that

- (i) $A \setminus A = \phi$
- (ii) $A \setminus B \subseteq A$.

Example 1.18. Let $A = \{a, b, c, d, e, f, g, h, i\}$, $B = \{a, c, e, g, i, o, u\}$,
 $C = \{p, q, r, s\}$. Then
 $A \setminus B = \{b, d, f, h\}$
 $A \setminus A = \{\} = \phi$
 $B \setminus A = \{o, u\}$
 $A \setminus C = \{a, b, c, d, e, f, g, h, i\} = A$
 $C \setminus A = \{p, q, r, s\} = C$.

Definition 1.13. (Complement of a set): The complement of a set A is the difference of U and A . It is the complement of A in U , the universal set. It is denoted by A' , A^c , \bar{A} or $U \setminus A$.

We shall use A^c for the complement of A . In the Venn diagram, the shaded region represents A^c .



Example 1.19. Let U be the set of natural numbers
 $A =$ set of all multiples of 3
 $B =$ set of all prime numbers.
Then $A^c = \{x \in \mathbb{N} \mid x \text{ is not a multiple of } 3\}$.
Clearly $19 \in A^c$, $30 \notin A^c$
 $A^c = \{1, 2, 4, 5, 7, 8, 10, 11, \dots\}$
 $B^c = \{x \in \mathbb{N} \mid x \text{ is not a prime number}\}$
Clearly $1 \in B^c$ ($\because 1$ is not a prime), $2 \notin B^c$.

The following results hold for complementation.

Theorem 1.5. Let A and B be any two sets. Then

- (i) $(A^c)^c = A$
- (ii) $A \cup A^c = U$
- (iii) $A \cap A^c = \phi$
- (iv) $\phi^c = U$
- (v) $U^c = \phi$
- (vi) $A \setminus B = A \cap B^c$
- (vii) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$
- (viii) $(A \cup B)^c = A^c \cap B^c$, De Morgan's law
- (ix) $(A \cap B)^c = A^c \cup B^c$, De Morgan's law
- (x) $(A \setminus B) \setminus C = A \cap (B \cup C)^c$

Proof: We shall prove only (ix)

$$\begin{aligned} x \in (A \cup B)^c &\iff x \notin A \cup B \\ &\iff x \notin A \text{ and } x \notin B \\ &\iff x \in A^c \text{ and } x \in B^c \end{aligned}$$

$\iff x \in A^c \cap B^c$

Hence $(A \cup B)^c = A^c \cap B^c$. □

The following results are useful in solving problems as they give the number of elements in the union, intersection and complements of finite sets in terms of the number of elements of the sets.

Theorem 1.6. *If A, B, C are finite subsets of a universal set U , then*

- (i) $o(A \cup B) = o(A) + o(B)$, if A and B are disjoint sets.
- (ii) $o(A \setminus B) = o(A) - o(A \cap B)$.
- (iii) $o(A \cup B) = o(A) + o(B) - o(A \cap B)$.
- (iv) $o(A \cup B \cup C) = o(A) + o(B) + o(C) - o(A \cap B) - o(B \cap C) - o(C \cap A) + o(A \cap B \cap C)$.
- (v) $o(A^c) = o(U) - o(A)$, if U is finite.

Proof:

- (ii) Since $A = (A \setminus B) \cup (A \cap B)$ and $(A \setminus B) \cap (A \cap B) = \phi$.
 \therefore By (i)

$$\begin{aligned} o(A) &= o(A \setminus B) + o(A \cap B) \\ \text{or } o(A \setminus B) &= o(A) - o(A \cap B) \end{aligned}$$

- (iii) $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$
and the sets $A \setminus B, B \setminus A$ and $A \cap B$ are mutually pairwise disjoint.
Applying (ii) the result follows.

- (iv)

$$\begin{aligned} o(A \cup B \cup C) &= o(A \cup B) + o(C) - o((A \cup B) \cap C) \quad \text{using (iii)} \\ &= o(A) + o(B) - o(A \cap B) + o(C) - o((A \cap C) \cup (B \cap C)) \\ &= o(A) + o(B) + o(C) - o(A \cap B) - (o(A \cap C) - o(B \cap C) + o((A \cap C) \cap (B \cap C))) \quad \text{using (iii)} \\ &= o(A) + o(B) + o(C) - o(A \cap B) - o(A \cap C) - o(B \cap C) + o(A \cap B \cap C) \end{aligned}$$

Hence proved.

- (i) and (v) are left to the reader □

Problem 1.2. *If B is a finite set and $A \subseteq B$ such that $o(A) = o(B)$, then $A = B$.*

Solution: Let $o(B) = n$.

Let, if possible $A \neq B$.

Then $B \setminus A \neq \phi$ so that $o(B \setminus A) \geq 1$.

Now,

$$\begin{aligned} o(B \setminus A) &= o(B) - o(A \cap B) \\ &= o(B) - o(A) \quad (\because A \cap B = A, \text{ as } A \subseteq B) \\ &= 0 \end{aligned}$$

which contradicts the fact that $o(B \setminus A) \geq 1$. Hence our assumption is wrong, so that $A = B$.

The above result fails to hold if A and B are infinite.

Consider

$A =$ set of even integers

$B =$ set of integers

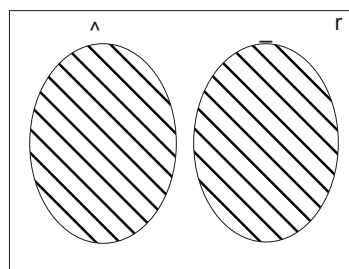
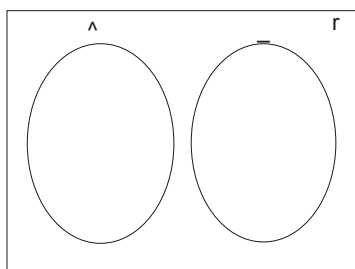
Then $A \subset B$, A and B are both infinite sets.

Definition 1.14. (Symmetric difference of two sets): Let A and B be two sets. The symmetric difference of A and B is the set of elements which are in A or B but not in both. It is denoted by $A \Delta B$.

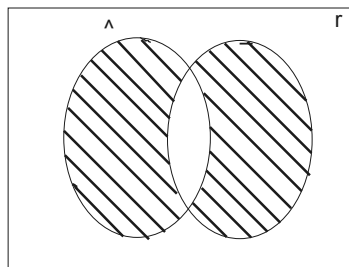
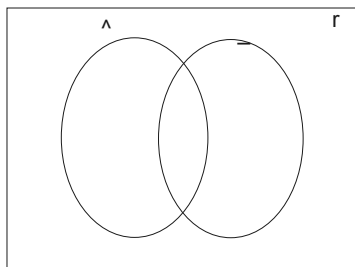
Symbolically, $A \Delta B = \{x \mid x \in A \cup B, x \notin A \cap B\}$.

Thus $A \Delta B = (A \cup B) \setminus (A \cap B)$.

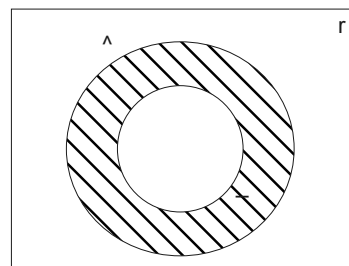
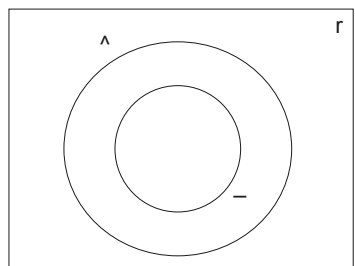
Using Venn diagram, the shaded region represents $A \Delta B$.



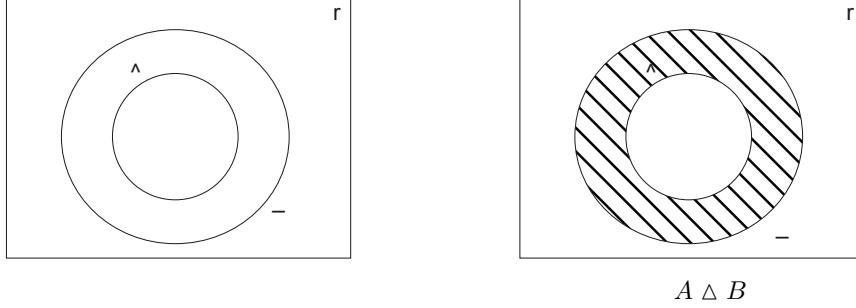
$A \Delta B$



$A \Delta B$



$A \Delta B$



From the definition, we get

$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$

Example 1.20. In \mathbb{N} , let

$$A = \{x \mid x \text{ is a multiple of } 4\}$$

$$B = \{x \mid x \text{ is a multiple of } 6\}$$

$$\begin{aligned} \text{Then } A \Delta B &= \{x \mid x \text{ is a multiple of } 4 \text{ or } 6, \text{ but not of both}\} \\ &= \{4, 6, 8, 16, 18, 20, 28, 30, \dots\} \end{aligned}$$

Problem 1.3. If A, B and C are any sets, then

$$(i) \quad A \Delta B = B \Delta A$$

$$(ii) \quad (A \Delta B) \Delta C = A \Delta (B \Delta C)$$

Solution:

$$\begin{aligned} (i) \quad A \Delta B &= (A \cup B) \setminus (A \cap B) \\ &= (B \cup A) \setminus (B \cap A) \\ &= B \Delta A \end{aligned}$$

(ii) Let $A, B, C \in \mathcal{P}(S)$

$$\begin{aligned} (A \Delta B) \Delta C &= ((A \cap B') \cup (B \cap A')) \Delta C \\ &= [\{(A \cap B') \cup (B \cap A')\} \cap C'] \cup [C \cap \{(A \cap B') \cup (B \cap A')\}] \\ &= [((A \cap B') \cap C') \cup ((B \cap A') \cap C')] \cup [C \cap \{(A' \cup B) \cap (B' \cup A)\}] \\ &= [(A \cap B' \cap C') \cup (B \cap A' \cap C')] \cup [C \cap \{(A' \cup B) \cap (B' \cup A)\}] \end{aligned}$$

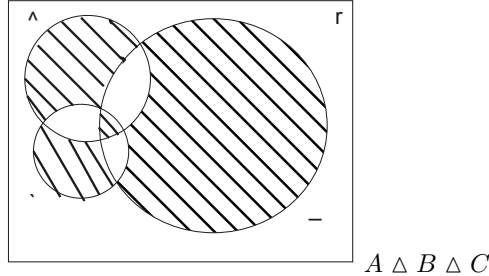
Now,

$$\begin{aligned} (A' \cup B) \cap (B' \cup A) &= \{(A' \cup B) \cap B'\} \cup \{(A' \cup B) \cap A\} \\ &= (A' \cap B') \cup (B \cap B') \cup (A' \cap A) \cup (B \cap A) \\ &= (A' \cap B') \cup (B \cap A) \end{aligned}$$

$$\begin{aligned} \text{Hence } (A \Delta B) \Delta C &= (A \cap B' \cap C') \cup (B \cap A' \cap C') \cup [C \cap \{(A' \cap B') \cup (B \cap A)\}] \\ &= (A \cap B' \cap C') \cup (B \cap A' \cap C') \cup (C \cap A' \cap B') \cup (C \cap B \cap A) \end{aligned}$$

$$\begin{aligned} \text{Similarly } A \Delta (B \Delta C) &= (A \cap B' \cap C') \cup (B \cap A' \cap C') \cup (C \cap A' \cap B') \cup (A \cap B \cap C) \\ \text{Hence } (A \Delta B) \Delta C &= A \Delta (B \Delta C). \end{aligned}$$

In view of the above result, we need not put any parenthesis while writing the symmetric difference of 3 sets. Using Venn diagram, the shaded portion shows $A \Delta B \Delta C$.



Thus the parentheses can be dropped while writing the symmetric difference of n sets.

More generally, it can be proved that the symmetric difference of n sets A_1, A_2, \dots, A_n , written as $A_1 \Delta A_2 \Delta \dots \Delta A_n$, is the set of those elements which are members of an odd number of the sets $A_i, i=1, 2, \dots, n$.

Definition 1.15. (Cartesian product of sets): Let A and B be two sets. Then the cartesian product of A and B is the set $\{(a, b) \mid a \in A, b \in B\}$. It is written as $A \times B$. We read it as A cross B . If one of the sets A or B is the null set then $A \times B$ is defined as the null set.

The elements of $A \times B$ are called ordered pairs. This is because the order of the elements is important. Thus if $a \neq b, (a, b) \neq (b, a)$.

If $(a_1, b_1), (a_2, b_2) \in A \times B$, then

$$(a_1, b_1) = (a_2, b_2)$$

if and only if $a_1 = a_2$ and $b_1 = b_2$.

More generally, the Cartesian product of n sets A_1, A_2, \dots, A_n is

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i, 1 \leq i \leq n\}.$$

Example 1.21. Let $A = \{a, e, i\}, B = \{p, q\}$,
then $A \times B = \{(a, p), (a, q), (e, p), (e, q), (i, p), (i, q)\}$
 $B \times A = \{(p, a), (p, e), (p, i), (q, a), (q, e), (q, i)\}$
 $B \times B = \{(p, p), (p, q), (q, p), (q, q)\}$

Thus, we see that

- (i) $A \times B \neq B \times A$
- (ii) number of elements in $A \times B = 6 = 3 \times 2 = o(A) \times o(B)$
- (iii) number of elements in $B \times A = 6 = 2 \times 3 = o(B) \times o(A)$
- (iv) number of elements in $B \times B = 4 = 2 \times 2 = o(B) \times o(B)$

More generally, if A and B are finite sets with m and n elements respectively, then number of elements in $A \times B = mn = o(A) \times o(B)$.

Theorem 1.7. If A, B, C are three sets, then

$$\begin{aligned} A \times (B \cup C) &= (A \times B) \cup (A \times C) \\ A \times (B \cap C) &= (A \times B) \cap (A \times C) \end{aligned}$$

Proof: Let $(x, y) \in A \times (B \cup C)$.

$\therefore x \in A$ and $y \in B \cup C$

$\Rightarrow x \in A$ and $y \in B$ or $y \in C$

$\Rightarrow (x, y) \in A \times B$ or $(x, y) \in A \times C$

$\Rightarrow (x, y) \in (A \times B) \cup (A \times C)$

Hence $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C) \dots (1)$.

Let $(a, b) \in (A \times B) \cup (A \times C)$.

$\therefore (a, b) \in A \times B$ or $(a, b) \in A \times C$

$\Rightarrow a \in A, b \in B$ or $a \in A, b \in C$

$\Rightarrow a \in A$, and $b \in B \cup C$

$\Rightarrow (a, b) \in A \times (B \cup C)$

Hence $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C) \dots (2)$

(1) and (2) $\Rightarrow A \times (B \cup C) = (A \times B) \cup (A \times C)$.

The proof of other part is left to the reader. \square

Problem 1.4. Let X, A and B be three sets such that $X \cap A = X \cap B$ and $X \cup A = X \cup B$. Prove that $A = B$.

Solution: We show that $A \subseteq B$

Let $x \in A$

Two cases arise:

Case 1: $x \in X$

Then $x \in A \cap X = X \cap B$

$\Rightarrow x \in B$

Case 2: $x \notin X$

Then $x \in A \Rightarrow x \in A \cup X = B \cup X$

$\Rightarrow x \in B$ ($\because x \notin X$)

Hence in each case $x \in A \Rightarrow x \in B$. So that $A \subseteq B$.

Similarly $B \subseteq A$. Hence $A = B$.

Problem 1.5. For any sets A and B prove that $A \cap B = A$ if and only if $A \subseteq B$

Solution: We first prove that $A \cap B = A \Rightarrow A \subseteq B$.

Since $A \cap B \subseteq B$, $\therefore A \subseteq B$.

Conversely, we prove that $A \subseteq B \Rightarrow A \cap B = A$.

By definition $A \cap B \subseteq A$.

If $x \in A$ then $x \in B$ ($\because A \subseteq B$).

Hence $x \in A \cap B$, so that

$A \subseteq A \cap B$.

Thus $A \cap B = A$.

Problem 1.6. Prove or disprove the following:

(i) $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$

(ii) $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$

(iii) $\mathcal{P}(A \setminus B) = \mathcal{P}(A) \setminus \mathcal{P}(B)$

Solution:

(i) $X \in \mathcal{P}(A \cap B)$

$\Leftrightarrow X \subseteq A \cap B$

$\Leftrightarrow X \subseteq A$ and $X \subseteq B$

$\Leftrightarrow X \in \mathcal{P}(A)$ and $X \in \mathcal{P}(B)$

$\Leftrightarrow X \in \mathcal{P}(A) \cap \mathcal{P}(B)$

Hence $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$, so the result is proved.

(ii) The result is not true.

Let $A = \{1, 2, 3\}$, $B = \{2, 3, 4\}$

Then $A \cup B = \{1, 2, 3, 4\}$, $C = \{1, 4\} \in \mathcal{P}(A \cup B)$
 but $C \notin \mathcal{P}(A)$ and $C \notin \mathcal{P}(B)$
 so that $C \notin \mathcal{P}(A) \cup \mathcal{P}(B)$.

$\therefore \mathcal{P}(A \cup B) \neq \mathcal{P}(A) \cup \mathcal{P}(B)$.

(iii) The result is not true. Choose A and B as in (ii) above. Then $A \setminus B = \{1\}$. So $\mathcal{P}(A \setminus B) = \{\phi, \{1\}\}$, $\{1, 2\} \in \mathcal{P}(A) \setminus \mathcal{P}(B)$ but $\{1, 2\} \notin \mathcal{P}(A \setminus B)$.

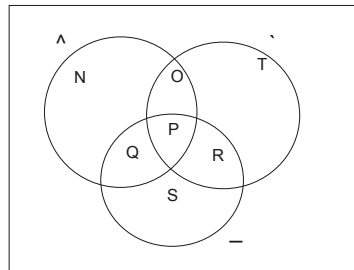
Hence $\mathcal{P}(A \setminus B) \neq \mathcal{P}(A) \setminus \mathcal{P}(B)$.

1.4 Exercise

- Let A denote the set of letters of the word 'mathematics', B denote the set of letters of the word 'algebra' and C denote the letters of the word 'analysis'.
 Find $A \cap C$, $A \cup B$, $(A \cup B) \cap C$, $A \Delta B$, $(B \setminus C) \setminus A$, $A \setminus (B \setminus C)$, $(A \cup B \cup C)^c$, $C \times (A \cap B)$.
- If $A = (-8, 2)$, $B = (-1, 5)$, write $A \cup B$, $A \cap B$, $A \setminus B$, $(A \cup B)^c$ as an interval.
- Let $X = \{x \in \mathbb{N} \mid x < 8\}$
 $Y = \{x \in \mathbb{Z} \mid |x + 1| \leq 5\}$
 $Z = \{x \in \mathbb{R} \mid x^5 - 3x^3 - 4x = 0\}$.
 Find $X \cap Y$, $X \cup Z$, $(X \cap Y) \times Z$, $(X \setminus Y) \setminus Z$, $X \Delta Z$, $(Y \Delta Z)$.
- If $A = \{1, 2, 4, 6\}$ and $B = \{1, 2, 3\}$,
 find $(A \times B) \cup (B \times A)$, $(A \times B) \cap (B \times A)$, $(A \times B) \setminus (B \times A)$.
- For any two sets A and B , prove that $(A \setminus B) \cap (B \setminus A) = \phi$.
- Using $(A \cup B)^c = A^c \cap B^c$, prove that $(A \cap B)^c = A^c \cup B^c$.
- Find the necessary and sufficient conditions for
 - $A \cup B = A$
 - $A \setminus B = A$
 - $A \Delta B = A$
 - $A \cap B = A \cup B$
- Let A , B , X and Y be sets such that $A \cup B = X \cup Y$, $A \cap B = X \cap Y = \phi$. Show that $X = \phi$ if and only if $B = (X \cap A) \cup (Y \cap B)$.
- If A , B , C are sets show that
 - $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
 - $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.
- Prove or disprove the following statements

- (i) $(A \cup B) \times C = (A \cup C) \times (B \cup C)$
 - (ii) $(A \cap B) \times C = (A \cap C) \times (B \cap C)$
 - (iii) If $A \subseteq C$ and $B \subseteq D$, then $A \times B \subseteq C \times D$
 - (iv) If $A \times B \subseteq C \times D$, then $A \subseteq C$ and $B \subseteq D$
 - (v) $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$
 - (vi) $A^c \times B^c = (A \times B)^c$
 - (vii) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$
 - (viii) $(A \cup B) \times C = (A \times C) \cup (B \times C)$.
 - (ix) $(A \cap B) \times C = (A \times C) \cap (B \times C)$.
 - (x) $(A \Delta B) \times C = (A \times C) \Delta (B \times C)$.
11. Let \mathbb{N} denote the set of natural numbers, \mathbb{Z} the set of integers, E the set of even integers and P the set of all prime numbers. Express the following statements in set theoretic notation.
- (i) Not every natural number is prime.
 - (ii) 2 is an even number which is also prime.
 - (iii) 3 is an odd prime.
 - (iv) Every natural number is an integer but not vice versa.
 - (v) There exists an integer which is not a natural number.
 - (vi) Every prime is odd.
12. If A and B are two sets having m and n elements respectively, prove that $A \times B$ has mn elements.
13. If L is a straight line and E is an ellipse in a plane, what are all the possible values of $o(L \cap E)$?
14. If $A = \{n \in \mathbb{N}: n \text{ is a multiple of } 12\}$
 $B = \{n \in \mathbb{N}: n \text{ is a multiple of } 18\}$
 Find $A \cup B$, $A \cap B$, $(A \cup B) \setminus (B \cap A)$, $A \times B$.
15. Let $U =$ set of all quadrilaterals in a plane, P , R , T and S be the subsets of U defined as follows:
 $P =$ set of all parallelograms
 $R =$ set of all rhombus
 $T =$ set of all rectangles
 $S =$ set of all squares
 Find the relationships between P , R , T and S in terms of containment.
16. In a survey of 100 delegates attending a conference, the number of delegates who knew one or more of the 3 languages Tamil, Punjabi, and Bangla was as follows: Tamil 28, Punjabi 30, Bangla 42; Tamil and Bangla 10; Tamil and Punjabi 8; Punjabi and Bangla 5. Only 3 people know all the three languages.
- (i) How many did not know any language at all?
 - (ii) How many knew only Bangla?
17. In a class 70% of the students like mango, 80% like bananas, 75% like apples, 85% like grapes and $x\%$ like all the four fruits. Find the minimum value that is possible for x .

18. Rakesh and Geeta are husband and wife. Geeta has 7 married friends and Rakesh has 5 married friends. They arrange a party and invite their friends with their partners. If all the friends come to the party, what is the maximum and minimum number of
- guests in the party.
 - common guests in the party.
19. For 3 sets A, B, C regions are labelled as below. The sets A, B, C are described as below:
 A : set of all women
 B : set of lawyers
 C : set of cricket lovers



Express the following regions in the terms of the sets A, B, C :

- region labelled 1
 - region labelled 2
 - region labelled 3
 - region labelled 5 or 7
 - region labelled 1, 4 or 6
20. Describe the persons represented by the regions in Q.19.

1.5 Binary Relation

Given two sets A and B at times, we are interested in associating elements of A with elements of B . The pairs of associated elements form a subset of $A \times B$. This motivates the following definition:

Definition 1.16. (Binary Relation): Let A and B be two sets. A binary relation from A to B is a subset of $A \times B$. A subset of $A \times A$ is called a binary relation on A .

The empty set (called void or null relation) and the entire cartesian product $A \times B$ (called universal relation) are always binary relations from A to B , though they are not as interesting as certain non-empty proper subsets of $A \times B$. If $R \subseteq A \times B$ and if $(a, b) \in R$ we say that 'a is R related to b' and we may write aRb .

Example 1.22. Let $A =$ set of all students of St. Xaviers school

$B = \{\text{hockey, football, badminton, cricket, volley ball, table tennis, basket ball}\}$

$R_1 = \{(a, b) \in A \times B \mid \text{student } a \text{ plays game } b\}$

$R_2 = \{(a, b) \in A \times B \mid \text{student } a \text{ plays lawn tennis}\}$

$R_3 = \{(a, b) \in A \times B \mid \text{student } a \text{ plays hockey or cricket}\}$.

Then R_1, R_2, R_3 are relation from A to B . Note that $R_2 = \phi$.

Example 1.23. $R_1 = \{(a, b) \in \mathbb{Z} \times \mathbb{N} \mid b = a^2\}$

$= \{(1, 1), (-1, 1), (-2, 4), (2, 4), \dots\}$

$R_2 = \{(a, b) \in \mathbb{Z} \times \mathbb{N} \mid b = |a|\}$

$= \{(-1, 1), (1, 1), (-2, 2), (2, 2), \dots\}$

are relations from \mathbb{Z} to \mathbb{N} .

Example 1.24. $R_3 = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid b = a+1\}$

$= \{(1, 2), (-1, 0), \dots\}$. $R_4 = \{(1, -1), (2, 3), (-29, 341)\}$. Then R_3 and R_4 define a relation on \mathbb{Z} .

Graph of a Relation

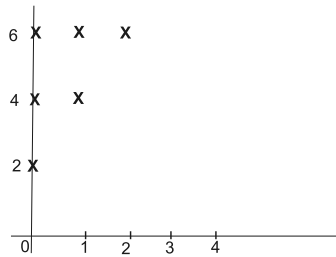
A relation can be represented graphically also and this helps us to understand it better. If R is a relation from A to B , then to draw the graph of R we proceed as follows:

Take two perpendicular lines OX and OY . Represent the elements of A by points on OX and the elements of B by points on OY . Plot the members of R as points in the XOY plane. This is the graph of R .

Example 1.25. Let $A = \{0, 1, 2, 3, 4\}$, $B = \{0, 2, 4, 6\}$.

Define $R = \{(a, b) \mid b > 2a\}$. Then $R = \{(0, 2), (0, 4), (0, 6), (1, 4), (1, 6), (2, 6)\}$.

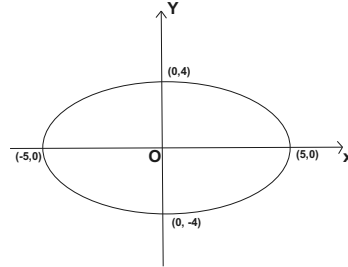
The graph of R is as shown by points marked \times .



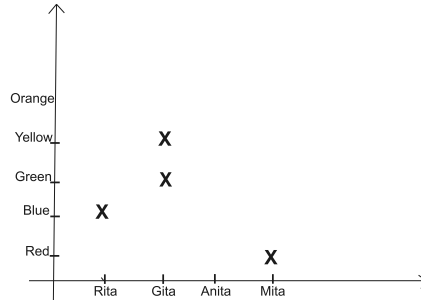
Example 1.26. Let R_1 be a relation defined on \mathbb{R} as follows:

$$R_1 = \{(x, y) \mid 16x^2 + 25y^2 = 400\}$$

The graphical representation of this relation is the set of points in the plane of \mathbb{R}^2 satisfying $16x^2 + 25y^2 = 400$ i.e. $\frac{x^2}{25} + \frac{y^2}{16} = 1$ which is an ellipse centered at origin with major axis of length 10 along x -axis and minor axis of length 8 along y -axis.



Example 1.27. Let $A = \{Rita, Gita, Anita, Mita\}$
 $B = \{Red, Blue, Green, Yellow, Orange\}$
and $R = \{(Rita, Blue), (Gita, Green), (Gita, Orange), (Mita, Red)\}$.
The graph of R is:



Definition 1.17. (Inverse of a relation): Let A and B be two sets and a relation R from A to B . The inverse of R is the set

$$\{(b, a) \in B \times A \mid (a, b) \in R\}$$

and is a relation from B to A , and is denoted by R^{-1} .

Example 1.28. 1. Let $A = \{1, 2, 3, 4, 5\}$, $B = \{a, b, c, d\}$

$$R = \{(1, a), (1, d), (3, b), (2, c)\}$$

$$\text{Then } R^{-1} = \{(a, 1), (d, 1), (b, 3), (c, 2)\}$$

2. Consider the relation from \mathbb{Z} to \mathbb{N} defined by

$$S = \{(a, b) \in \mathbb{Z} \times \mathbb{N} \mid b = a^2\}$$

$$= \{(1, 1), (-1, 1), (2, 4), (-2, 4), (3, 9), (-3, 9), \dots\}$$

$$\text{Then } S^{-1} = \{(b, a) \in \mathbb{N} \times \mathbb{Z} \mid b = a^2\}$$

$$= \{(1, 1), (1, -1), (4, 2), (4, -2), (9, 3), (9, -3), \dots\}$$

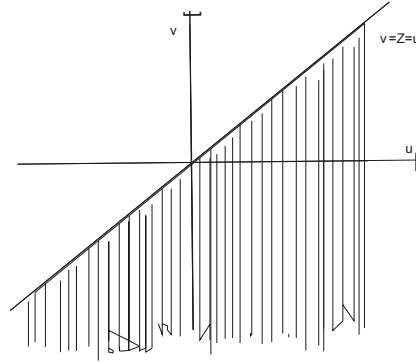
It is important to note that every relation has an inverse. The graph of R^{-1} can be obtained from the graph of R by reflecting it in the line $y = x$.

Properties of Binary Relation on a Set

Some binary relations on a set have certain properties which make them special. We shall study these properties.

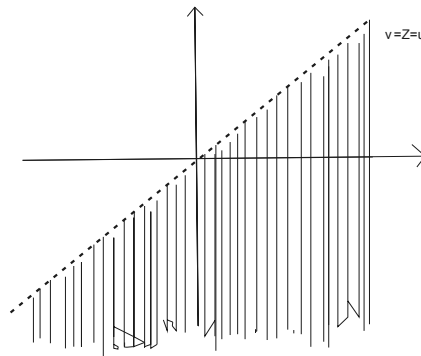
Definition 1.18. A binary relation R on a set A is said to be reflexive if and only if $(a, a) \in R \forall a \in A$.

Example 1.29. 1. $R_1 = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \geq y\}$ is a relation on \mathbb{R} . Since $x \geq x \forall x \in \mathbb{R}$, so that $(x, x) \in R_1 \forall x \in \mathbb{R}$. Hence R_1 is a reflexive relation. The graph of the relation is as shown:



It is the shaded region including the line $y = x$.

2. $R_2 = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x > y\}$. Since $2 \not> 2 \therefore (2, 2) \notin R_2$
Hence R_2 is not reflexive. The graph of the relation is as shown:



It is the shaded region excluding the line $y = x$.

3. $R_3 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 > 0\}$
 $(x, x) \in R_3$ for all $x \neq 0$.
 $(0, 0) \notin R_3$, so that R_3 is not reflexive.
 This shows that $(x, x) \in R_3 \forall x \in \mathbb{R}$ for the relation to be reflexive. If it fails even for one x , the relation is not reflexive.

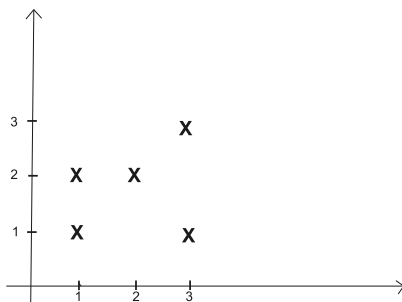
By looking at the graph of the relation can we say that it is reflexive? Before doing this, let us define the diagonal of $A \times A$.

Definition 1.19. Let A be any non-empty set. Then $D = \{(a, a) \mid a \in A\}$ is called the diagonal set of $A \times A$.

Graphically, a relation R on a set A is reflexive if and only if the diagonal of $A \times A$ is contained in R .

A relation R on a set A for which $R = D$ is called the identity relation.

Example 1.30. Let $A = \{1, 2, 3\}$
 $R = \{(1, 1), (2, 2), (3, 3), (1, 2), (3, 2)\}$.
 The graph of R is

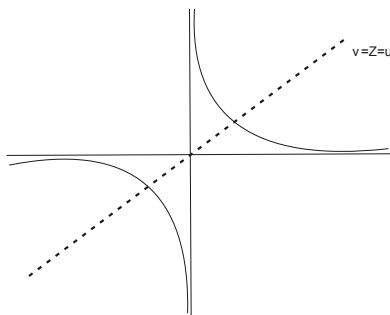


Since the graph of R contains the diagonal, therefore R is reflexive.

Definition 1.20. A binary relation R on a set A is symmetric if and only if $(a, b) \in R$ implies that $(b, a) \in R$.

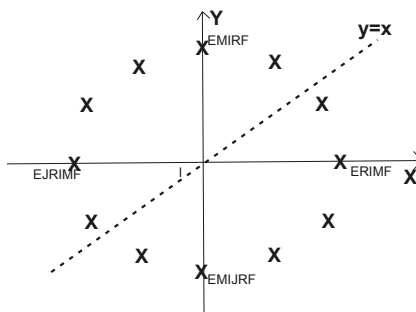
Example 1.31. 1. $R_4 = \{(x, y) \in \mathbb{R}^2 \mid xy = 1\}$ is a relation on \mathbb{R} . If $(x, y) \in \mathbb{R}_4$ then $xy = 1$ so that $yx = 1$. $\therefore (y, x) \in \mathbb{R}_4$. Hence R_4 is symmetric relation.

The graph of this relation is



It is symmetric about the line $y = x$.

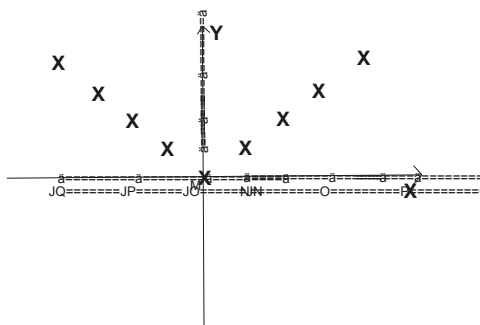
2. $R_5 = \{(x, y) \in \mathbb{Z}^2 \mid x^2 + y^2 = 25\}$ is a relation on \mathbb{Z} .
 If $(x, y) \in \mathbb{R}_5$ then $x^2 + y^2 = 25$ so that $y^2 + x^2 = 25$.
 Hence $(y, x) \in \mathbb{R}_5$. Thus R_5 is symmetric.
 The graph of R_5 is



Observe that the graph is symmetrical about $y = x$.

3. $R_6 = \{(x, y) \in \mathbb{Z}^2 \mid y = |x|\}$

Some points on R_6 are $(1, 1)$, $(-1, 1)$, $(-3, 3)$ etc. Observe that the 2nd component is always positive. $(-1, 1) \in R_6$ but $(1, -1) \notin R_6$. Hence R_6 is not symmetric. The graph of R_6 is



Though the graph looks symmetrical but still R_6 is not symmetric, because symmetry is about y -axis, and not the line $y = x$.

Just by looking at the graph of the relation can we say that the relation is symmetric? Yes, of course. If the graph of the relation is symmetric about the line $y = x$, then if (x, y) belongs to the graph, so will (y, x) . Hence the relation will be symmetric.

A relation R is symmetric if and only if $R = R^{-1}$ i.e. R and R^{-1} have identical graphs.

Definition 1.21. A binary relation R on a set A is transitive if and only if $(a, b), (b, c) \in R$ implies that $(a, c) \in R$.

Example 1.32. 1. Consider the relation R_1 defined in Example 1.29. If $(a, b), (b, c) \in R_1$ then $a \geq b$ and $b \geq c$ so that $a \geq c$. Hence $(a, c) \in R_1$ so that R_1 is a transitive relation.

2. Consider the relation R_6 defined in Example 1.31. If $(x, y), (y, z) \in R_6$ then $y = |x|$ and $z = |y|$. Thus $z = |x|$ so that $(x, z) \in R_6$. Thus R_6 is a transitive relation.

3. Consider the relation R_4 defined in Example 1.31

$(3, \frac{1}{3}), (\frac{1}{3}, 3) \in R_4$, but $(3, 3) \notin R_4$, so that R_4 is not transitive.

Definition 1.22. A binary relation R on a set A is antisymmetric if and only if both $(a, b), (b, a) \in R$ implies that $a = b$.

Remark 1.1. As the name suggests antisymmetric is ‘against symmetric’, so the graph will not be symmetric. It is not just the negation of symmetric.

In fact, no pair $(a, b) \in R$, with $a \neq b$ is such that $(b, a) \in R$. So, not symmetric does not imply antisymmetric and not antisymmetric does not imply symmetric.

Example 1.33. 1. Consider the relation R_4 defined in Example 1.31. $(3, \frac{1}{3}), (\frac{1}{3}, 3) \in R_4$, but $3 \neq \frac{1}{3}$. Hence the relation R_4 is not antisymmetric.

2. Consider the relation R_1 defined in Example 1.29. Let $(x, y), (y, x) \in R_1$. Then $x \geq y$ and $y \geq x$ so that $x = y$. Hence R_1 is an antisymmetric relation.

3. Let A be any set and $\mathcal{P}(A)$ the power set of A . On $\mathcal{P}(A)$ define a relation as follows:

$$R_7 = \{(X, Y) \in \mathcal{P}(A) \times \mathcal{P}(A) \mid X \subseteq Y\}$$

If $(X, Y), (Y, X) \in R_7$ then $X \subseteq Y$ and $Y \subseteq X$, so that $X = Y$. Hence R_7 is antisymmetric.

By looking at the graph of a relation can we conclude whether it is antisymmetric or not? Yes.

The graph of an antisymmetric relation is such that no pair of points other than on the main diagonal are symmetrically located about $y = x$. So it is possible that some points are symmetrically located whereas some points are not. Such a relation is neither symmetric nor antisymmetric. Further the graph consisting of points on the main diagonal only is both symmetric and antisymmetric.

In the following example, we verify the properties of relations.

Example 1.34. Let $A = \{1, 2, 3, 4, 5\}$

$$R_1 = \{(1, 1), (2, 2), (1, 2), (2, 1)\}$$

$$R_2 = \{(4, 5), (5, 4), (1, 2), (2, 3), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}$$

$$R_3 = \{(1, 1), (5, 5), (2, 3), (4, 5)\}$$

$$R_4 = \{(1, 2), (2, 1), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}$$

Then R_1 is not reflexive as $(3, 3) \notin R_1$

R_1 is symmetric as whenever $(a, b) \in R_1 \Rightarrow (b, a) \in R_1$. R_1 is transitive as $(a, b), (b, c) \in R_1 \Rightarrow (a, c) \in R_1$. R_1 is not antisymmetric, as $(1, 2), (2, 1) \in R_1$ but $1 \neq 2$.

Thus R_1 is a symmetric and transitive relation but not reflexive and antisymmetric.

Since $(a, a) \in R_2 \forall a \in A$. $\therefore R_2$ is reflexive. R_2 is not symmetric, as $(1, 2) \in R_2$ $(2, 1) \notin R_2$, R_2 is not transitive, as $(1, 2), (2, 3) \in R_2$ but $(1, 3) \notin R_2$, $(4, 5), (5, 4) \in R_2$ but $4 \neq 5$ so R_2 is not antisymmetric. Thus R_2 is reflexive only.

R_3 is antisymmetric because no two points of the form (x, y) and (y, x) for $x \neq y$ belong to R_3 . It is not reflexive as $(1, 1) \notin R_3$, it is not symmetric as $(2, 3) \in$

R_3 but $(3, 2) \notin R_3$. It is transitive. Thus R_3 is an antisymmetric and transitive relation. It is neither symmetric nor reflexive.

R_4 is reflexive, symmetric and transitive. It is not antisymmetric as $(1, 2), (2, 1) \in R_4$ but $1 \neq 2$.

Equivalence Relation

We have studied four different types of properties of binary relations and each of them is totally independent of the other. But relations satisfying a certain combination of these properties form an important class of relations studied in Mathematics.

Definition 1.23. Let R be a binary relation on a set A . Then R is called a partial order on A if it is reflexive, antisymmetric and transitive and the system (A, R) is called a partially ordered set (Poset).

$(\mathbb{Z}, \geq), (\mathbb{N}, \leq)$ are partially ordered sets.

Definition 1.24. Let R be a binary relation on a set A . Then R is called an equivalence relation if it is reflexive, symmetric and transitive.

Example 1.35. Suppose A is the set of all points on the surface of the earth. On A , define $R = \{(a, b) \in A \times A \mid a \text{ and } b \text{ have the same longitude}\}$

Clearly $(a, a) \in R \forall a \in A$, so that R is reflexive. Let $(a, b) \in R$ then a, b have the same longitude so that $(b, a) \in R$. Hence R is symmetric.

Let $(a, b), (b, c) \in R$. Then a, b have the same longitude and b, c have same longitude. Thus a, c have the same longitude, so $(a, c) \in R$. Thus R is transitive.

Since R is reflexive, symmetric and transitive therefore it is an equivalence relation.

Sometimes we do not talk of a specific relation on A . We denote it by \sim . We write $a \sim b$. The symbol \sim is read as 'wobble'. We read it as 'a is related to b' or 'a wobble b'.

Example 1.36. On \mathbb{Z} define a relation \sim as follows $a \sim b$ if $(a - b)$ is divisible by 4. Clearly for any $a \in \mathbb{Z}$, $a - a = 0$ which is always divisible by 4.

$\therefore a \sim a \forall a \in \mathbb{Z}$ so that R is reflexive.

Let $a \sim b$. Then $(a - b)$ is divisible by 4 so that $(b - a)$ is also divisible by 4.

Hence $a \sim b$ implies that $b \sim a$.

$\therefore \sim$ is symmetric.

Let $a, b, c \in \mathbb{Z}$ such that $a \sim b$ and $b \sim c$, then $(a - b)$ and $(b - c)$ are both divisible by 4. $\therefore (a - b) + (b - c) = (a - c)$ is divisible by 4, hence $a \sim c$, so that \sim is transitive.

Thus \sim is an equivalence relation on \mathbb{Z} .

In the above example we observe the following:

$0 \sim 4k, 1 \sim 1+4k, 2 \sim 4k+2$ and $3 \sim 4k+3$ for any $k \in \mathbb{Z}$. Any $n \in \mathbb{Z}$ is either of the form $4k, 4k+1, 4k+2, 4k+3$ for some $k \in \mathbb{Z}$.

Thus every integer is related to 0 or 1 or 2 or 3. Consider the sets

$$A_0 = \{\dots, -12, -8, -4, 0, 4, 8, 12, \dots\}$$

$$A_1 = \{\dots, -11, -7, -3, 1, 5, 9, 13, \dots\}$$

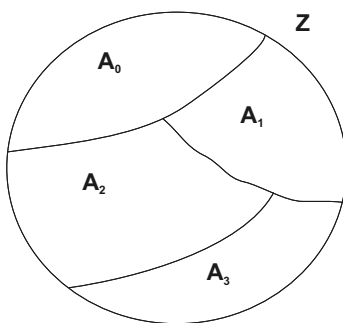
$$A_2 = \{\dots, -10, -6, -2, 2, 6, 10, 14, \dots\}$$

$$A_3 = \{\dots, -9, -5, -1, 3, 7, 11, 15, \dots\}$$

Any two elements of A_0 are related to each other and each of them is related to 0. Similarly each element of A_1 is related to 1, each element of A_2 is related to 2 and each element of A_3 is related to 3. Every integer n belongs to exactly one of the sets A_0, A_1, A_2 or A_3 . That is

$$A_0 \cup A_1 \cup A_2 \cup A_3 = \mathbb{Z}$$

and any two A_i 's are disjoint, $i = 0, 1, 2, 3$.



Thus the set \mathbb{Z} is divided into 4 mutually disjoint subsets such that any two members of the same set are related and any two members from different sets are not related.

This motivates us to see whether this is possible for every equivalence relation.

Definition 1.25. (Equivalence class): Let A be any set and \sim is an equivalence relation on A . For any $a \in A$, the set $\{x \in A : a \sim x\}$, of all elements of A which are related to 'a' is called the equivalence class of 'a'.

It is denoted by \bar{a} or $[a]$ or $cl(a)$. We shall use the notation $[a]$ for the equivalence class of 'a'.

Definition 1.26. (Quotient set): Given an equivalence relation \sim on a set A , the set of all equivalence classes is called the quotient set of $A \text{ mod } \sim$. It may be denoted by A/\sim .

Thus in Example 1.36 the equivalence classes of 0, 1, 2, 3 are A_0, A_1, A_2, A_3 respectively.

Thus $[0] = A_0, [1] = A_1, [2] = A_2, [3] = A_3$.

In fact $[4] = A_0$ so that perhaps we may say that if $x \in [0]$ then $[x] = [0]$.

The quotient set of $\mathbb{Z} \text{ mod } \sim$ is $\{A_0, A_1, A_2, A_3\}$.

We now prove that any two elements which are related give rise to the same equivalence class.

Theorem 1.8. Let \sim be an equivalence relation on a set A . Let $a \in A$. Then for any $b \in A$, $b \sim a$ if and only if $[b] = [a]$.

Proof: $[a] = \{x \in A \mid a \sim x\}$.

Let $b \in A$ such that $b \sim a$. We show that $[a] = [b]$.

Let $x \in [a]$. Then $a \sim x$. Now $b \sim a, a \sim x \Rightarrow b \sim x$ ($\because \sim$ is transitive)

$\Rightarrow x \in [b]$. Hence $[a] \subseteq [b]$(1)

If $x \in [b]$, then $b \sim x$.

Now $b \sim a \Rightarrow a \sim b$ ($\because \sim$ is symmetric).

Since $a \sim b$ and $b \sim x$, therefore $a \sim x$ by transitivity.

$\Rightarrow x \in [a]$. $\therefore [b] \subseteq [a]$...(2)

(1) and (2) $\Rightarrow [a] = [b]$.

Conversely let $[a] = [b]$.

$b \in [b]$ ($\because b \sim b$) $\Rightarrow b \in [a] \Rightarrow a \sim b \Rightarrow b \sim a$ ($\because \sim$ is symmetric). \square

Theorem 1.9. *Let \sim be an equivalence relation on a set A and $a, b \in A$. Then the equivalence classes $[a]$ and $[b]$ are either identical or disjoint.*

Proof: In case $[a] = [b]$, then proof is complete.

Suppose $[a] \neq [b]$, we prove that $[a] \cap [b] = \phi$.

Let, if possible, $x \in [a] \cap [b]$. Then $x \in [a]$ and $x \in [b]$. $\therefore a \sim x$ and $x \sim b$, so that $a \sim b$. Then $[a] = [b]$ by Theorem 1.8, which contradicts our assumption.

Hence $[a] \cap [b] = \phi$. \square

The above theorem is generally stated as “two equivalence classes are either identical or disjoint”. Thus, an equivalence relation gives rise to a partition of the underlying set. Before proving this, we give a formal definition of a partition.

Definition 1.27. *A partition of a set A is a collection of subsets $\{A_\alpha : \alpha \in \Lambda\}$ such that*

- (i) $\bigcup_{\alpha \in \Lambda} A_\alpha = A$
- (ii) $A_\alpha \cap A_\beta = \phi$ for $\alpha \neq \beta, \alpha, \beta \in \Lambda$.

If A is a finite set, then the partition will be finite. If A is an infinite set, the partition may be finite or infinite.

Example 1.37. 1. *Let E = set of even integers*

and O = set of odd integers

Then $\{E, O\}$ is a finite partition of \mathbb{Z} .

2. *Referring to Examples 1.36*

the set $\{A_0, A_1, A_2, A_3\}$ forms a partition of \mathbb{Z} . Thus an infinite set \mathbb{Z} has a finite partition.

3. *On \mathbb{Z} , define $A_0 = \{0\}$, for $n > 0$ $A_n = \{n, -n\}$. Then $\{A_n | n \in \mathbb{N} \cup \{0\}\}$ is an infinite partition of \mathbb{Z} .*

Theorem 1.10. *Let A be any set. Every equivalence relation on A gives rise to a partition of A . Conversely, corresponding to every partition of A , there is defined an equivalence relation on A .*

Proof: Let \sim be an equivalence relation on A . Since each element of A belongs to some equivalence class, namely $[a]$, and any two equivalence classes are either identical or disjoint, therefore $\{[a] : a \in A\}$ forms a partition of A . Conversely let $\{A_\alpha : \alpha \in \Lambda\}$ be a partition of A , where Λ is some index set. Then $A = \bigcup_{\alpha \in \Lambda} A_\alpha$ and $A_\alpha \cap A_\beta = \phi$ for $\alpha \neq \beta, \alpha, \beta \in \Lambda$. We define a relation \sim on A as follows:

For $a, b \in A$, $a \sim b$ if and only if a, b belong to the same set A_α , for some $\alpha \in \Lambda$. Let $a, b, c \in A$. Then

1. Since $A = \bigcup_{\alpha \in \Lambda} A_\alpha$, there exists $\alpha \in \Lambda$ such that $a \in A_\alpha$. Hence $a \sim a$, so that \sim is reflexive.

2. Let $a \sim b$. Then, there exists $\alpha \in \Lambda$ such that $a, b \in A_\alpha$.
 $\therefore b, a \in A_\alpha$, so that $b \sim a$. Hence \sim is symmetric.

3. Let $a \sim b$ and $b \sim c$,
 $a \sim b \Rightarrow \exists \alpha \in \Lambda$ such that $a, b \in A_\alpha$.
 $b \sim c \Rightarrow \exists \beta \in \Lambda$ such that $b, c \in A_\beta$.
 If $\alpha \neq \beta$ then $A_\alpha \cap A_\beta = \phi$ but $b \in A_\alpha \cap A_\beta$, so that $\alpha = \beta$ i.e. $A_\alpha = A_\beta$. Thus $a, c \in A_\alpha$. Hence $a \sim c$. So \sim is transitive.

From 1, 2, 3 it follows that \sim is an equivalence relation on A . □

Graph of an Equivalence Relation

Consider a set A consisting of 10 elements, say

$$A = \{a_1, a_2, \dots, a_{10}\}$$

Suppose R is an equivalence relation on A , whose equivalence classes are

$$C_1 = [a_1] = \{a_1, a_2, a_5\}, C_2 = [a_3] = \{a_3, a_4, a_7, a_9\}, C_3 = [a_6] = \{a_6, a_8\}, C_4 = [a_{10}] = \{a_{10}\}.$$

Then every element of C_i is related to each other, $i = 1, 2, 3, 4$ (by definition of equivalence relation) and no two elements of C_i and C_j are related, for $i \neq j$.

$$\therefore R = (C_1 \times C_1) \cup (C_2 \times C_2) \cup (C_3 \times C_3) \cup (C_4 \times C_4)$$

$$= R_1 \cup R_2 \cup R_3 \cup R_4, \text{ where } R_i = C_i \times C_i, i = 1, 2, 3, 4.$$

Thus R_i is the universal relation on C_i . Also $R_i \cap R_j = \phi, i \neq j, i, j = 1, 2, 3, 4$.

$$o(R_i) = o(C_i \times C_i) = [o(C_i)]^2$$

$$o(R) = \sum_{i=1}^4 o(R_i) = \sum_{i=1}^4 [o(C_i)]^2 = 3^2 + 4^2 + 2^2 + 1^2 = 30.$$

In fact, $R = \{(a_1, a_1), (a_1, a_2), (a_1, a_5), (a_2, a_1), (a_2, a_2), (a_2, a_5), (a_5, a_1), (a_5, a_2), (a_5, a_5), (a_3, a_3), (a_3, a_4), (a_3, a_7), (a_3, a_9), (a_4, a_3), (a_4, a_4), (a_4, a_7), (a_4, a_9), (a_7, a_3), (a_7, a_4), (a_7, a_7), (a_7, a_9), (a_9, a_3), (a_9, a_4), (a_9, a_7), (a_9, a_9), (a_6, a_6), (a_6, a_8), (a_8, a_6), (a_8, a_8), (a_{10}, a_{10})\}$.

The graph of the R is shown in the Figure 1.

If we rearrange the elements of A on the axes so that the elements of an equivalence class occur together then the graph of R appears as in Figure 2. Another way of rearranging the elements is shown in the Figure 3.

Thus the graph of an equivalence relation can be rearranged as square blocks, put diagonally.

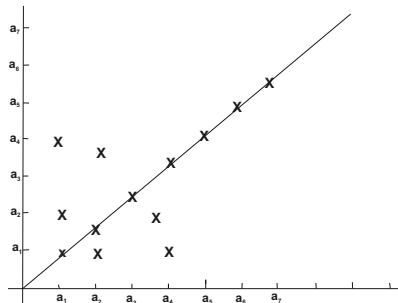


Fig 1

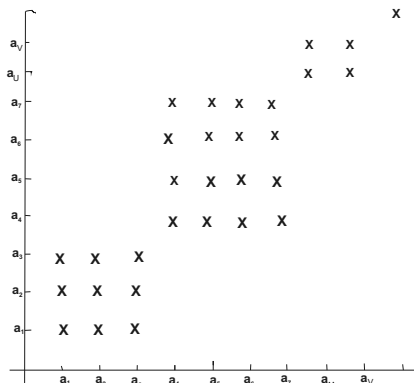


Fig 2

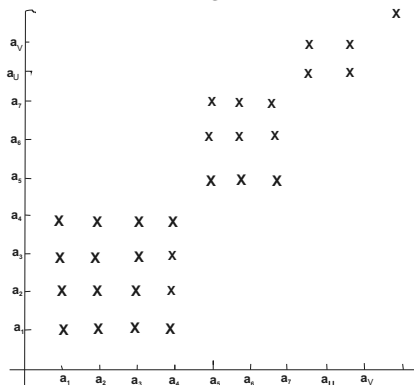


Fig 3

Thus, we can generalize the above result as follows:

Let A be a finite set having n elements and let R be an equivalence relation on A . Let $C_1, C_2, C_3, \dots, C_k$ be the equivalence classes of R and $o(C_i) = n_i, 1 \leq i \leq k$. Then $A = \bigcup_{i=1}^k C_i$.

Since any two elements of C_i are related and no two elements of C_i and C_j are related for $i \neq j$, therefore if R_i is the universal relation, then

$$R = \bigcup_{i=1}^k R_i, R_i \cap R_j = \phi, i \neq j.$$

$$o(R) = \sum_{i=1}^k o(R_i) = \sum_{i=1}^k o(C_i)^2 = n_i^2.$$

If the elements of A are written on the axes in the order of the elements of $C_1, C_2, C_3, \dots, C_k$, then the graph of R can be put as $n_i \times n_i$ square blocks, along the diagonal $i = 1, 2, \dots, k$.

Problem 1.7. On \mathbb{R}^2 , define a binary relation as follows:

$$\mathcal{R} = \{((a, b), (c, d)) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid a^2 + b^2 = c^2 + d^2\}$$

Prove that \mathcal{R} an equivalence relation. Find the equivalence classes of \mathcal{R} .

Solution: Since $a^2 + b^2 = a^2 + b^2, \forall (a, b) \in \mathbb{R}^2$

$\therefore ((a, b), (a, b)) \in \mathcal{R} \quad \forall (a, b) \in \mathbb{R}^2$. Hence the relation \mathcal{R} is reflexive.

Let $((a, b), (c, d)) \in \mathcal{R}$

$$\therefore a^2 + b^2 = c^2 + d^2$$

$$\Rightarrow c^2 + d^2 = a^2 + b^2$$

$$\Rightarrow ((c, d), (a, b)) \in \mathcal{R}$$

$\Rightarrow \mathcal{R}$ is a symmetric relation.

Let $((a, b), (c, d)), ((c, d), (e, f)) \in \mathcal{R}$

$$\therefore a^2 + b^2 = c^2 + d^2 \text{ and } c^2 + d^2 = e^2 + f^2$$

Thus $a^2 + b^2 = e^2 + f^2$

so that $((a, b), (e, f)) \in \mathcal{R}$

Thus, \mathcal{R} is transitive relation.

Hence, \mathcal{R} is an equivalence relation.

Let us now find the equivalence classes of $(a, b) \in \mathbb{R}^2$. $[(a, b)] = \{(x, y) \in \mathbb{R}^2 | (x, y) \sim (a, b)\} = \{(x, y) \in \mathbb{R}^2 | (x^2 + y^2 = a^2 + b^2)\}$.

Thus $[(a, b)]$ is a circle with center at the origin and passing through (a, b) . Hence the equivalence classes are the concentric circles with center at the origin.

Problem 1.8. Let A be a set having 5 elements.

- (i) How many binary relations can be defined on A ?
- (ii) How many reflexive binary relations can be defined on A ?
- (iii) How many symmetric binary relations can be defined on A ?
- (iv) How many equivalence relations can be defined on A ?

Solution: Since $o(A) = 5 = n$ (say)

$$\therefore o(A \times A) = 5^2 = 25.$$

Let $A = \{a, b, c, d, e\}$, Then $D = \{(x, x) | x \in A\}$.

- (i) Since a binary relation on A is precisely a subset of $A \times A$ and $o(\wp(A \times A)) = 2^{25}$

\therefore There are 2^{25} binary relations on A .

- (ii) Since a reflexive relation on R always contains the diagonal D , and there are 5(=n) elements in the diagonal.

\therefore Number of subsets of $A \times A$ which always contains $D = 2^{25-5} = 2^{20} (= 2^{n^2-n})$.

- (iii) Let R be a symmetric binary relation on A . Then $(a, b) \in R \Rightarrow (b, a) \in R$, for $a \neq b$. Also any number of elements of the form (x, x) for $x \in A$ may be in R . Thus we see that the choice of elements of R has to be made from $5 + 4 + 3 + 2 + 1 = 15$ elements.

\therefore Number of subsets of $A \times A$ which always contain (b, a) whenever it contains $(a, b) = 2^{15}$. Number of symmetric relations on $A = 2^{15}$.

- (iv) Since every partition of a set gives rise to an equivalence relation on the set, therefore the equivalence relations on a set with 5 elements is equal to the number of partitions of a set with 5 elements.

Let $A = \{a, b, c, d, e\}$.

Number of partitions of A into subsets of the form $\{a\}, \{b\}, \{c\}, \{d\}, \{e\} =$ number of ways in which 5 sets containing one element each can be chosen = 1.

Number of partitions of A into subsets of the form $\{a\}, \{b\}, \{c\}, \{d, e\} =$ number of ways in which a set containing 2 elements can be chosen $= {}^5C_2 =$

10.

Number of partitions of A into subsets of the form $\{a\}, \{b\}, \{c, d, e\}$ = number of ways in which a set containing 3 elements can be chosen $= {}^5C_3 = 10$.

Number of partitions of A into subsets of the form $\{a\}, \{b, c, d, e\}$ = number of ways in which a set containing 4 elements can be chosen $= {}^5C_4 = 5$.

Number of partitions of A into subsets of the form $\{a\}, \{b, c\}, \{d, e\}$ = number of ways in which a set containing 2 elements and another set containing 2 of the remaining 3 elements can be chosen $= {}^5C_2 \times {}^3C_2 = 10 \times 3 = 30$.

Number of partitions of A into subsets of the form $\{a, b\}, \{c, d, e\}$ = number of ways in which a set containing 2 elements can be chosen $= {}^5C_2 = 10$.

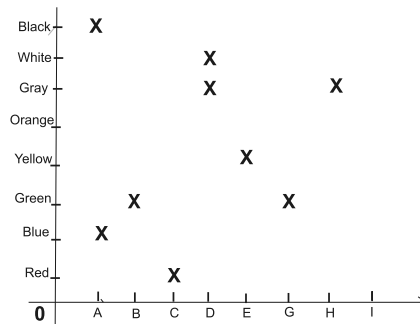
Number of partitions of A into subsets of the form $\{a, b, c, d, e\}$ = number of ways in which a set containing 5 elements can be chosen $= {}^5C_5 = 1$.

Total number of partitions of $A = 1 + 10 + 10 + 5 + 30 + 10 + 1 = 67$.

\therefore Number of equivalence relations on $A =$ number of partitions of $A = 67$.

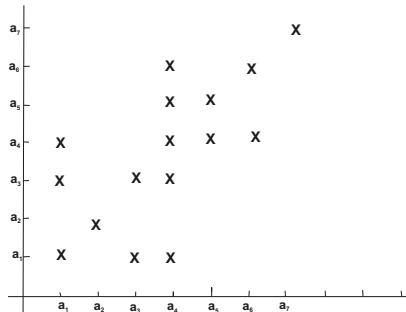
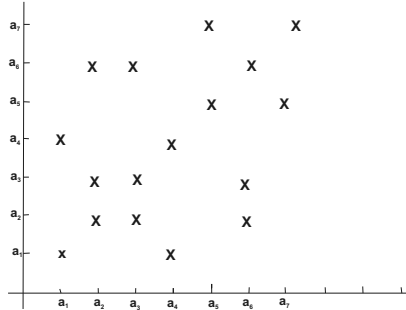
1.6 Exercise

- Let C be the set of all children in Delhi in the age group 3 to 10 years and S the set of all schools in Delhi. Define 3 binary relations from C to S .
- If A and B are sets such that $o(A) = 5$ and $o(B) = 3$, then
 - How many binary relations are there from A to B .
 - How many binary relations are there on A .
 - How many binary relations are there on B .
- Draw the graphs of the following binary relations.
 - $A = \{a, b, c, d, e\}$, $B = \{x, y, z\}$
 $R = \{(a, x), (b, x), (c, x), (d, y), (e, z)\}$
 - $A = \{\text{SVC, LBC, MC, JMC, DR, IP, HR, LSR, DB}\}$
 $B = \{\text{B.Sc., BA, BBE, B.Com, MA, M.Com}\}$
 $R = \{(\text{SVC, B.Sc.}), (\text{SVC, MA}), (\text{LBC, B.Com}), (\text{MC, BA}), (\text{LSR, BBE}), (\text{JMC, BA}), (\text{DR, M.Com})\}$
- Write the relation whose graph is the following:



5. Let \mathcal{R} be a relation on the set of reals defined as follows:
 $\mathcal{R} = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y^2 = 4x\}$
 Draw the graph of this relation.
6. Find the inverses of the relations in Q3 and draw their graphs.
7. If \mathcal{R} is a relation on a set A . Prove that
 (i) \mathcal{R} is reflexive $\Leftrightarrow \mathcal{R}^{-1}$ is reflexive.
 (ii) \mathcal{R} is symmetric $\Leftrightarrow \mathcal{R}^{-1}$ is symmetric.
 (iii) \mathcal{R} is transitive $\Leftrightarrow \mathcal{R}^{-1}$ is transitive.
 (iv) \mathcal{R} is antisymmetric $\Leftrightarrow \mathcal{R}^{-1}$ is antisymmetric.
8. Check whether the following relations are reflexive, symmetric, transitive and antisymmetric.
 (i) S is the set of all students of IIT, Delhi.
 $R_1 = \{(a, b) \in S \times S \mid A \text{ and } B \text{ study a common course}\}$
 (ii) W is the set of all words of the English language.
 $R_2 = \{(x, y) \in W \times W \mid \text{words } x \text{ and } y \text{ have no letter in common}\}$
 (iii) P is the set of all points on the earth.
 $R_3 = \{(p, q) \in P \times P \mid p \text{ and } q \text{ have the same latitude}\}$
 (iv) X is the set of all women in India.
 $R_4 = \{(a, b) \in X \times X \mid a \text{ is mother of } b\}$
 (v) X is the set of all people living in India.
 $R_5 = \{(x, y) \in X \times X \mid x \text{ and } y \text{ have the same mother tongue}\}$
9. On each of the sets defined in Q8, define a relation different from the one already given.
10. Determine which of the following relations are reflexive, symmetric, transitive and antisymmetric.
 (i) L is the set of all lines in a plane
 $R_1 = \{(l_1, l_2) \in L \times L \mid l_1 \text{ is perpendicular to } l_2\}$
 $R_2 = \{(l_1, l_2) \in L \times L \mid l_1 \text{ is parallel to } l_2\}$
 $R_3 = \{(l_1, l_2) \in L \times L \mid l_1 \text{ intersects } l_2 \text{ in one point}\}$
 (ii) On \mathbb{Q} , the set of rationals, define
 $R_4 = \{(a, b) \in \mathbb{Q} \times \mathbb{Q} \mid |a - b| < \frac{1}{2}\}$
 $R_5 = \{(a, b) \in \mathbb{Q} \times \mathbb{Q} \mid |a| = |b|\}$
 $R_6 = \{(\frac{a}{b}, \frac{c}{d}) \in \mathbb{Q} \times \mathbb{Q} \mid ad = bc\}$
 (iii) On \mathbb{N} , the set of natural numbers, define
 $R_7 = \{(a, b) \in \mathbb{N} \times \mathbb{N} \mid a \text{ divides } b\}$
 (iv) On \mathbb{Z} , define
 $R_8 = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid |a| = |b|\}$
11. For the relations defined in Q1, determine which of the properties: reflexivity, symmetry, transitivity and antisymmetry do they possess?
12. Which of the relations in Q8 are equivalence relations. Determine the equivalence classes of the equivalence relations.
13. Determine the set of equivalence classes of the equivalence relations in Q10.

14. Determine a binary relation on $S = \{1, 2, 3, 4, 5\}$ with the help of a graph, which satisfies the following properties:
- Symmetric and antisymmetric, not reflexive.
 - Symmetric and reflexive, not antisymmetric
 - Symmetric, reflexive and antisymmetric.
 - Reflexive and antisymmetric, but not symmetric.
 - Neither reflexive, nor symmetric nor antisymmetric.
15. Draw the graph of the following relation on \mathbb{N} . Also find the number of elements in the relation.
 $A = \{(a, b) \in \mathbb{N} \times \mathbb{N} \mid a \leq b \leq 15\}$
16. Let $A \subseteq \mathbb{N} \times \mathbb{N}$ defined as:
- $(1, 1) \in A$, $(a, b) \in A \Rightarrow (a, b+1)$ and $(a+1, b+1) \in A$
 Draw the graph of A .
 - If $B = \{(a, b) \in \mathbb{N} \times \mathbb{N} \mid a \geq b\}$, find $A \cap B$
17. Construct examples of relations on $S = \{a, b, c, d, e\}$ which satisfy the following properties:
- Reflexive but neither symmetric nor transitive.
 - Symmetric but neither reflexive nor transitive
 - Transitive but neither reflexive nor symmetric.
 - Reflexive and symmetric but not transitive.
 - Reflexive and transitive but not symmetric.
 - Symmetric and transitive but not reflexive.
 - Neither symmetric, nor reflexive, nor transitive.
 - Reflexive, symmetric and transitive.
 - Symmetric but not antisymmetric.
 - Antisymmetric but not symmetric.
 - Antisymmetric and symmetric.
 - Neither antisymmetric nor symmetric.
18. What are the equivalence classes of
- The identity relation on $\{2, 4, 6, 8, 10, 12\}$
 - The universal relation on the set $\{1, 2, 3, \dots, 12\}$?
19. Construct an equivalence relation on the set $\{1, 2, 3, \dots, 10\}$ having exactly
- 3 equivalence classes
 - 5 equivalence classes
 - 11 equivalence classes
 - 10 equivalence classes.
20. Given below are the graphs of a binary relation on a set. By looking at the graph, can you tell whether the relation is an equivalence relation? What are the equivalence classes.



21. Let $A = \{a, b, c, d, e, f\}$. Given the following partition of A , find the equivalence relation corresponding to it.

(i) $\{\{a, b, d, e\}, \{c, f\}\}$

(ii) $\{\{a, d, f\}, \{b, c\}, \{e\}\}$

How many elements will there be in the equivalence relation?

22. Let $A = \{1, 2, 3, 4, 5, 6, 7\}$. Given the equivalence classes of the equivalence relation R on A , find the partition of A where R has the equivalence classes given by

(i) $[1] = \{1, 3, 5, 7\}; [2] = \{2, 4, 6\}$

(ii) $[1] = \{1\}, [2] = \{2, 3, 5, 7\}, [4] = \{4, 6\}$

Also find the equivalence relation. How many elements does the equivalence relation contain?

23. Let $A = \{1, 2, 3, 4, 5\}$. On A can you define an equivalence relation having exactly

(i) 5 elements

(ii) 7 elements

(iii) 8 elements

(iv) 17 elements

(v) 18 elements.

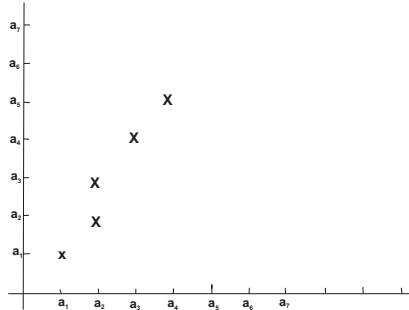
1.7 Supplementary Exercises

1. State whether the following are true or false. Justify the false ones.
 - (i) The null set is a superset of every set.
 - (ii) Every set is a superset of the universal set.
 - (iii) If $A = \{a, b, c\}$ then $c \subseteq A$.
 - (iv) If $S = \{1, 2, 3\}$ then $\{1\} \in S$.
 - (v) $\phi \subseteq \{\phi, \{\phi\}\}$
 - (vi) $5 \in \{3a + 4b \mid a \in \{-1, 0, 1, 2\}, b \in \{0, 1, 2, 3\}\}$
 - (vii) If A has 3 elements, then $\mathcal{P}(\mathcal{P}(A))$ has 27 elements.
 - (viii) If A and B are unequal sets, then $A \cap B \subsetneq A \cup B$.
 - (ix) If A and B are sets then $A \Delta B \neq \phi$.
 - (x) If A and B are non-empty sets then $A \times B = B \times A$.
 - (xi) $A \subseteq B \Rightarrow A^c \subseteq B^c$.
 - (xii) $(A \times B)^c = A^c \times B^c$.
 - (xiii) If a relation R on a set A is symmetric then it is not anti-symmetric.
 - (xiv) Every relation on a set is reflexive.
 - (xv) If a relation R is symmetric, so is R^{-1} .
 - (xvi) Every relation on a set A is either symmetric or anti-symmetric.
 - (xvii) If $A = \phi$, then $\mathcal{P}(A) = A$.
 - (xviii) The number of subsets of A which contain neither 1 nor 5, where $A = \{1, 2, 3, 4, 5, 6\}$, are $2^6 - 2^2$.
 - (xix) The number of equivalence relations on a set with 3 elements is 3.
2. If $A = \{a, b, \phi\}$, $B = \{\phi\}$ list the elements of $A \cup B$, $A \cap B$, $A \setminus B$, $A \Delta B$ and $\mathcal{P}(A)$.
3. If $A = \{a, b, c, \{a, b\}\}$, find
 - (i) $A \setminus \{a, b\}$
 - (ii) $\{a, b, c\} \setminus A$
 - (iii) $(\{a, b, c\} \cup \{A\}) \setminus A$
 - (iv) $A \setminus \{A\}$
 - (v) $\mathcal{P}(A)$
4. A_1, A_2, \dots, A_k are sets such that $A_i \subseteq A_{i+1}$, $i = 1, 2, \dots, k-1$. Find $A_1 \cap A_2 \cap \dots \cap A_k$ and $A_1 \cup A_2 \cup \dots \cup A_k$.
5. For any real number α , let

$$A_\alpha = \{a \in \mathbb{R} \mid a \leq \alpha\}$$
 - (i) write A_α as an interval.
 - (ii) $\bigcup_{\alpha \in \mathbb{R}} A_\alpha$.
 - (iii) $A_\alpha \cap A_\beta^c$.
6. Let $A = \{n \in \mathbb{Z} \mid n \text{ is even}\}$, $B = \{n \in \mathbb{Z} \mid n \text{ is odd}\}$,
 $C = \{n \in \mathbb{Z} \mid n \text{ is a multiple of 3}\}$,
 $D = \{n \in \mathbb{Z} \mid n \text{ is a multiple of 4}\}$.
 Find (i) $A \cap C$, (ii) $A \cup C$, (iii) $C \cup D$, (iv) $B \cap D$, (v) $B \cap C$ (vi) $A \cap B \cap C \cap D$.
7. Prove that $A \subseteq B$ if and only if $(B \cap X) \cup A = B \cap (X \cup A)$ for every set X .

8. Write the power set of the set A where
- $A = \phi$
 - $A = \{\phi\}$
 - $A = \{\phi, \{\phi\}\}$
 - $A = \mathcal{P}(A)$, where A is as in (iii)
 - $A = \mathcal{P}(B)$, $B = \{x, y\}$.
9. Write 3 elements of the set
 $\{(a, b) \in \mathbb{R} \times \mathbb{R} \mid \frac{a}{b} \in \mathbb{Q}, a, b \text{ are distinct irrational numbers}\}$.
10. If A is any set, when can we have
- $o(\mathcal{P}(A)) = o(A)$
 - $\mathcal{P}(A) = A$.
11. A and B are finite sets with $o(A) = m$ and $o(B) = n$.
- How many binary relations can be defined from A to B ?
 - How many binary relations can be defined on A ?
12. How many reflexive binary relation can be defined on a set with n elements?
13. How many symmetric binary relation can be defined on a set with n elements ?
14. Prove or disprove the following, "Every symmetric and antisymmetric binary relation on a set A is reflexive."
15. On the set A , the following relations are defined. Check whether they are equivalence relations or not. If not, give reasons. If yes, find the equivalence classes.
- $A = \mathbb{Z}$. For $a, b \in \mathbb{Z}$, $a \sim b$ if $|a|=|b|$.
 - $A = \mathbb{Z}$. For $a, b \in \mathbb{Z}$, $a \sim b$ if $ab > 0$.
 - $A = \mathbb{R}$. For $a, b \in \mathbb{R}$, $a \sim b$ if $b = 2a + 3$.
 - $A = \mathbb{Q}$. For $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$, $\frac{a}{b} \sim \frac{c}{d}$ if $\frac{a}{b}$ and $\frac{c}{d}$ are equivalent to a rational number with common denominator.
 - $A = \mathbb{Z}$. For $a, b \in \mathbb{Z}$, $a \sim b$ if $2a + 3b = 10$.
16. On \mathbb{R}^2 , the following relations are defined. Check whether they are equivalence relations or not. If not, give reasons. If yes, find the equivalence classes.
- $(a, b) \sim (c, d)$ iff both the points lie on the same curve $4x + 5y = k$, for some $k \in \mathbb{R}$.
 - $(a, b) \sim (c, d)$ iff both the points lie on the same curve $x^2 + y^2 = k^2$, for some $k \in \mathbb{R}$.
 - $(a, b) \sim (c, d)$ iff both the points lie on the same curve $9x^2 + 16y^2 = k^2$, for some $k \in \mathbb{R}$.
 - $(a, b), (c, d) \in \mathbb{R}^2$ such that $b, d > 0$: $(a, b) \sim (c, d)$ iff $\frac{d}{b} = 2^{c-a}$.
 - $(a, b) \sim (c, d)$ iff $ad = bc$.

- 17.
- (i) On \mathbb{R}^* define the relation \sim as follows: $a \sim b$ if $\frac{a}{b}$ is a rational number. Is \sim an equivalence relation? If yes, find the equivalence classes.
 - (ii) On \mathbb{R} define the relation \sim as follows: $a \sim b$ if $a-b \in \mathbb{Z}$. In case \sim is an equivalence relation on \mathbb{R} , find the equivalence classes of '0', $\frac{1}{4}$ and $\sqrt{2}$, a where $0 \leq a \leq 1$.
 - (iii) On \mathbb{R} define the relation \sim as follows: $x \sim y$ if $y = mx + c$. For what value of m and c is the relation symmetric?
 - (iv) On \mathbb{R} define the relation \sim as follows: $x \sim y$ if $y = -x$. Is the relation an equivalence relation? Justify. Is it antisymmetric?
18. If R_1, R_2 are two equivalence relations on a set A , then are the following also equivalence relations on A
- (i) $R_1 \cap R_2$ and (ii) $R_1 \cup R_2$ (iii) R_1^{-1} .
19. How many equivalence relations can be defined on a set with n elements, where
- (i) $n = 3$
 - (ii) $n = 4$
20. Graph the relations
- (i) $A = \{(a, b) \in \mathbb{N} \times \mathbb{N} | b \leq a\}$.
 - (ii) Let $A \subset \mathbb{N} \times \mathbb{N}$ defined by
 - (a) $(1, 1) \in A$
 - (b) $(a, b) \in A \Rightarrow (a+1, b), (a+1, b+1) \in A$.
- Draw the graph of A . If $B = \{(a, b) \in \mathbb{N} \times \mathbb{N} | b < a\}$
 $C = \{(a, b) \in \mathbb{N} \times \mathbb{N} | b \geq a\}$
 Find $A \cap B, A \cap C, B \cap C, A \cap B \cap C$.
21. Let $R = \{(a, a), (b, c), (a, b)\}$ be a relation on the set $\{a, b, c\}$. Add the minimum number of element to R so that R becomes
- (i) reflexive
 - (ii) symmetric
 - (iii) transitive
 - (iv) antisymmetric
 - (v) equivalence relation.
22. Complete the graph of the following relation to define the smallest relation which is:



- (i) reflexive
 - (ii) symmetric
 - (iii) transitive
 - (iv) antisymmetric
 - (v) equivalence relation.
23. Draw the graph of the following relation on \mathbb{N} . For a fixed $n \in \mathbb{N}$, $A_n = \{(a, b) \in \mathbb{N} \times \mathbb{N} \mid a \leq b \leq n\}$. How many elements are there in A_n ?
24. Find the flaw, if any, in the following argument:
- (i) A is any set and R a symmetric and transitive relation defined on A .
 $(a, b) \in R \Rightarrow (b, a) \in R$ since R is symmetric.
 Now $(a, b), (b, a) \in R \Rightarrow (a, a) \in R$, as R is transitive.
 Thus $(a, a) \in R$ so that R is reflexive.
 Thus a symmetric and transitive relation is reflexive.
 - (ii) Let A be any set and R is symmetric and antisymmetric relation on A .
 $(a, b) \in R \Rightarrow (b, a) \in R$ since R is symmetric.
 $(a, b), (b, a) \in R \Rightarrow a=b$ as R is antisymmetric.
 Thus $(a, a) \in R$ so that R is reflexive.
 Thus a symmetric and antisymmetric relation is reflexive.

1.8 Answers to Exercises

Exercise - (5.2)

1.
 - (i) $\{1\}$
 - (ii) $\{5, 10, 15, \dots\}$
 - (iii) ϕ
 - (iv) $\{0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5\}$
 - (v) $\{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8\}$
2.
 - (i) $A = \{n \in \mathbb{N} \mid n \text{ is a multiple of } 3, n < 21\}$
 - (ii) $B = \{x \in \mathbb{N} \mid x \leq 2\}$
 - (iii) $C = \{x^2 + 1 \mid x \in \mathbb{N}\}$
 - (iv) $D = \{x \in \mathbb{Z} \mid |x| \leq 3\}$
 - (v) $E = \{x \in \mathbb{C} \mid x^4 - 1 = 0\}$

3. Other solutions are possible
- (i) rice, mice, dice
 - (ii) $2 + 5\sqrt{7}, -2 + 5\sqrt{7}, 2 - 5\sqrt{7}$
 - (iii) $2, -2, 2\sqrt{2}$
 - (iv) $1, -1, 7$
 - (v) $-2, 0, -1$
 - (vi) $-1, 1, 2$
4. Other answers are also possible
- (i) $\{2, 4, 10\}$
 - (ii) $\{2\}$
 - (iii) $\{4, 6, 8\}$
 - (iv) $\{4\}$
 - (v) $\{8\}$
 - (vi) $\{2, 4, 10\}$
 - (vii) $\{2, 4\}$
5. (i) $\{q, r\}$, (ii) $\{p, q, r\}$, (iii) $\{r, s, t\}$, (iv) $\{p, q\}$, (v) ϕ
- 6.
- (i) $\{\phi\}$
 - (ii) $\{\phi, \{\phi\}\}$
 - (iii) $\{\phi, \{w\}, \{x\}, \{y\}, \{z\}, \{w, x\}, \{x, y\}, \{y, z\}, \{z, w\}, \{w, y\}, \{x, z\}, \{w, x, y\}, \{w, x, z\}, \{x, y, z\}, \{y, z, w\}, \{w, x, y, z\}\}; 2^4 = 16$
7. (i) F , (ii) T null set is a subset of every set,
 (iii) F , $p \in A$ (iv) T (v) F , $A \subseteq A$
 (vi) T (vii) T (viii) T (ix) T (x) F , $\{q, r\} \subseteq A$ (xi) F (xii) T

Exercise (5.4)

1. $\{a, s, i\}, \{m, a, t, h, e, i, c, s, l, g, b, r\}, \{a, l, s, i\}, \{m, t, h, i, c, s, l, g, b, r\},$
 $\{g, e, b, r\}, \{m, a, t, h, i, c, s\}, \{d, f, j, k, o, p, q, u, v, w, x, z\}, \{(a, a),$
 $(a, e), (n, a), (n, e), (l, a), (l, e), (y, a), (y, e), (s, a), (s, e), (i, a), (i, e)\},$
2. $A \cup B = (-8, 5)$, $A \cap B = (-1, 2)$, $A \setminus B = (-8, -1]$, $(A \cup B)^c =] - \infty, -8] \cup [5, \infty)$
3. $X \cap Y = \{1, 2, 3, 4\}$, $X \cup Z = \{-2, 0, 1, 2, \dots, 7\}$,
 $(X \cap Y) \times Z = \{(1, -2), (1, 0), (1, 2), (2, -2), (2, 0), (2, 2), (3, -2), (3, 0),$
 $(3, 2), (4, -2), (4, 0), (4, -2)\},$
 $(X \setminus Y) \setminus Z = \{5, 6, 7\},$
 $X \Delta Z = \{-2, 0, 1, 3, 4, 5, 6, 7\},$
 $Y \Delta Z = \{-6, -5, -4, -3, -1, 1, 3, 4\}$
4. $(A \times B) \cup (B \times A) = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (4, 1), (4, 2),$
 $(4, 3), (6, 1), (6, 2), (6, 3), (1, 4), (1, 6), (2, 4), (2, 6), (3, 1), (3, 2), (3,$
 $4), (3, 6)\}$
 $(A \times B) \cap (B \times A) = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$
 $(A \times B) \setminus (B \times A) = \{(1, 3), (2, 3), (4, 1), (4, 2), (4, 3), (6, 1), (6, 2), (6,$
 $3)\}$
6. In the given relation, replacing A by A^c and B by B^c .

7. (i) $B \subseteq A$, (ii) $A \cap B = \phi$, (iii) $B = \phi$, (iv) $A = B$

8. *Hint:* Suppose $B = (X \cap A) \cup (Y \cap B)$.

$$B = B \cap (A \cup B) = B \cap (X \cup Y) = (B \cap X) \cup (B \cap Y).$$

$$\therefore (X \cap A) \cup (Y \cap B) = (B \cap X) \cup (Y \cap B)$$

But $X \cap A \subset A$, $Y \cap B \subset B$ and $A \cap B = \phi$.

$\therefore X \cap A = X \cap B = \phi$. Now use $X = X \cap (A \cup B)$.

10. Only (i), (ii), (v) and (vi) are false.

11. (i) $\mathbb{N} \not\subseteq P$, (ii) $2 \in \mathbb{E} \cap P$, (iii) $3 \in P \setminus E$, (iv) $\mathbb{N} \subset \mathbb{Z}$, (v) $\mathbb{Z} \setminus \mathbb{N} \neq \phi$, (vi) $P \subset E^c$

13. 0, 1 or 2

14. $\{x \in \mathbb{N} \mid x \text{ is a multiple of 12 or 18}\}$
 $\{x \in \mathbb{N} \mid x \text{ is a multiple of 36}\}$
 $\{x \in \mathbb{N} \mid x \text{ is either a multiple of 12 or 18 but not of 36}\}$
 $\{(x, y) \mid x \text{ is a multiple of 12, } y \text{ is a multiple of 18}\}$

15. $S \subset R \subset P$
 $S \subset T \subset P$

16. (i) 20, (ii) 30

17. 10

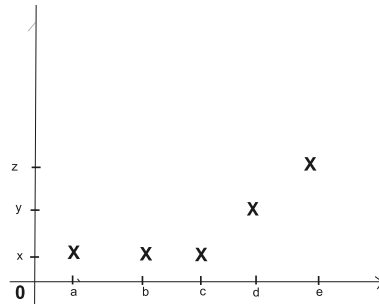
18. (i) 24, 14 (ii) 10, 0

19.
 (i) $A \setminus (B \cup C)$
 (ii) $(A \cap C) \setminus B$
 (iii) $A \cap B \cap C$
 (iv) $C \setminus A$
 (v) $(A \cup B) \setminus C$

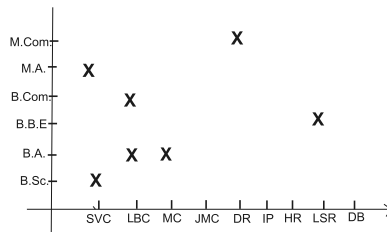
20.
 (i) women who are neither lawyers nor cricket lovers
 (ii) women who love cricket but are not lawyers
 (iii) women who are lawyers and love cricket
 (iv) all men who love cricket
 (v) women or lawyers who do not love cricket

Exercise (1.6)

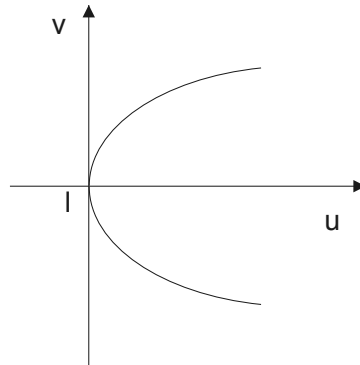
1.
 - (i) $a \sim b$ if a studies in school b .
 - (ii) $a \sim b$ if the distance of school b from the residence of a is less than 5 km.
 - (iii) $a \sim b$ if the school bus of school b comes within 1 km. of the residence of child a .
2. (i) 2^{15} , (ii) 2^{25} , (iii) 2^9
3. (i)



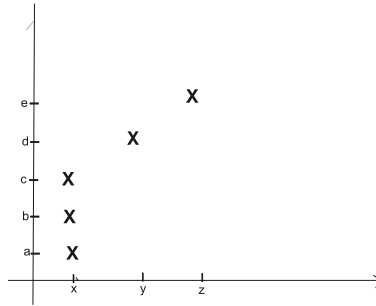
(ii)



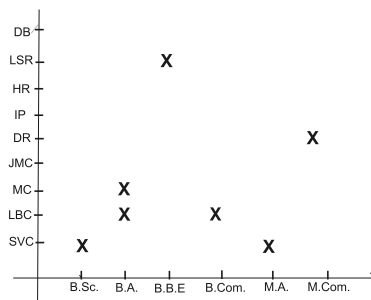
4. $\{(A, \text{Blue}), (C, \text{Red}), (A, \text{Black}), (B, \text{Green}), (D, \text{Grey}), (D, \text{White}), (E, \text{Yellow}), (G, \text{Green}), (H, \text{Grey}), \}$
- 5.



6. (i)



(ii)



8.

- (i) RS
- (ii) S
- (iii) RST
- (iv) A
- (v) RST

10.

- (i) $R_1 : S, R_2 : RST, R_3 : S$
- (ii) $R_4 : RS, R_5 : RST, R_6 : RST$
- (iii) $R_7 : RTA,$
- (iv) $R_8 : RST$

12. R_3 and R_5 .13. R_2, R_5, R_6 and R_8 are equivalence relations.

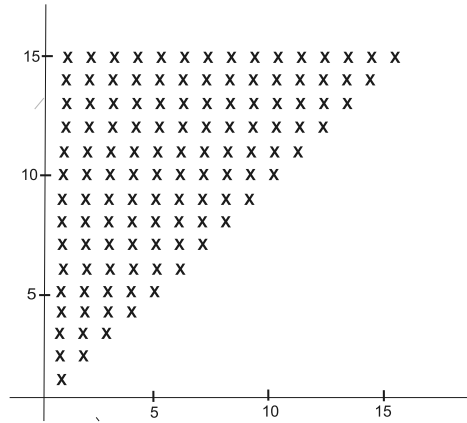
For R_2 : If l_θ is a line through O making an angle θ with the X -axis, then $\{[l_\theta] : 0 \leq \theta \leq 2\pi\}$.

$R_5 : \{[a] | a \in \mathbb{Q}^+ \cup \{0\}\}$.

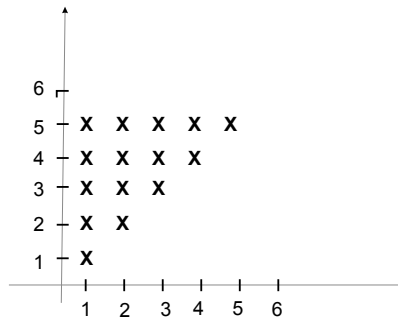
$R_6 : \{[\frac{a}{b}] | a, b \in \mathbb{Z}, b \neq 0, (a, b) = 1\} \cup \{[0]\}$.

$R_8 : \{[a] | a \in \mathbb{Z}^+ \cup \{0\}\}$.

15. 120 elements, graph is:



16.



(ii) $A \cap B = \{(a, a) | a \in N\}$

18.

- (i) $\{2\}, \{4\}, \{6\}, \{8\}, \{10\}, \{12\}$
- (ii) $\{1, 2, 3, \dots, 12\}$

19. Other answers are possible

- (i) $a \sim b$ if 3 divides $(a - b)$.
- (ii) The relation with the equivalence classes $\{c_1, c_2, c_3, c_4, c_5\}$ where $c_1 = \{1, 2, 3\}, c_2 = \{4\}, c_3 = \{5, 6, 7\}, c_4 = \{8\}, c_5 = \{9, 10\}$.
- (iii) not possible.
- (iv) identity relation.

20.

- (i) Yes. Equivalence classes are $\{[a_1], [a_2], [a_5]\}$ where $[a_1] = \{a_1, a_4\}, [a_2] = \{a_2, a_3, a_6\}, [a_5] = \{a_5, a_7\}$.
- (ii) Not an equivalence relation.

- 21.
- (i) $\{(a, a), (a, b), (a, d), (a, e), (b, a), (b, b), (b, d), (b, e), (d, a), (d, b), (d, d), (d, e), (e, a), (e, b), (e, d), (e, e), (c, c), (c, f), (f, c), (f, f)\}$, 20.
 - (ii) $\{(a, a), (a, d), (a, f), (d, a), (d, d), (d, f), (f, a), (f, d), (f, f), (b, b), (b, c), (c, b), (d, a), (c, c), (e, e)\}$, 14.
- 22.
- (i) Partition is $\{C_1, C_2\}$ where $C_1 = \{1, 3, 5, 7\}$, $C_2 = \{2, 4, 6\}$, 25.
 - (ii) $\{C_1, C_2, C_3\}$ where $C_1 = \{1\}$, $C_2 = \{2, 3, 5, 7\}$, $C_3 = \{4, 6\}$, 21.
23. If an equivalence relation has n elements then $n = m_1^2 + m_2^2 + \dots + m_k^2$, where m_i is the order of the i -th equivalence class.
- (i) Yes, $5 = 1^2 + 1^2 + 1^2 + 1^2 + 1^2$
 - (ii) Yes, $7 = 1^2 + 1^2 + 1^2 + 2^2$
 - (iii) No
 - (iv) Yes, $17 = 4^2 + 1^2$
 - (v) No

Supplementary Exercise

- 1.
- (i) False
 - (ii) False
 - (iii) False, $\{c\} \subseteq A$
 - (iv) False, $1 \in S$
 - (v) True
 - (vi) True
 - (vii) False
 - (viii) True
 - (ix) False, $A \Delta B \neq \phi$ when $A \neq B$.
 - (x) False, $A \times B \neq B \times A$ when $A \neq B$.
 - (xi) False, $B^c \subseteq A^c$
 - (xii) False, $(A \times B)^c \supset A^c \times B^c$.
 - (xiii) False, it can be both
 - (xiv) False
 - (xv) True
 - (xvi) False
 - (xvii) False
 - (xviii) False, 2^4 .
 - (xix) False, it is 5
2. $A \cup B = \{a, b, \phi\}$, $A \cap B = \{\phi\}$, $A \setminus B = \{a, b\}$, $A \Delta B = \{a, b\}$
 $\mathcal{P}(A) = \{\phi, \{a\}, \{b\}, \{\phi\}, \{a, b\}, \{a, \phi\}, \{b, \phi\}, \{a, b, \phi\}\}$
- 3.
- (i) $\{c, \{a, b\}\}$
 - (ii) ϕ
 - (iii) $\{A\}$
 - (iv) A
 - (v) $\{\phi, \{a\}, \{b\}, \{c\}, \{\{a, b\}\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, \{a, b\}\}, \{b, \{a, b\}\}, \{c, \{a, b\}\}, \{a, b, c\}, \{a, b, \{a, b\}\}, \{a, c, \{a, b\}\}, \{b, c, \{a, b\}\}, A\}$

4. A_1, A_k

5.

- (i) $A_\alpha = (-\infty, \alpha]$
- (ii) \mathbb{R}
- (iii) if $\beta < \alpha$, $(\beta, \alpha]$
if $\beta \geq \alpha$, ϕ

6.

- (i) $\{n \in \mathbb{Z} \mid n \text{ is a multiple of } 6\}$
- (ii) $\{n \in \mathbb{Z} \mid n \text{ is a multiple of either } 2 \text{ or } 3 \text{ or both}\}$
- (iii) $\{n \in \mathbb{Z} \mid n \text{ is a multiple of either } 3 \text{ or } 4 \text{ or both}\}$
- (iv) ϕ
- (v) $\{n \in \mathbb{Z} \mid n \text{ is an odd multiple of } 3\}$
- (vi) ϕ .

11. (i) 2^{mn} (ii) 2^{m^2} 12. 2^{n^2-n} 13. $2^{n \frac{(n+1)}{2}}$ 14. F, If $A = \{a, b, c\}$, $R = \{(a, a), (b, b)\}$.

15.

- (i) Yes, $\{\{a, -a\} \mid a \in \mathbb{Z}\}$
- (ii) No, Not reflexive
- (iii) No
- (iv) Yes, $\{\frac{[p]}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}, (p, q) = 1\}$
- (v) No

16.

- (i) Yes, $\{\text{lines with equations } 4x + 5y = k \mid k \in \mathbb{R}\}$
- (ii) Yes, concentric circles with centre at the origin.
- (iii) Yes, $\{\text{ellipse with equation } 9x^2 + 16y^2 = k^2 \mid k \in \mathbb{R}\}$
- (iv) Yes, $\{\text{curves with equation } y = k2^x \mid k \in \mathbb{R}^+\}$
- (v) Yes, $\{\text{all lines through the origin, punctured at the origin}\} \cup \{(0, 0)\}$

17.

- (i) Yes, $\mathbb{Q}^* \{[r] \mid r \text{ is irrational}\}$ where $[r] = \{kr \mid k \in \mathbb{Q}^*\}$.
- (ii) \mathbb{Z} , $\mathbb{Z} + \frac{1}{4}$, $\mathbb{Z} + \sqrt{2}$, $\mathbb{Z} + a$.
- (iii) $m = -1$, c can be take any value
 $m = 1$, $c = 0$
- (iv) Only symmetric.

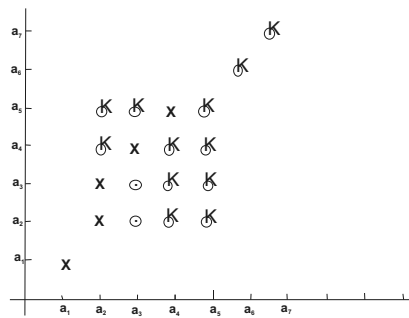
18.

- (i) Yes
- (ii) No, $R_1 : a \equiv b \pmod{4}$; $R_2 : a \equiv b \pmod{3}$.
Then $(1, 5), (5, 8) \in R_1 \cup R_2$, but $(1, 8) \notin R_1 \cup R_2$.
- (iii) Yes

19.

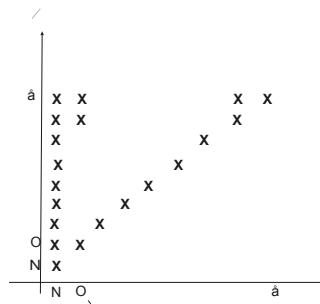
- (i) 5
- (ii) 18

20. (i) diagonal and below the diagonal
 (ii) diagonal and below the diagonal.
 $B, \{(a, a) | a \in \mathbb{N}\}, \phi, \phi.$
21. (i) $\{(b, b), (c, c)\}$
 (ii) $\{(c, b), (b, a)\}$
 (iii) $\{(a, c)\}$
 (iv) ϕ
 (v) $\{(b, b), (c, c), (c, b), (b, a), (a, c), (c, a)\}$
22. (v)



23. $\frac{n(n+1)}{2}$

NM



NM NR

24. (i) $(a, a) \in \mathbb{R}$ only when there exists some $b \in \mathbb{R}$ such that $(a, b) \in \mathbb{R}$. Such an (a,b) may not exist.
 (ii) $(a, b) \in \mathbb{R}$ may not exist.

Chapter 2

Binary Operations

We shall now extend the concept of addition and multiplication of numbers to binary operations on other sets, like set of matrices, polynomials, functions, etc. Properties of these binary operations will be studied. Finally, to illustrate this, we shall discuss the symmetries of regular plane figures, for example, the symmetries of an equilateral triangle, square rectangle etc...

2.1 Definition and Examples

The idea of binary operation may be illustrated by the usual operation of addition in \mathbb{Z} . For every ordered pair of integers (m, n) , there is associated an unique integer $m + n$. We may therefore think of addition as a mapping from $\mathbb{Z} \times \mathbb{Z}$ into \mathbb{Z} , where the image of $(m, n) \in \mathbb{Z} \times \mathbb{Z}$ is denoted by $m + n$. Generalizing this concept we have the following definition:

Definition 2.1. *Let S be a non empty set. Any mapping $\circ : S \times S \rightarrow S$ is called a binary operation on S . The image of $(a, b) \in S \times S$ under the operation \circ , is denoted by $a \circ b$.*

Various symbols used for binary operations are $+$, \times , \circ , \star , $$, \odot , juxtaposition etc.*

Remark 2.1. *The adjective binary is used because our rule combines two elements at a time.*

Example 2.1. *The following are binary operations:*

1. \circ defined by $m \circ n = m$ on \mathbb{N}
2. \star defined by $m \star n = m + n + 1$ on \mathbb{N}
3. $*$ defined by $m * n = |m - n|$ on \mathbb{Z}

4. Let S be a non empty set and $\mathcal{P}(S)$ the power set of S . Define

- (i) $*$ defined by $A * B = A \cup B$ on $\mathcal{P}(S)$,
(ii) \circ defined by $A \circ B = A \cap B$ on $\mathcal{P}(S)$.

The following are not binary operations:

- (i) \circ defined by $m \circ n = m - n$ on \mathbb{N}
 $\because 1, 2 \in \mathbb{N}$ but $1 \circ 2 = 1 - 2 = -1 \notin \mathbb{N}$
(ii) $*$ defined by $m * n = m \div n$ on \mathbb{N}
 $\because 1 * 2 = 1 \div 2 = \frac{1}{2} \notin \mathbb{N}$
(iii) \odot defined by $a \odot b = a \div b$ on \mathbb{Q}
 $1 \odot 0 = 1 \div 0$ is not defined.

Example 2.2. : Let \mathbb{Q}^* be the set of non-zero rational numbers. Check whether division (denoted by \div) is a binary operation in \mathbb{Q}^* . Verify whether the following statements hold for all $a, b, c \in \mathbb{Q}^*$.

- (i) $a \div b = b \div a$
(ii) $a \div (b \div c) = (a \div b) \div c$

Solution: If $a, b \in \mathbb{Q}^*$ then $a = \frac{m}{n}, b = \frac{p}{q}$, where m, n, p, q are non-zero integers. Then $a \div b = \frac{m}{n} \div \frac{p}{q} = \frac{mq}{np} \in \mathbb{Q}^*$ because mq, np are non-zero integers. Hence \div is a mapping from $\mathbb{Q}^* \times \mathbb{Q}^*$ into \mathbb{Q}^* .

- (i) This statement is false. This is because if we take $a = 1, b = 2$ then $a, b \in \mathbb{Q}^*$ but $a \div b = 1 \div 2 = \frac{1}{2}, b \div a = \frac{2}{1} = 2$ since $\frac{1}{2} \neq 2$, therefore $a \div b \neq b \div a$.
(ii) This statement is false. Take $a = 1, b = 2, c = 3$ then $a \div (b \div c) = 1 \div (2 \div 3) = 1 \div \frac{2}{3} = \frac{3}{2}, (a \div b) \div c = (1 \div 2) \div 3 = \frac{1}{2} \div 3 = \frac{1}{6}$, therefore $a \div (b \div c) \neq (a \div b) \div c$.

Example 2.3. Let \mathbb{R} denote the set of real numbers and $*$ a binary operation on \mathbb{R} defined by $a * b = a + b + ab$. Verify that for all $a, b, c \in \mathbb{R}$

- (i) $a * b = b * a$
(ii) $a * (b * c) = (a * b) * c$

Solution:

- (i) $a * b = a + b + ab = b + a + ba = b * a$ using the commutative property of addition and multiplication in \mathbb{R} .
(ii) $a * (b * c) = a * (b + c + bc) = a + (b + c + bc) + a(b + c + bc) = a + b + c + bc + ab + ac + abc$,
 $(a * b) * c = (a + b + ab) * c = a + b + ab + c + (a + b + ab)c = a + b + c + ab + ac + bc + abc$
Hence $a * (b * c) = (a * b) * c$.

The multiplication table (Cayley table)

If S is a finite set consisting of n elements, then a binary operation \star on S can be described by means of a table consisting of n rows and n columns. The rows and columns are headed by the elements of S . The entry at the intersection of a row headed by an element $x \in S$ and column headed by an element $y \in S$

is $x \star y$. Such a table is called a binary operation table, multiplication table, composition table or Cayley table.

We may define a binary operation by giving the multiplication table or else, having defined the binary operation by some rule, we may write the multiplication table for it. This is illustrated in the following examples.

Example 2.4. 1. Let $S = \{x, y, z\}$. Let \star be a binary operation on S defined by the multiplication table

\star	x	y	z
x	x	y	x
y	y	y	z
z	x	z	z

Reading the table, we find,

$$x \star x = x, x \star y = y, x \star z = x, y \star x = y, y \star y = y, y \star z = z, z \star x = x, z \star y = z, z \star z = z.$$

2. Let $S = \{1, -1, i, -i\}$, where $i = \sqrt{-1}$, with the usual multiplication as the binary operation.

The composition table is given below

\cdot	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

3. Let $S = \{1, 2, 3, 4, 5\}$ and \circ a binary operation on S defined by $a \circ b = \gcd(a, b)$. The multiplication table is given below

\circ	1	2	3	4	5
1	1	1	1	1	1
2	1	2	1	2	1
3	1	1	3	1	1
4	1	2	1	4	1
5	1	1	1	1	5

Properties of binary operations

A non-empty set equipped with one or more binary operations is called an algebraic structure.

The algebraic structure consisting of a set S and binary operations \star, \circ on S is denoted by (S, \star, \circ) . $(\mathbb{N}, +)$, $(\mathbb{Z}, +, \cdot)$, (\mathbb{Q}, \cdot) are algebraic structures. According to the properties of binary operations, the algebraic structures are grouped into different classes. We shall now discuss different types of binary operations. In algebra, we come across various mathematical systems which give rise to such type of binary operations.

Definition 2.2. (Associative operation):

Let \star be a binary operation on a set S . Then \star is said to be associative if and only if $(a \star b) \star c = a \star (b \star c) \quad \forall a, b, c \in S$.

Note that \star will not be associative if there exists even one triad x, y, z of elements of S such that $(x \star y) \star z \neq x \star (y \star z)$. In case \star is associative on S , we say that the algebraic structure (S, \star) is associative or S is associative with respect to \star .

Example 2.5. *The addition and multiplication on any set of numbers are associative. Thus $(\mathbb{N}, +)$, (\mathbb{N}, \cdot) , $(\mathbb{Z}, +)$, (\mathbb{Z}, \cdot) , $(\mathbb{Q}, +)$, (\mathbb{R}, \cdot) , (\mathbb{R}^*, \cdot) are all associative algebraic structures.*

$(\mathcal{P}(S), \cup)$, $(\mathcal{P}(S), \cap)$ are associative algebraic structures.

(\mathbb{Q}^, \div) , $(\mathbb{Z}, -)$ are non-associative algebraic structures.*

$(\mathbb{R}^, *)$ in example 2.3 is associative.*

In example 2.4(1), (S, \star) is not associative as

$$(x \star y) \star z \neq x \star (y \star z).$$

Operation with Identity Element

Let S be a non-empty set and \star a binary operation on S . If there exists some element $e \in S$ such that $x \star e = e \star x = x$, $\forall x \in S$, then e is said to be an identity element (neutral element) with respect to \star and (S, \star) is called algebraic structure with identity element.

Example 2.6. 1. $(\mathbb{Z}, +)$ is an algebraic structure with identity element 0.

2. (\mathbb{N}, \cdot) is an algebraic structure with identity element 1.

3. $(\mathcal{P}(S), \cup)$ is an algebraic structure with identity element, the null set.

4. $(\mathbb{N}, +)$ is an algebraic structure without identity element as $0 \notin \mathbb{N}$.

Theorem 2.1. *(Uniqueness of identity element)*

Let (S, \star) be an algebraic structure. If an identity element exists, it is unique.

Proof: Let, if possible there be two identity elements e_1 and e_2 . Then $x \star e_1 = e_1 \star x = x \quad \forall x \in S \dots$ (i)

$$x \star e_2 = e_2 \star x = x \quad \forall x \in S \dots$$
 (ii)

In (i) taking $x=e_2$, we get $e_1 \star e_2 = e_2$

In (ii) taking $x = e_1$, we get $e_1 \star e_2 = e_1$

Hence $e_1 = e_2$. □

Definition 2.3. (Invertible Elements):

Let (S, \star) be an algebraic structure with identity element e . An element $x \in S$ is said to be invertible with respect to \star if there exists some $y \in S$ such that $x \star y = y \star x = e$, and y is called an inverse of x .

Note that the inverse of identity element is itself. Also if y is an inverse of x then x is an inverse of y .

Example 2.7. 1. In (\mathbb{N}, \cdot) , 1 is the identity element and no element other than 1 has an inverse.

2. In $(\mathbb{Z}, +)$, every element has an inverse. In fact if $x \in \mathbb{Z}$, its inverse is $-x$.
3. In (\mathbb{Z}, \cdot) , only 1 and -1 have inverses.
4. In $(\mathcal{P}(S), \cap)$, the identity element is S , and S is the only invertible element, its inverse being S .
5. In $(\mathbb{R}, +)$, every element has an inverse.
6. In (\mathbb{R}, \cdot) , every non-zero element has an inverse.

Theorem 2.2. In an associative algebraic structure with identity element, the inverse of an element, if it exists, is unique.

Proof: Let (S, \star) be an associative algebraic structure with identity element e . Let, if possible an element $x \in S$ have two inverses y and z . Then,
 $x \star y = y \star x = e \dots$ (i)

$$x \star z = z \star x = e \dots \text{ (ii)}$$

Since \star is associative, therefore

$y \star (x \star z) = (y \star x) \star z$, so that, by using (i) and (ii), we get $y \star e = e \star z$, that is $y = z$. \square

Example 2.8. Let $S = \{1, 2, 3, 4\}$. Define a binary operation \star on S by the table

\star	1	2	3	4
1	1	2	3	4
2	2	1	1	1
3	3	1	1	4
4	4	2	3	4

Is \star associative? Does it have an identity element? If it does, find which elements are invertible.

Solution: \star is not associative, because $2 \star (3 \star 4) = 1$, $(2 \star 3) \star 4 = 4$, so that $2 \star (3 \star 4) \neq (2 \star 3) \star 4$. Clearly 1 is the identity element. From the table, 2 has two inverses, namely 2 and 3. Inverse of 1 is 1. Also $2 \star 4 = 1$ but $4 \star 2 = 2$. Hence 4 does not have an inverse. Thus 1, 2 and 3 are the invertible elements.

Remark 2.2. The above example shows that some elements may be invertible whereas others may not be. Moreover, if the binary operation is not associative, inverse of an element may not be unique.

Definition 2.4. (Commutative Operation):

Let S be a non-empty set and \star a binary operation on S . Then \star is said to be commutative if and only if $x \star y = y \star x \quad \forall x, y \in S$.

In case \star is commutative on S , we say that the algebraic structure (S, \star) is commutative or S is commutative with respect to \star .

Example 2.9. $(\mathbb{N}, +)$, (\mathbb{Z}, \cdot) , $(\mathbb{Q}, +)$, (\mathbb{R}, \cdot) , $(\mathcal{P}(S), \cup)$, $(\mathcal{P}(S), \cap)$ are all commutative algebraic structures.

$(\mathbb{Z}, -)$, (\mathbb{Q}^*, \div) , where \mathbb{Q}^* is the set of non-zero rational numbers, are not commutative algebraic structures.

Note that if the multiplication table is symmetric about the main diagonal then the binary operation is commutative and vice versa.

2.2 Exercise

1. Verify whether the following definitions of \star is a binary operations on the given set.

(i) \star defined by $a \star b = 2a + 3b$ on \mathbb{N} .

(ii) \star defined by $a \star b = a - b$ on \mathbb{N} .

(iii) \star defined by $a \star b = |a - b|$ on \mathbb{N} .

(iv) \star defined by $a \star b = a - b$ on \mathbb{Z} .

(v) \star defined by $a \star b = \sqrt{|a - b|}$ on \mathbb{Z} .

- (i) Let $S = \{1, 2, 3\}$. Write the multiplication table for the following binary operations on S.

(ii) \star defined by: $(1, 1) \rightarrow 2, (1, 2) \rightarrow 3, (1, 3) \rightarrow 1, (2, 1) \rightarrow 1, (2, 2) \rightarrow 2, (2, 3) \rightarrow 2, (3, 1) \rightarrow 1, (3, 2) \rightarrow 2, (3, 3) \rightarrow 3$

(iii) \circ defined by: $(a, b) \rightarrow 1 \quad \forall a, b \in S$

(a) $(a, b) \rightarrow \min(a, b) \quad \forall a, b \in S$.

2. Does the following table define a binary operation \star on (i) $S = \{1, 2, 3\}$ (ii) $P = \{1, 2, 3, 4\}$?

Justify your answer.

\star	1	2	3
1	1	3	4
2	2	1	3
3	1	3	2

3. Verify whether the following operations on S are commutative and associative:

(i) $S = \{1, 2\}$, \circ is defined by $1 \circ 1 = 2, 1 \circ 2 = 2, 2 \circ 1 = 2, 2 \circ 2 = 1$

(ii) $S = \mathbb{Z}$, \circ is defined by $a \circ b = a + b - ab$

(iii) $S = \mathbb{Z}$, \circ is defined by $a \circ b = 2a + 3b$

(iv) $S = \mathbb{R}$, \circ is defined by $a \circ b = a$

(v) $S = \mathbb{R}^*$, \circ is defined by $a \circ b = \frac{a}{b}$, where \mathbb{R}^* denotes the set of non-zero real numbers.

(vi) $S = \mathbb{Z}$, \circ is defined by $a \circ b = a - b - ab$.

4. How many different binary operations can be defined on a set S, if S has (i) 2 elements (ii) 4 elements (iii) 8 elements (iv) n elements?

5. Give examples of the following types of binary operations:

(i) commutative but not associative

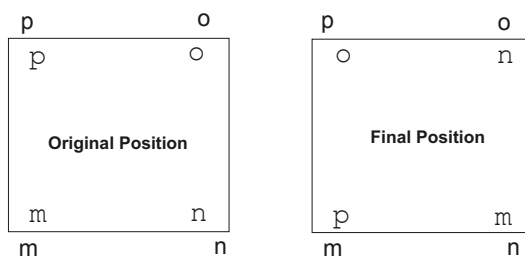
(ii) associative but not commutative

(iii) neither commutative nor associative.

2.3 Introduction to Groups

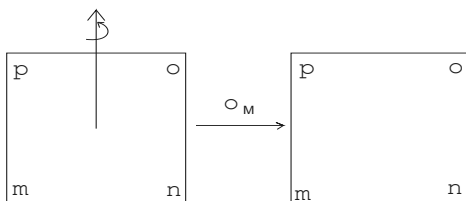
Suppose that a square is removed from a piece of cardboard and fitted back in the original space after moving it. Though it occupies the same space but the position may be different in the sense that the vertices may occupy different positions. Let us consider all the different possible movements of the square. We would like to describe the relationship between the starting position and the final position in the terms of motions.

Cut out a square from a piece of cardboard and name the vertices as P, Q, R, S . Also mark the corners of the board from where it has been cut as P, Q, R, S .

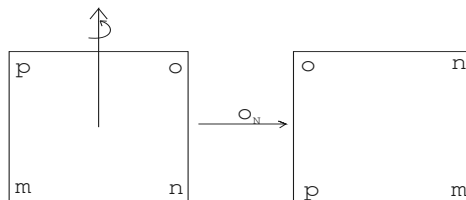


The final position of the square can be obtained from the original position by the rotation of the square about the axis through the centre, perpendicular to the plane, through an angle of 90° anticlockwise. Let the plane of the square be horizontal. Consider the following possible motions

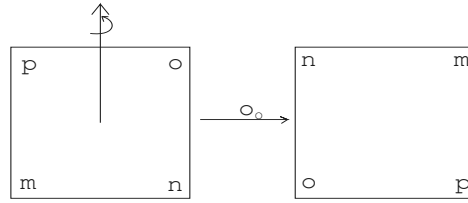
1. R_0 = Rotation of 0° about vertical axis in the plane of the square (no change in position)



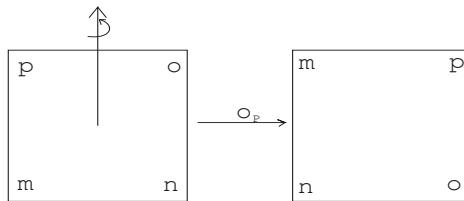
2. R_1 = Rotation of 1 right angle anticlockwise about the vertical axis perpendicular to the plane of the square



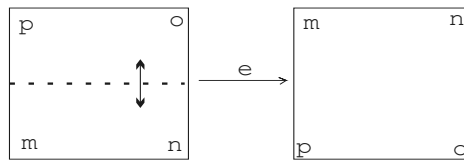
3. R_2 = Rotation of 2 right angles anticlockwise, about the vertical axis perpendicular to the plane of the square.



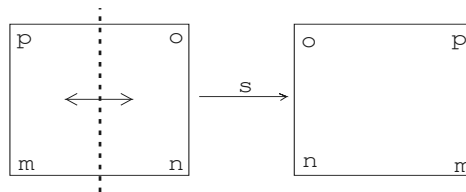
4. R_3 = Rotation of 3 right angles anticlockwise, about the vertical axis perpendicular to the plane of the square.



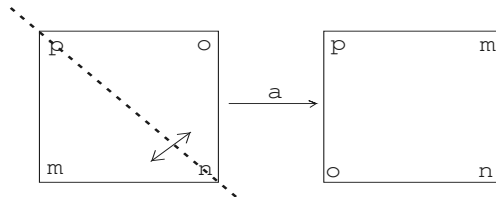
5. H = Rotation of 180° anticlockwise about horizontal axis in the plane of the square.



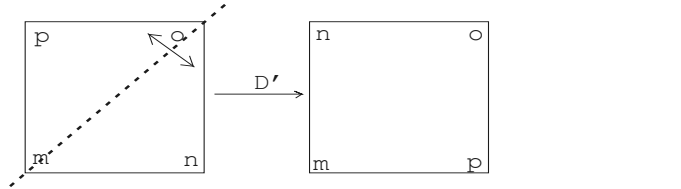
6. Rotation of 180° anticlockwise about vertical axis in the plane of the square.



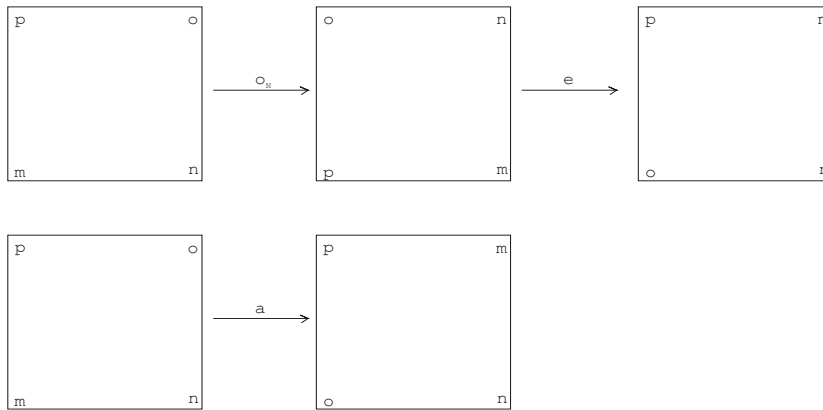
7. Rotation of 180° anticlockwise about the main diagonal.



8. Rotation of 180° anticlockwise about the other diagonal.



Two motions are equivalent if their net effect is the same. For example, a rotation through one right angle clockwise is equivalent to a rotation through three right angles anticlockwise. Moreover the effect of R_1 , followed by H is equivalent to D , as is shown below:



It can be verified that any motion of the square which makes it fit back into the original space is equivalent to one of the above eight motions.

Let $D_8 = \{R_0, R_1, R_2, R_3, H, V, D, D'\}$. On D_8 define a binary operation as follows:

For $a, b \in D_8$ $(ab)\square = a(b\square)$ where $a\square$ means the effect of 'a' on the square $ABCD$. The composition table of D_8 is as follows

\cdot	R_0	R_1	R_2	R_3	H	V	D	D'
R_0	R_0	R_1	R_2	R_3	H	V	D	D'
R_1	R_1	R_2	R_3	R_0	D'	D	H	V
R_2	R_2	R_3	R_0	R_1	V	H	D'	D
R_3	R_3	R_0	R_1	R_2	D	D'	V	H
...
H	H	D	V	D'	R_0	R_2	R_1	R_3
V	V	D'	H	D	R_2	R_0	R_3	R_1
D	D	V	D'	H	R_3	R_1	R_0	R_2
D'	D'	H	D	V	R_1	R_3	R_2	R_0

We observe that for $a, b \in D_8, ab \in D_8$. This property is called closure property of D_8 .

Also note that if $a \in D_8$ then $aR_0 = R_0a = a$.

Thus combining any element of D_8 with R_0 on either side yields back the element. An element R_0 with this property is called identity element (no effect element). Also for each $a \in D_8$, there is exactly one $b \in D_8$ such that $ab = ba = R_0$. Such an element b is called inverse of a and vice versa. For example, R_1, R_3 are inverses of each other, whereas R_0, R_2, H, V, D, D' are their own inverses. If a and b are inverses of each other then b “undoes” whatever a “does” in the sense that a and b taken together in any order produce the “no effect” element, that is, the identity element R_0 .

Observe that the eight motions describe above in Fig. 1 are mappings of $\{P, Q, R, S\}$ onto itself, and the operation is the composition of mappings. Since the composition of mapping is associative, therefore, $a(bc) = (ab)c$ for all $a, b, c \in D_8$

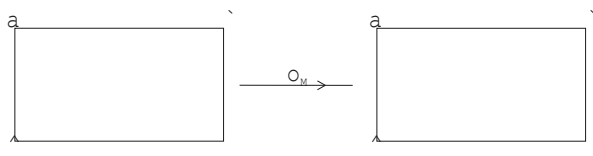
From the table, observe that $R_1V = D$ and $VR_1 = D'$, so that $R_1V \neq VR_1$. Thus the binary composition is not commutative on D_8 .

2.4 Symmetries

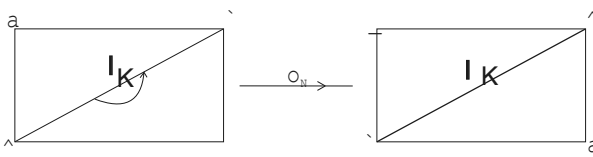
Symmetries of Non-square Rectangle

We study the symmetries of a rectangle which is not a square. Consider a rectangle $ABCD$ with centre O . Take O as origin and line through O parallel to AB and BC as X-axis and Y-axis respectively. Consider the following motions.

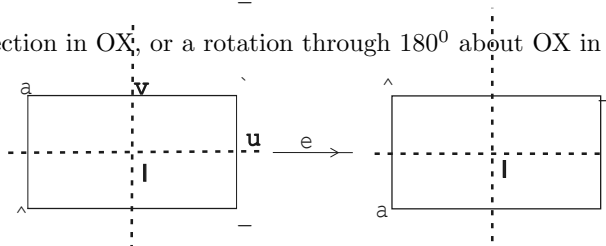
R_0 = Rotation through 0° , i.e. no motion at all.



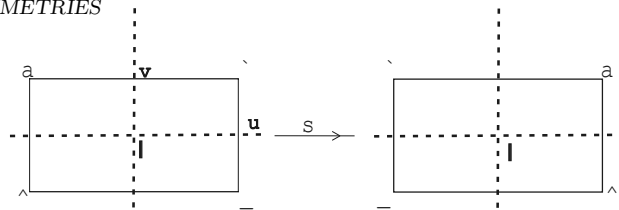
R_1 = rotation through π anticlockwise about the line through O perpendicular to the plane of rectangle.



H = Reflection in OX , or a rotation through 180° about OX in space.



V = Reflection in OY , or a rotation through 180° about OY in space.



These are the only different symmetries of a rectangle. We note that a reflection about any diagonal is not a symmetric motion. It would be a good idea to take a cutout of a rectangle and observe the motions, as was done in the case of a square. The binary operation is the composition of motions. The multiplication table is:

	R_0	R_1	H	V
R_0	R_0	R_1	H	V
R_1	R_1	R_0	V	H

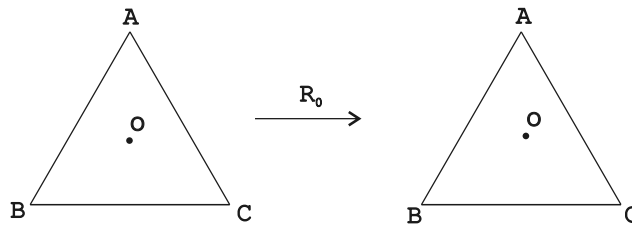
H	H	V	R_0	R_1
V	V	H	R_1	R_0

Let $V_4 = \{R_0, R_1, H, V\}$. From the table we observe that V_4 is closed with respect to composition of motion. R_0 is the identity element and each element is its own inverse. Moreover, the table is symmetric about the main diagonal, V_4 is commutative.

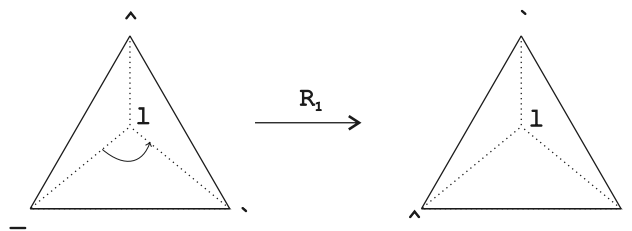
Symmetries of an Equilateral Triangle

Let us now consider the set of symmetries of an equilateral triangle. Let ABC be an equilateral triangle with centre O. Consider the following motions:

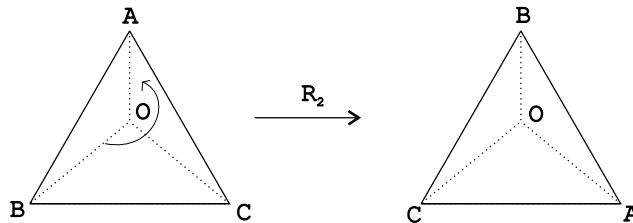
R_0 : Rotation about the centre through 0^0



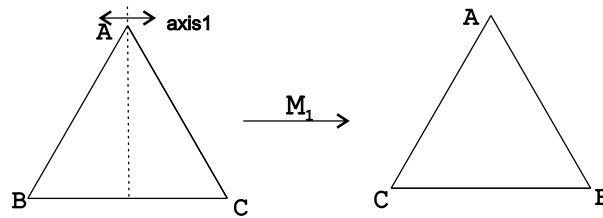
R_1 : Rotation through $\frac{2}{3}\pi$ anticlockwise about the line through O perpendicular to the plane of the triangle



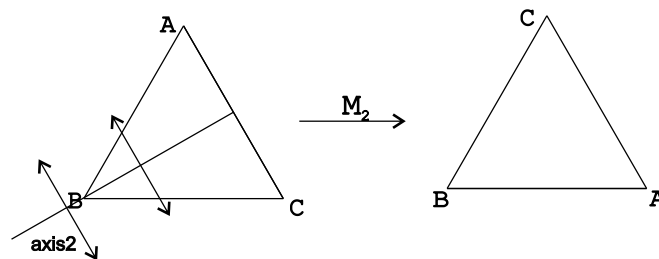
R_2 : Rotation through $\frac{4}{3}\pi$ anticlockwise about the line through O perpendicular to the plane of the triangle.



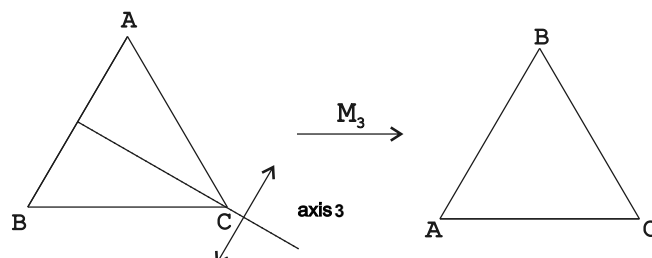
M_1 : Reflection in the axis 1.



M_2 : Reflection in the axis 2.



M_3 : Reflection in the axis 3.



Let $D_6 = \{R_0, R_1, R_2, M_1, M_2, M_3\}$. The operation considered is the composition of motions. The multiplication table is given as follows:

	R_0	R_1	R_2	M_1	M_2	M_3
R_0	R_0	R_1	R_2	M_1	M_2	M_3
R_1	R_1	R_2	R_0	M_3	M_1	M_2
R_2	R_2	R_0	R_1	M_2	M_3	M_1

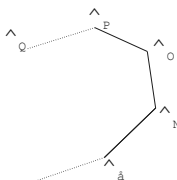
M_1	M_1	M_2	M_3	R_0	R_1	R_2
M_2	M_2	M_3	M_1	R_2	R_0	R_1
M_3	M_3	M_1	M_2	R_1	R_2	R_0

Observe that the 6 motions described above are mapping of $\{A, B, C\}$ onto itself and operation is the composition of mappings. Clearly D_6 is closed. Also associative law holds because the composition of mapping is associative.

R_0 is the identity element. R_1, R_2 are the inverses of each other whereas all other elements are their own inverses. From the table it is clear that $R_2M_1 = M_3, M_1R_2 = M_2$, so that $R_2M_1 \neq M_1R_2$. Thus the binary composition is not commutative on D_6 .

Dihedral group

Let us now study the symmetries of a regular polygon of n sides (n -gon). Consider a regular n -gon A_1, A_2, \dots, A_n . Take a copy of this n -gon and move it in any manner. Now place it on the original n -gon so as to cover it completely. A motion of this nature is called a symmetry of the n -gon.



If $\alpha = \frac{2\pi}{n}$ let R_k denotes the anticlockwise rotation of the polygon about a line through its centre and perpendicular to the plane of the polygon through an angle $k\alpha, k=0, 1, 2, 3, \dots, n-1$. These are the n rotations. There are also n reflections, through the n lines of symmetry $L_i, i=1, 2, 3, \dots, n$.

If n is odd, each line of symmetry passes through a vertex and the mid-point of the opposite side. If n is even, there are two types of lines of symmetry, one type passing through two opposite vertices (and these are $\frac{n}{2}$ in number) and the other type are the perpendicular bisectors of two opposite sides (these are also $\frac{n}{2}$ in number). Thus, in this case also, there are n lines of symmetry.

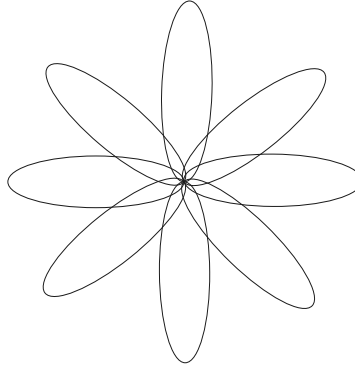
Let D_{2n} be the set of all these rotations and reflections of the regular n -gon. It has $2n$ elements. Let us define a binary operation on D_{2n} . For $A, B \in D_{2n}$, by AB we mean the symmetry obtained by first applying B and then A to the regular n -gon. There are only two types of symmetries, rotations and reflections. Clearly rotation followed by a rotation is a rotation, reflection followed by a reflection is rotation, and rotation followed by reflection (or vice versa) is a reflection. Hence the closure property holds in D_{2n} . Since we are viewing

symmetries as functions on the vertices of the n -gon, AB is just the function composition. Since composition of function is associative, the binary operation is associative. The identity of D_{2n} is the identity symmetry R_0 . The inverse of the rotation R_i is the rotation R_{n-i} , $i=1, 2, \dots, n$ and the inverse of R_0 is R_0 . The inverse of a reflection is itself.

Do it Yourself

Take a piece of cardboard and cut out a regular n -gon from it (you can take a specific value of n , say 6). Paint one face red and the other blue. Take a reflection of this hexagon about any axis. The result of reflection is that if the red face was on the top, the blue face comes on the top. The net result of a reflection is that not only the order of vertices is changed but the face is also reversed. Now apply a rotation to the hexagon. The effect of this is that the face remains of the same colour, only the order of the vertices is changed.

Problem 2.1. Describe all the symmetries of the figure given below:



Solution: *Rotational symmetries*

Rotations about O , the centre of the circle, through angles $0, \frac{\pi}{4}, \frac{2\pi}{4}$ and $\frac{3\pi}{4}$ are the 4 rotational symmetries.

Reflectional symmetries

Reflections about the axis AE, BF, CG and DH are the 4 reflectional symmetries.

Hence there are 8 symmetries in all.

Problem 2.2. Describe the symmetries of the following:

(i) X X X X

(ii) an infinitely long strip of the alphabet X i.e. \dots X X X X X \dots

Solution:

(i) Let O be the centre of the figure. X X · X X

Rotation about the axis through the point O perpendicular to the plane through an angle 0° (i.e. no motion) and 180° .

The reflections are:

(a) Reflection about an axis through O in the plane of the paper.

X X : X X

- (b) Reflection about the horizontal line passing through O as shown.

$$\dots X \dots X O X \dots X \dots$$

- (ii) The rotational symmetry is as in (i), where O is any point in the centre of any two consecutive X 's.

$$\dots X X X O X \dots$$

The reflections are:

- (a) Reflection about an axis midway between any two consecutive X 's in the plane of the paper

$$\dots X X \vdots X X \dots$$

- (b) Reflection about an axis through the centre of any X , in the plane of the paper.

$$\begin{array}{c} \vdots \\ \dots X X X X \dots \\ \vdots \end{array}$$

- (c) Reflection about the horizontal line passing through the centres of all the X 's as shown.

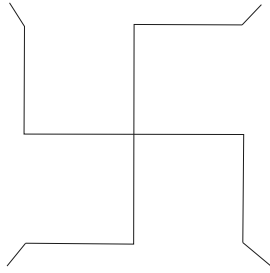
$$\dots\dots X X X X \dots\dots$$

2.5 Exercise

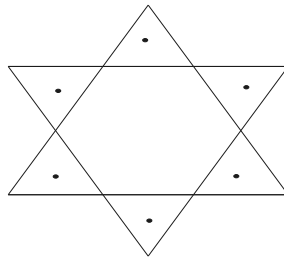
- Describe all the symmetries of the following:
 - circle.
 - isosceles triangle which is not equilateral.
 - scalene triangle.
- Are the following motions of a rectangle (which is not a square) symmetries?
 - reflection about a diagonal.
 - rotation about an axis through the centre perpendicular to the plane, through one right angle.
 - rotation as above through 2 right angles.
- Consider an infinitely long strip of equally spread alphabets. Describe the symmetries of these strips.
 - $\dots O O O O O \dots$
 - $\dots M M M M M \dots$
 - $\dots N N N N N \dots$
 - $\dots T T T T T \dots$
 - $\dots D D D D D \dots$

4. List all the symmetries of the following:

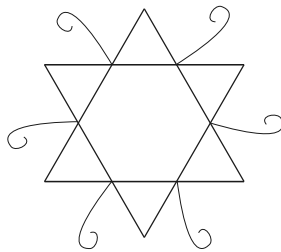
(i)



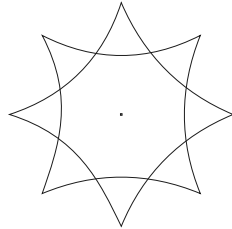
(ii)



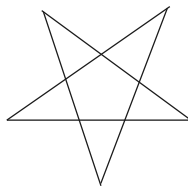
(iii)



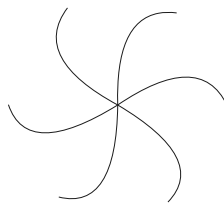
(iv)



(v)



(vi)



2.6 Solved Problems

Problem 2.3. *How many binary operations can be defined on a set with five elements*

Solution:

Let $S = \{a, b, c, d, e\}$.

The Cayley table of S looks like

\odot	a	b	c	d	e
a	*	*	*	*	*
b	*	*	*	*	*
c	*	*	*	*	*
d	*	*	*	*	*
e	*	*	*	*	*

The Cayley table has $5 \times 5 = 25$ entries. Each entry has 5 choices, namely a, b, c, d, e . Moreover all the choices are independent of each other.

\therefore Number of ways in which the table can be completed $= 5^{25}$. Since a binary operation on S corresponds to one way of completing the table.

\therefore Number of binary operations = number of ways in which the table can be completed $= 5^{25}$.

Problem 2.4. *How many commutative binary operations can be defined on a set with five elements*

Solution: Let $S = \{a, b, c, d, e\}$.

Number of binary operations on $S = 5^{25}$

\odot	a	b	c	d	e
a	*	*	*	*	*
b		*	*	*	*
c			*	*	*
d				*	*
e					*

If the binary operation is commutative, then in the Cayley table, the entries below the diagonal are reflection with respect to the diagonal of the entries above the diagonal. The $*$ entries can be chosen arbitrarily from S . These are $1+2+3+4+5=15$ entries. Hence, the number of ways of choosing them $= 5 \times 5 \times 5 \dots 15 \text{ times} = 5^{15}$.

There are 5^{15} commutative binary operations.

Problem 2.5. *How many binary operations having an identity element can be defined on a set with 5 elements*

Solution: Let $S = \{a, b, c, d, e\}$.

Number of binary operations on $S = 5^{25}$. If the binary operation has an identity element, say b , then the Cayley table of S looks like

		identity				
		↓				
		a	b	c	d	e
identity \rightarrow b	a	*	a	*	*	*
	b	a	b	c	d	e
	c	*	c	*	*	*
	d	*	d	*	*	*
	e	*	e	*	*	*

Thus nine elements in the Cayley table have been fixed. The remaining $25 - 9 = 16$ elements, marked $*$ can be chosen arbitrarily from the elements of S .

Thus, number of such binary operations $= 5 \times 5 \times 5 \dots 16 \text{ times} = 5^{16}$.

Problem 2.6. *How many commutative binary operations having an identity element can be defined on a set with 5 elements*

Solution: Let $S = \{a, b, c, d, e\}$.

Without loss of generality, we can take 'a' to be the identity element. If the operation is commutative also, then in the Cayley table only the entries marked \star have to be chosen. The Cayley table is:

\odot	a	b	c	d	e
a	a	b	c	d	e
b	b	\star	\star	\star	\star
c	c		\star	\star	\star
d	d			\star	\star
e	e				\star

Each entry can be chosen as any one from the elements of S.

Number of entries to be chosen = $1 + 2 + 3 + 4 = 10$.

Number of choices for each entry = 5.

\therefore Number of required binary operation = 5^{10} .

Problem 2.7. *In D_{2n} , explain geometrically why a rotation followed by a rotation must be a rotation.*

Solution: A rotation changes the order of the vertices, while the top face remains the same. Thus a rotation followed by a rotation means the top face will remain the same, only the order of vertices will change, so that it will be a rotation.

Problem 2.8. *In D_{2n} explain geometrically why a reflection followed by a reflection must be a rotation.*

Solution: Reflection means reversing the face of the regular n -gon. Thus a reflection followed by a reflection means face reversed twice, that is, the same face up. Only the order of the vertices may change. Thus it is a rotation.

Problem 2.9. *In D_{2n} , explain geometrically why a rotation and a reflection taken together in either order must be a reflection.*

Solution: Reflection means reversing the face of the regular n -gon, whereas rotation keeps the top face same. Thus a reflection and rotation taken in any order means reversal of the face and hence it is a reflection.

Problem 2.10. *Associate the number +1 with a rotation and number -1 with a reflection. Describe an analogy between multiplying these two numbers and multiplying elements of D_{2n} .*

Solution: A rotation followed by a rotation gives a rotation. Also we have

$$(+1)(+1) = +1.$$

A reflection followed by a reflection means face reversed twice, that is same face up. The net result is a rotation. Also $(-1)(-1) = +1$. A reflection followed by a rotation (or otherwise) gives a reflection (by the above example), Hence its like $(-1)(+1) = -1$ or $(+1)(-1) = -1$.

2.7 Supplementary Exercises

1. State whether the following statements are true or false. Justify your answer.
 - (i) A binary operation on a set S assigns at least one element of S to each ordered pair of elements of S .
 - (ii) A binary operation on a set S assigns not more than one elements of S to each ordered pair of elements of S .
 - (iii) A binary operation on a set S assigns exactly one element of S to each ordered pair of elements of S .
 - (iv) Every binary operation on a set consisting of 2 elements is commutative.
 - (v) If \star is a commutative binary operation on S , then

$$a \star (b \star c) = (c \star b) \star a \quad \forall a, b, c \in S.$$
 - (vi) Let S be the set of all 2×2 matrices over \mathbb{Z} and \star the usual matrix multiplication. Then
 - (a) \star is associative
 - (b) \star is commutative
 - (c) (S, \star) has identity element
 - (d) elements of S with non-zero determinant are invertible.
 - (vii) Addition on the set of odd integers is a binary operation.
 - (viii) On the set of even integers, \star defined by $a \star b = \frac{a+b}{2}$ is a binary operation.
 - (ix) A rectangle has a symmetry about both the diagonals.
 - (x) A rectangle has a rotational symmetry through an angle of 90° .
 - (xi) The symmetries of an equilateral triangle commute.
 - (xii) The symmetries of a rectangle commute.
 - (xiii) A regular pentagon has 10 symmetries.
 - (xiv) The English alphabet X has 2 symmetries.
 - (xv) Two mutually perpendicular lines has 6 symmetries.
2. Which of the following are binary operations on \mathbb{N} ?
 - (i) $m \star n = m - n$
 - (ii) $m \star n = m \div n$
 - (iii) $m \star n = n$
 - (iv) $m \star n = m + n + m^2$
 - (v) $m \star n = 4m + 5n$
 - (vi) $m \star n = m + n - 1$
 - (vii) $m \star n = m + n - mn$
 - (viii) $m \star n = mn - (m + n)$
 - (ix) $m \star n = \begin{cases} m + n - mn, & \text{if } m = 1 \text{ or } n = 1 \\ mn - m - n, & \text{if } m \neq 1 \text{ and } n \neq 1 \end{cases}$

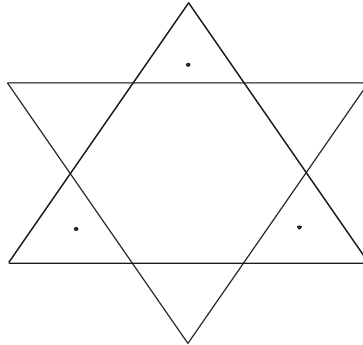
For the binary operations, check which of them are commutative, associative and have identity element.

3. Give an example of an infinite set S and a binary operation \star on S such that
 - (i) exactly one element of (S, \star) has an inverse
 - (ii) exactly two elements of (S, \star) have inverses
 - (iii) every element of S , except one, has an inverse.
4. Write the Cayley table for all the binary operations which can be defined

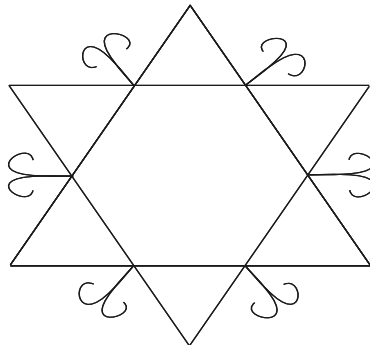
- on $S = \{a, b\}$
- (i) How many of them are commutative
 - (ii) How many have an identity element
 - (iii) How many are commutative and have an identity element
5. Write the Cayley table for 6 binary operations on $S = \{a, b, c\}$, which
- (i) are commutative
 - (ii) have an identity element
 - (iii) are commutative and have an identity element.
6. Let S be a non-empty set and \star a binary operation on S . An element $e \in S$ is called left (right) identity with respect to \star .
 if $e \star a = a \quad \forall a \in S$, ($a \star e = a \quad \forall a \in S$). Give example of a binary operation \star on a set S which has
- (i) left identity but not right identity
 - (ii) right identity but not left identity
 - (iii) right identity and left identity. Are they different?
Can you generalize your answer?
7. Give examples of a set and a binary operation \star on S such that
- (i) (S, \star) has two distinct left identities
 - (ii) (S, \star) has two distinct, right identities
 - (iii) every element of S is a left identity
 - (iv) (S, \star) has some invertible element and some elements which do not have an inverse.
8. How many binary operations can be defined on a finite set S with
- (i) 2 elements
 - (ii) 3 elements
 - (iii) 4 elements
 - (iv) n elements?
9. How many commutative binary operations can be defined on a finite set S with
- (i) 2 elements
 - (ii) 3 elements
 - (iii) 4 elements
 - (iv) n elements?
10. How many binary operations having an identity element can be defined on a finite set S with
- (i) 2 elements
 - (ii) 3 elements
 - (iii) 4 elements
 - (iv) n elements?
11. How many commutative binary operations, having an identity element, can be defined on a finite set S with
- (i) 2 elements
 - (ii) 3 elements
 - (iii) 4 elements
 - (iv) n elements?

12. Give two examples of figures for each of the following:
- having only rotational symmetry
 - having only reflection symmetry
 - having both rotational and reflection symmetry
 - having neither rotational nor reflection symmetry.
13. Describe all the symmetries of the following:
- parallelogram which is neither a rectangle nor a rhombus
 - rhombus which is not a square
 - ellipse which is not a circle
 - hyperbola
 - right angled isosceles triangle
 - right angled triangle which is not isosceles
14. List all the symmetries of following figures:

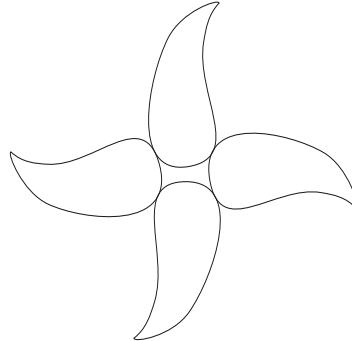
i



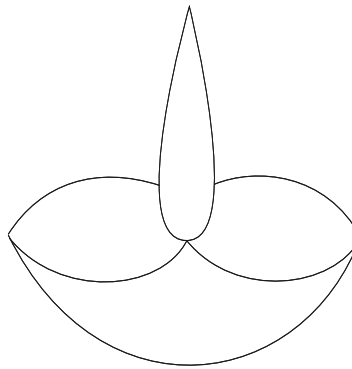
ii



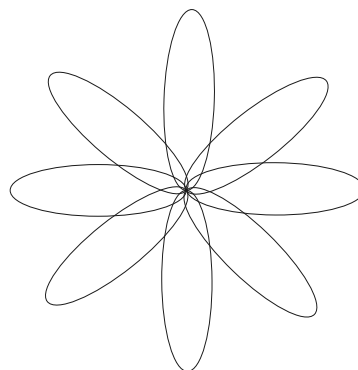
iii



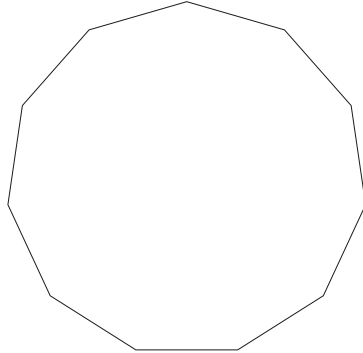
iv



v



15. Find the symmetries of the Indian 2-rupee coin shown below. Disregard the printing and figure on the coin.



2.8 Answers to Exercises

Exercise - 2.2

1. (i) Yes (ii) No (iii) No (iv) Yes (v) No.

2.

(i)	\star	1	2	3
	1	2	3	1
	2	1	2	2
	3	1	2	3

(ii)	\circ	1	2	3
	1	1	1	1
	2	1	1	1
	3	1	1	1

(iii)	\min	1	2	3
	1	1	1	1
	2	1	2	2
	3	1	2	3

3. (i) No, $\because 4 \notin S$ (ii) No, $\because 4 * 1$ is not defined.

4. i commutative, not associative

ii commutative, associative

iii not commutative, not associative

iv not commutative, associative

v not commutative, not associative

vi not commutative, not associative

5. (i) 2^4 (ii) 4^{16} (iii) 8^{64} (iv) n^{n^2}

6. i On \mathbb{Z} , define $a \star b = |a - b|$.

ii On the set of 2×2 matrices over \mathbb{Z} , define $A \star B = AB$, the usual multiplication of matrices.

iii On \mathbb{Z} define, $a \star b = a - b$.

Exercise - 2.5

1. (i) Reflection about every diameter. Rotation about an axis through the centre perpendicular to the plane of the circle, through an angle α , where α is any real number.

(ii) Identity motion, reflection in the median through the vertex.

(iii) Identity motion (i.e. no motion).

2. (i) No (ii) No (iii) Yes

3. (i) Reflection about an axis

(a) between any two O 's in the plane of the paper.

$$\dots O O | O O O \dots$$

(b) through the centre of any O in the plane of the paper.

$$\begin{array}{c} \vdots \\ \dots O O O O O \dots \\ \vdots \end{array}$$

(c) through the centre of all the O 's

$$\dots O O O O O \dots$$

Rotation about the axis L through the mid point of the centres of any two consecutive O 's through an angle (a) 0° (b) 180°
Hence there are infinitely many symmetries

- (ii) Reflection about an axis

(a) midway between any two consecutive M 's.

$$\dots M M | M M \dots$$

(b) through the middle of any M .

$$\begin{array}{c} \vdots \\ \dots M M M M \dots \\ \vdots \end{array}$$

Rotation about the axis L through the midpoints of any two consecutive M 's through an angle of 0° i.e. no motion.

- (iii) Only the 'no motion' rotation.

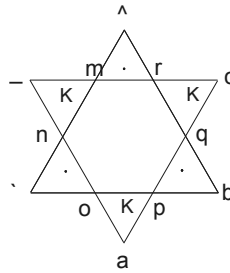
(iv) Similar to (ii).

- (v) Reflection about the horizontal line bisecting the D 's

$$\dots D D D D \dots$$

The 'no motion' rotation.

4. (i)



Rotation about the axis through O perpendicular to the plane through an angle (a) 0° (b) 90° (c) 180° and (d) 270° .

(ii) 6 rotations about an axis through O perpendicular to the plane through angles 0° , 60° , 120° , 180° , 240° , 300°

Reflection about lines

(a) AD (b) BE (c) CF (d) PS (e) QT (f) RU .

Thus there are 12 symmetries in all

(iii) Only 6 rotations

(iv) 8 rotations about the centre, through angles $k\frac{\pi}{4}$, $k = 0, 1, \dots, 7$
8 reflections.

(v) 5 rotations and 5 reflections.

(vi) 6 rotations.

Supplementary Exercise

1.

(i) F. It assigns exactly one element

(ii) F. It assigns exactly one element

(iii) T

(iv) F, let $S = \{a, b\}$. Define \star by

\star	a	b
a	b	a
b	b	b

(v) T

(vi) (a) T, (b) F, (c) T, (d) F, true when the determinant is ± 1

(vii) F, $3+5=8$, 8 is not odd.

(viii) F, $6 \star 8=7$ which is not even.

(ix) F

(x) F

(xi) F, $\because R_1 M_2 \neq M_2 R_1$

(xii) T

(xiii) T

(xiv) F, it has 4 symmetries

(xv) F, it has 8 symmetries

2. (i) No (ii) No (iii) Yes (iv) Yes (v) Yes (vi) Yes (vii) No (viii) No (ix) No.

3. (i) (\mathbb{N}, \cdot) (ii) (\mathbb{Z}, \cdot) (iii) (\mathbb{Q}, \cdot)

4.

\cdot	a	b
a	\star	\star
b	\star	\star

Each \star has 2 choices a or b . Thus the number of possible tables=16.

(i) 2^3 (ii) 2 (iii) 2.

6. Let $S = \{a, b, c, d\}$

(i) Define $x * y = y, \forall x, y \in S$.

Then every element of S is a left identity. It does not have a right identity.

(ii) Define $x \odot y = x, \forall x, y \in S$.

Then every element of S is a right identity. There is no left identity.

(iii) Let $S =$ set of all 2×2 matrices with integral entries. Consider (S, \cdot) . Then

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is a right as well as left identity.

They are not different. If right and left identity both exist then they must be the same.

7. For (i), (ii), (iii) see 6(i) and (ii).

(iv) $M_{2 \times 2}(\mathbb{Z})$, the set of all 2×2 matrices over \mathbb{Z} . All elements with determinant ± 1 are invertible.

8. (i) 2^4 (ii) 3^9 (iii) 4^{16} (iv) n^{n^2} .

9. (i) 2^3 (ii) 3^6 (iii) 4^{10} (iv) $n^{n(n+1)/2}$.

10. (i) 2 (ii) 3^4 (iii) 4^9 (iv) $n^{(n-1)^2}$.

11. (i) 2 (ii) 3^3 (iii) 4^6 (iv) $n^{\frac{n(n-1)}{2}}$.

12.

(i) 2 rotational symmetries, through 0° and 180° .

(ii) 2 rotational and 2 reflections in diagonals.

(iii) 2 rotational and 2 reflections in the major and minor axes.

(iv) 2 rotational and 2 reflections in the transverse and conjugate axes.

(v) One rotational (through 0°) and 1 reflection about the median through the vertex.

(vi) One rotational (through 0°)

13.

(i) 2 rotations

(ii) 2 rotations, 2 reflections

(iii) 2 rotations, 2 reflections (one each in the major and minor axis)

(iv) 2 rotations, 2 reflections (one each in the transverse and conjugate axes).

(v) One rotation (no motion), one reflection in the median through the vertex.

(vi) One rotation (no motion) only.

14.

(a) 3 rotations, 3 reflections

(b) 6 rotations, 6 reflections

(c) 2 rotations,

(d) 1 reflection

(e) 8 rotations through angles $\frac{k\pi}{4}$, $0 \leq k \leq 7$

15. It is a regular 11-gon. So 11 rotations and 11 reflections.

Chapter 3

Functions

Function is a commonly used word in everyday life having different meanings. But, in Mathematics, the concept of a function is very basic and is of fundamental importance. Moreover, it has a very specific meaning. In this chapter we define mathematically a function and their various types and study their properties in detail. Operation on functions, conditions of invertibility and computation of the inverse of an invertible function will also be discussed.

3.1 Definition and Representation

Definition 3.1. Let A and B be two sets. A binary relation f from A to B is called a function (or mapping) from A to B if each element of A is related to exactly one element of B . In other words, f is a function from A to B if for each element $a \in A$ there exists exactly one element $b \in B$ such that $(a, b) \in f$, b is called the image of a under f and we write $b = f(a)$.

In the above definition, the set A is called the domain and the set B is called the codomain (or target) of f . The set $\{b \in B \mid (a, b) \in f, \text{ for some } a \in A\}$ is called the range of f . Thus, the range of f is the set of images of the elements of A . If for any $b \in B$ there exists an $a \in A$ such that $f(a) = b$, then a is called a preimage of b under f . An element $b \in B$ can have more than one preimages in A .

Example 3.1. Let $X = \{1, 2, 3, 4\}$, $Y = \{x, y, z\}$, $f = \{(1, z), (2, y), (3, x), (4, y)\}$. X is the domain of f , Y is the codomain (or target) of f . Image of 1 under f is z . We write $f(1) = z$. Similarly $f(2) = y$, $f(3) = x$, $f(4) = y$. Preimage of x is 3. $f(2) = y$ and $f(4) = y$ tells us that preimage of y is 2 as well as 4. Thus y has two preimages.

Let $g = \{(1, z), (2, y), (3, x), (4, y), (2, z)\}$. Since 2 is the first member of two elements of g , i.e $g(2) = y$ and $g(2) = z$, therefore 2 has two images, y and z , hence g is not a function.

Let $h = \{(1, z), (2, x), (3, y)\}$, Since 4 does not have any image, therefore h is not a function.

Notation: ‘ f is a function from set A to a set B ’ is written as $f : A \rightarrow B$ or $A \xrightarrow{f} B$. If the image of $a \in A$ under f is b we say that f maps a to b and we write $b = f(a)$. This is also written as $f : a \rightarrow b$. The domain of a function f will be denoted by $D(f)$ and the range of a function f will be denoted by $R(f)$. Here $D(f) = A$ and $R(f) \subseteq B$.

Example 3.2. Define $f : \mathbb{Z} \rightarrow \mathbb{Z}$ by $f(x) = x^2 + 5$. The domain and codomain of f are both \mathbb{Z} .

$$\begin{aligned} R(f) &= \{f(x) | x \in \mathbb{Z}\} \\ &= \{n \in \mathbb{Z} | n - 5 \text{ is a perfect square}\} \end{aligned}$$

Let us find the preimage of 105, Since $105 - 5 = 100 = 10^2$, $\therefore 10$ is a preimage of 105 and $f(10) = 105$. What is the preimage of 7? $7 - 5 = 2 \neq x^2$ for any $x \in \mathbb{Z}$, $\therefore 7$ does not have a preimage. So $7 \notin R(f)$.

When we define a function, sometimes there is some ambiguity, i.e., an element can have more than one images or the images do not lie in the codomain. In such a situation we say that the function is not well defined. In fact, a function $f : A \rightarrow B$ is well defined if

- (i) $f(a)$ is defined for each $a \in A$
- (ii) $f(a) \in B$ for all $a \in A$
- (iii) There is no ambiguity in determining $f(a)$.

This is illustrated in the following examples:

Example 3.3.

- (i) Define $f : \mathbb{Z} \rightarrow \mathbb{Z}$ by $f(n) = -n$ for all $n \in \mathbb{N}$. Here $f(m)$ is not defined, for $m \in \mathbb{Z}$, $m \leq 0$ so f is not a well-defined function.
- (ii) Define $g : \mathbb{N} \rightarrow \mathbb{N}$ by $f(n) = n - 5 \quad \forall n \in \mathbb{N}$. Then $f(1) = -4 \notin \mathbb{N}$, so that f is not a well defined function.
- (iii) Define $h : \mathbb{Q} \rightarrow \mathbb{Z}$ by $h(\frac{a}{b}) = a$. Here $h(\frac{1}{2}) = 1$.
Since $\frac{1}{2} = \frac{2}{4}$
 $\therefore h(\frac{2}{4}) = 2$.
Thus, $h(\frac{1}{2})$ is defined as 1 as well as 2 so that there is an ambiguity in defining $f(\frac{1}{2})$. Therefore, h is not a well-defined function.

Definition 3.2. (Equal functions): Two functions f and g are equal if they have

- (i) the same domain i.e $D(f) = D(g)$
- (ii) the same codomain
- (iii) $f(x) = g(x)$ for all $x \in D(f)$.

Arrow Diagram for Function

If A and B are finite sets of small orders, then a function from A to B can be defined by an arrow diagram also. This enhances the understanding and is very convenient. In Fig. 1, f is a function.

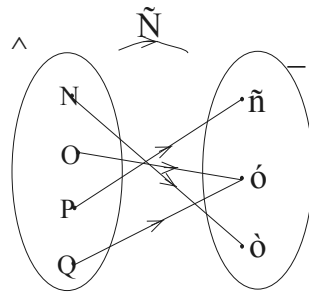


Fig. 1

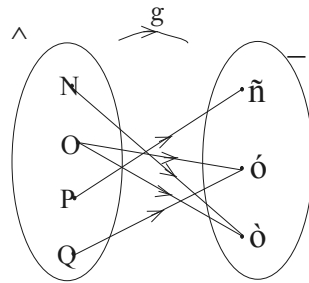


Fig. 2

In Fig. 2, under g , 2 is mapped to y as well as z , so that g is not a function.

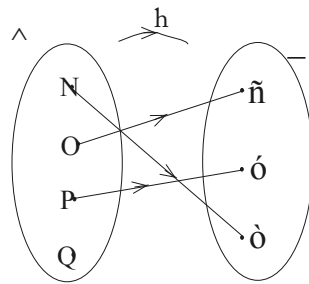


Fig. 3

In Fig. 3, under h , 4 does not have an image, so h is not a function. Observe that in this diagrammatic representation for a function from each point of the domain one and only one line should emerge. The arrow diagrams are specially useful in giving counter examples for functions.

Representation of a Function

In general there are four ways to represent a function.

- (i) Verbally (in words) y is the square of x
- (ii) Numerically (table of values)

x	1	2	3	4	5
y	1	4	9	16	25

Fig. 4(Tabular form of a function)

- (iii) Algebraically (formula) $y = x^2$
- (iv) Visually (graph or arrow diagram).

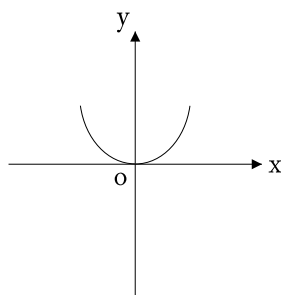


Fig. 5(Graphical representation of a function)

As shown above, the same function has been represented in all the four ways. If a single function can be represented in all four ways, it is often useful to go from one representation to another to gain insight into the function. Certain functions are described more naturally by one method than by another.

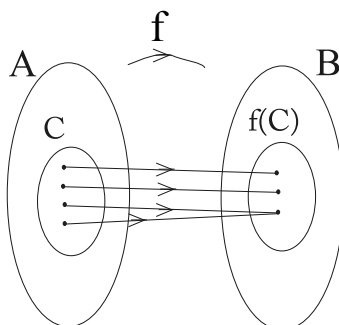
3.2 Images and Inverse Images

If $f : A \rightarrow B$ is a function and $C \subseteq A$, then the set of images of elements of C is called the image of C and is denoted by $f(C)$.

Symbolically, $f(C) = \{f(x) | x \in C\}$

Clearly, $f(C) \subseteq B$. We may also write $f(C)$ as:

$f(C) = \{y \in B \mid y = f(a) \text{ for some } a \in C\}$.



Theorem 3.1. Let $f : A \rightarrow B$ be a function. If A_1, A_2 are subsets of A , then

- (i) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$
- (ii) $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$
- (iii) $A_1 \subseteq A_2 \Rightarrow f(A_1) \subseteq f(A_2)$

Proof: Left to the reader. \square

Remark 3.1. (i) Equality may not hold in (ii).

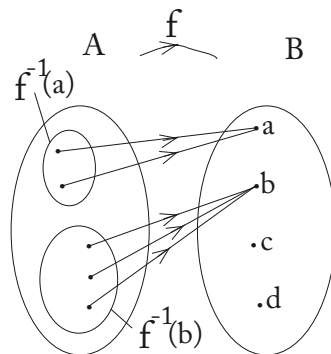
- (ii) Define $f : \mathbb{Z} \rightarrow \mathbb{Z}$
 by $f(z) = z^2, \forall z \in \mathbb{Z}$.
 Let $A_1 = \{1, 2, 3\}$ and $A_2 = \{-1, -2, 3, 4\}$.
 $f(A_1 \cap A_2) = \{9\}$, $f(A_1) \cap f(A_2) = \{1, 4, 9\}$
 Thus, $f(A_1 \cap A_2) \neq f(A_1) \cap f(A_2)$.

- (iii) If $f : A \rightarrow B$ is a function and $a \in A$, then $f(a) \in B$, whereas $f(\{a\}) \subseteq B$.
 Thus, $f(a)$ is not the same as $f(\{a\})$.

Inverse Images

Given a function f from A to B , \leftarrow we now explain what we mean by the inverse image of an element of B . Let $f : A \rightarrow B$ be a function and let $b \in B$. The set of all elements of A which are mapped to b is called the inverse image of b under f and is denoted by $f^{-1}(b)$. Symbolically, $f^{-1}(b) = \{a \in A \mid f(a) = b\}$.

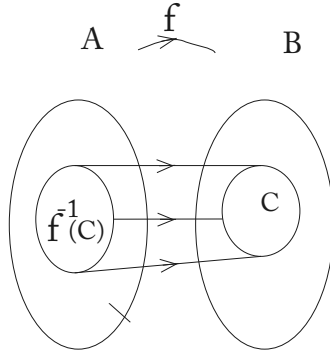
Example 3.4. Let $A = \{1, 2, 3, 4, 5\}$, $B = \{a, b, c, d\}$, $f = \{(1, a), (2, b), (3, a), (4, b), (5, b)\}$. Thus, $f(1) = f(3) = a$, $f(2) = f(4) = f(5) = b$. $f^{-1}(a) = \{1, 3\}$, $f^{-1}(b) = \{2, 4, 5\}$, $f^{-1}(c) = f^{-1}(d) = \phi$.



Inverse image of a set

Let $f : A \rightarrow B$ be a function and $C \subseteq B, C \neq \phi$. The set of all elements of A whose images belong to C is called the inverse image of C under f and is denoted by $f^{-1}(C)$. Thus $f^{-1}(C) = \{x \in A \mid f(x) \in C\}$.

We define $f^{-1}(\phi) = \phi$.



Clearly, $f^{-1}(B) = A$. In fact, $f^{-1}(R(f)) = A$.

Example 3.5. Consider the function in Example 3.4.

If $C = \{a, c\}$, then $f^{-1}(C) = \{1, 3\}$.

If $D = \{b, c, d\}$, then $f^{-1}(D) = \{2, 4, 5\}$.

If $E = \{a\}$, then $f^{-1}(E) = \{1, 3\}$.

In fact, $f^{-1}(\{a\}) = f^{-1}(a)$.

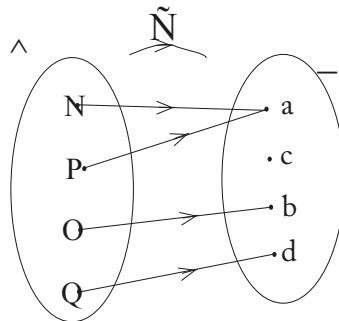
Remark 3.2. If f is a function, then $f^{-1}(a) = f^{-1}(\{a\})$.

Theorem 3.2. Let $f : A \rightarrow B$ be a function and B_1, B_2 subsets of B . Then

- (i) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$
- (ii) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$
- (iii) $B_1 \subseteq B_2 \Rightarrow f^{-1}(B_1) \subseteq f^{-1}(B_2)$
- (iv) $f(f^{-1}(B_1)) \subseteq B_1$, and equality may not hold
- (v) If $C \subseteq A$ then $C \subseteq f^{-1}(f(C))$
- (vi) $f^{-1}(B_1 \setminus B_2) = f^{-1}(B_1) \setminus f^{-1}(B_2)$
- (vii) $f^{-1}(B_1^c) = (f^{-1}(B_1))^c$.

Proof: We shall prove only (iv) and (v).

- (iv) Let $y \in f(f^{-1}(B_1))$. Then $y = f(x)$ for some $x \in f^{-1}(B_1)$.
 $x \in f^{-1}(B_1) \Rightarrow f(x) \in B_1 \Rightarrow y \in B_1$. $f(f^{-1}(B_1)) \subseteq B_1$.
 Let $A = \{1, 2, 3, 4\}$, $B = \{a, b, c, d\}$.



Let $B_1 = \{b, c\}$. Then $f^{-1}(B_1) = \{2\}$. $f(f^{-1}(B_1)) = \{b\} \neq B_1$. Hence proved.

(v) Let $x \in C$. Then $f(x) \in f(C)$. So that $x \in f^{-1}(f(C))$.

Hence $C \subseteq f^{-1}(f(C))$. Consider the function f as in (iv). Let $C = \{1, 2\}$. Then $C \subseteq A$. $f^{-1}(f(C)) = f^{-1}\{a, b\} = \{1, 2, 3\} \neq C$. \square

3.3 Types of Functions

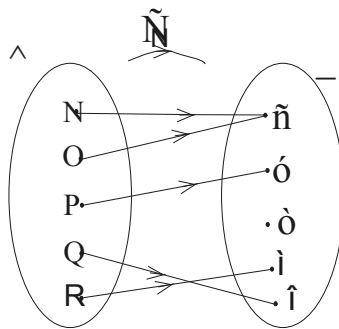
We now study functions with special properties.

Definition 3.3. Let $f : A \rightarrow B$ be a function. Then

- (i) f is said to be onto (or surjective) if $R(f) = B$, i.e for every $b \in B$ there exists some $a \in A$ such that $f(a) = b$.
- (ii) f is said to be one-one (or injective) if distinct elements of the domain have distinct images, i.e if $a, b \in A$ such that $a \neq b$ then $f(a) \neq f(b)$. If a function is not one-one it is many-one.
- (iii) f is bijective if it is both onto and one-one.

The condition for a function $f : A \rightarrow B$ to be onto is that for every $b \in B$, $f(x) = b$ has a solution in A . Usually, the contrapositive of the condition for one-one is used. That is, f is one-one if $a, b \in A$ such that $f(a) = f(b)$, then we must have $a = b$. We illustrate the concept of onto and one-one by using arrow diagram for functions.

Example 3.6. Let f_1 be a function defined by



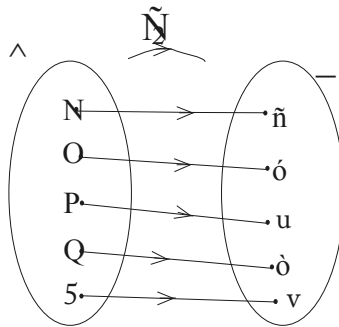
Neither one-one nor onto function.

Then f_1 is a function which is not onto, as no element of A is mapped to z , i.e $f_1(a) = z$ does not have a solution in A .

f_1 is not one-to-one as $1 \neq 2$ but $f_1(1) = f_1(2) = x$.

Thus f_1 is neither one-one nor onto function.

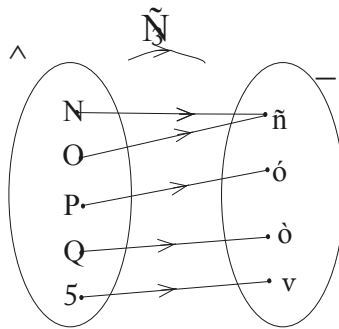
Let f_2 be a function defined by



One-one and onto function

Then f_2 is both one-to-one and onto.

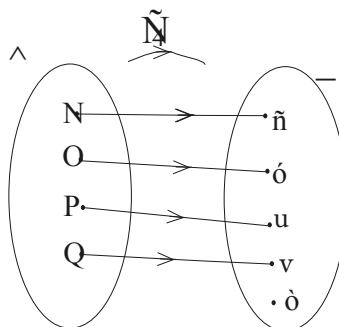
Let f_3 be a function defined by



Onto but not one-one function

Then f_3 is not one-to-one, because $1 \neq 2$ but $f_3(1) = x = f_3(2)$. It is onto.

Let f_4 be a function defined by



One-one but not onto function

f_4 is one-one, as distinct elements have distinct images. f_4 is not onto as $z \in B$ does not have a preimage. $R(f) = \{x, y, u, v\} \neq B$.

The above examples show that the property of being onto or one-one are independent of each other. A function may be neither one-one nor onto as f_1 , both one-one and onto as f_2 , onto but not one-one as f_3 or only one-one but not onto as f_4 .

Example 3.7. Define $f : \mathbb{Z} \rightarrow \mathbb{Z}$ by $f(x) = x + 4 \forall x \in \mathbb{Z}$. We show that f is one-one.

Let $x, y \in \mathbb{Z}$ (domain) such that $f(x) = f(y)$. Then $x + 4 = y + 4$, so that $x = y$. Thus f is one-to-one.

Now, we show that f is onto.

Let $y \in \mathbb{Z}$, the codomain. Suppose there exists $x \in \mathbb{Z}$, the domain, such that $f(x) = y$. Then $x + 4 = y$ so that $x = y - 4$. Thus for each $y \in$ codomain, there exists $y - 4 \in$ Domain, such that $f(y - 4) = y$. Hence f is onto.

So f is both one-one and onto function.

Example 3.8. Define $g : \mathbb{Z} \rightarrow \mathbb{Z}$ by $g(x) = 2x + 4$. g is one-one (prove it!).

It is not onto, $7 \in \mathbb{Z}$ the codomain. Suppose there is $x \in \mathbb{Z}$, the domain such that $g(x) = 7$. Then $2x + 4 = 7 \Rightarrow 2x = 3$, which has no solution in \mathbb{Z} . Hence g is not onto.

Example 3.9. Define $f : \mathbb{Z} \rightarrow \mathbb{Z}$ by $f(x) = x^2$. Then f is not onto for $7 \in \mathbb{Z}$ the codomain, suppose $x \in \mathbb{Z}$ such that $f(x) = 7 \Rightarrow x^2 = 7$. But this equation has no solution in \mathbb{Z} . Thus f is not onto. Let $B = \{x^2 \mid x \in \mathbb{Z}\}$.

Define $f : \mathbb{Z} \rightarrow B$ by $f(x) = x^2$. Let $b \in B$. Then $b = x^2$ for some $x \in \mathbb{Z}$. Thus $f(x) = x^2 = b$. or $f(x) = b$. So that f is onto.

The above example shows that by changing the codomain, an onto function may cease to be onto. Thus the property of being onto depends upon the codomain. The property of being onto also depends on the domain as shown in the Example 3.11.

Example 3.10. Define $f_1 : \mathbb{Q} \rightarrow \mathbb{Z}$ by $f_1(x) = 2x + 4$.

f_1 is onto. For, let $y \in \mathbb{Z}$, suppose there exists $x \in \mathbb{Z}$ such that $f_1(x) = y \Rightarrow 2x + 4 = y \Rightarrow x = (y - 4)/2$. Thus for each $y \in \mathbb{Z}$, there exists $x = (y - 4)/2 \in \mathbb{Q}$ such that $f_1(x) = y$. Hence f_1 is onto. We note that the function $f : \mathbb{Z} \rightarrow \mathbb{Z}$ defined by $f(x) = 2x + 4$ is not onto. f and f_1 have the same codomain, and same rule, only their domains are different.

The following examples show that the property of a function being one-one is dependent on the domain.

Example 3.11. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^2$. f is not one-one, because $1, -1 \in \mathbb{R}$ (the domain). But $f(1) = 1 = f(-1)$. Define $g : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}$ by $g(x) = x^2$. Let $x_1, x_2 \in \mathbb{R}^+ \cup \{0\}$ such that $g(x_1) = g(x_2)$ then $x_1^2 = x_2^2$, so that $x_1 = x_2$. Hence g is one-one. The only difference in the functions f and g is in their domain.

Definition 3.4. (Restriction of a function):

Let $f : B \rightarrow C$ be a function and $A \subseteq B$. The function $g : A \rightarrow C$ defined by $g(x) = f(x) \forall x \in A$, is called the restriction of f to A . It is denoted by $f|_A$.

Definition 3.5. (Extension of a function):

Let $f : A \rightarrow C$ be a function and $B \supseteq A$. A function $g : B \rightarrow C$ is called an extension of f if $g|_A = f$.

Example 3.12. Let $i^2 = -1$ and $\mathbb{Z}[i] = \{a + ib \mid a, b \in \mathbb{Z}\}$. Define $f : \mathbb{Z}[i] \rightarrow \mathbb{Z}$ by $f(a + ib) = a^2 + b^2$ and $\mathbb{Z} \subseteq \mathbb{Z}[i]$. Define $g : \mathbb{Z} \rightarrow \mathbb{Z}$ by $g(x) = x^2 \forall x \in \mathbb{Z}$. For $x \in \mathbb{Z}$, $g(x) = x^2 = x^2 + 0^2 = f(x + i0) = f(x)$. Hence $g(x) = f(x) \forall x \in \mathbb{Z}$, so that $f|_{\mathbb{Z}} = g$ i.e g is a restriction of f to \mathbb{Z} .

Example 3.13. Define a function $f : \mathbb{Z} \rightarrow \mathbb{Q}$ by $f(a) = |a|$. \mathbb{Z} is a subset of \mathbb{Q} . On \mathbb{Q} define $g : \mathbb{Q} \rightarrow \mathbb{Q}$ by $g(\frac{a}{b}) = \frac{|a|}{|b|}$. Then $g(a) = |a| \forall a \in \mathbb{Z}$. Thus $g(a) = f(a) \forall a \in \mathbb{Z}$, so that g is an extension of f .

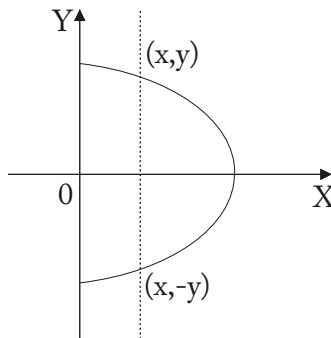
Define $h : \mathbb{Q} \rightarrow \mathbb{Q}$ by $h(x) = \begin{cases} |x| & \text{if } x \in \mathbb{Z}, \\ 0 & \text{otherwise} \end{cases}$

Then $h(x) = f(x)$ for all $x \in \mathbb{Z}$, so that h is also an extension of f . This shows that the g and h are two extensions of the function f . Thus, extension of a function need not be unique.

3.4 Real Valued Functions

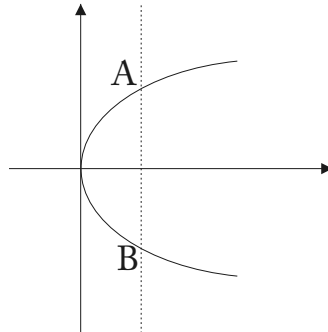
While defining a function $f(x)$, sometimes we want the variable x to take real values and the value of the function $f(x)$ should also be a real number. Thus we have: "If a function $f(x)$ is defined from a subset of \mathbb{R} to a subset of \mathbb{R} , then we say that $f(x)$ is a real valued function of a real variable." For such functions, the domain and range are both subsets of \mathbb{R} . Thus, if $f : \mathbb{R} \rightarrow \mathbb{R}$, then by the domain of f , we mean $\{x \in \mathbb{R} \mid f(x) \text{ is defined and real}\}$. This is called the natural domain of f . If the graph of a function $y = f(x)$ is drawn, then any line parallel to the Y -axis should intersect the graph in at most one point. If it intersects it in 2 or more points then it is not a function. This is because there are more than one value of the function for a given value in the domain. This is called the vertical line test.

Example 3.14. Consider the semi-circle $x^2 + y^2 = 1, x > 0$, i.e $y = \pm\sqrt{1 - x^2}$.



As shown, a line parallel to Y -axis meets the graph in two points, namely (x, y) and $(x, -y)$. Thus, y defined by above is not a function. The above test is called the vertical line test. This test is very useful, because from the graph we can determine whether it is a function or not.

Example 3.15. The graph of $y^2 = 4x$ is as shown.

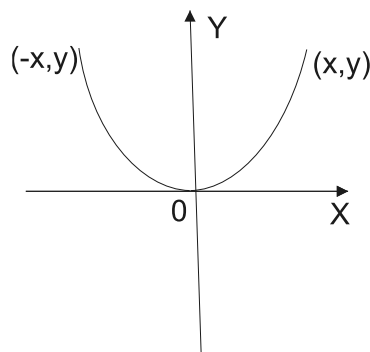


Does it define a function? Using the vertical line test we see that a line parallel to Y -axis intersects the curve in two points A and B . So it does not represent a function.

From the graph we can also determine whether a given function is one-to-one or not. If a line parallel to X -axis intersects the graph in two or more points, then it is not a one-to-one function.

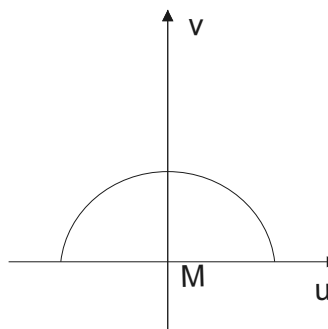
Example 3.16. Consider $y = x^2$. Is it a function? Is it one-to-one?

The graph is



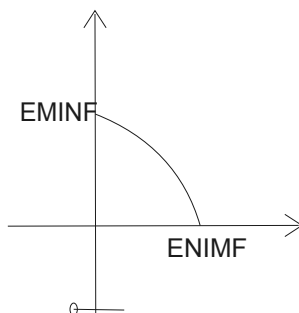
Using the vertical line test we see that it is a function. Now using the horizontal line test, we conclude that it is not a one-one function, as $x_1^2 = (-x_1)^2$.

Example 3.17. Consider $x^2 + y^2 = 1, y > 0$



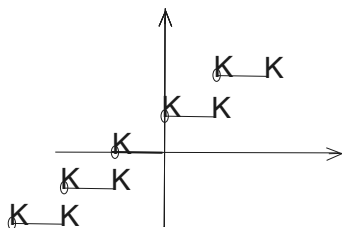
Using the vertical line test we find that it is a function. Horizontal line test tells us that the function is not one-to-one.

Example 3.18. Consider $x^2 + y^2 = 1, x \geq 0, y \geq 0$. The graph is



The vertical and horizontal line test tell us that it is a one-to-one function. The domain and range can also be obtained looking at the graph of the function. The domain is the projection of the graph on the X-axis. The range is the projection of the graph on the Y-axis. Here domain is $[0, 1]$ and the range is also $[0, 1]$.

Example 3.19. Consider the function whose graph is

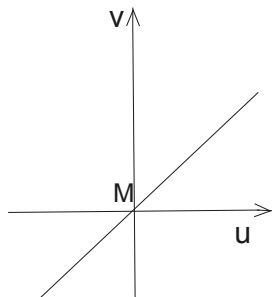


Taking projection on the X-axis and Y-axis, we find that domain = \mathbb{R} , range = \mathbb{Z} .

3.5 Some Functions on the Set of Real Numbers

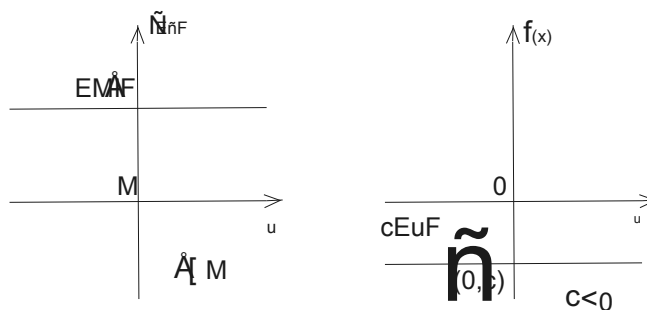
We shall now define some commonly used functions on the set of real numbers.

- (i) **Identity function:** The function from $i : \mathbb{R} \rightarrow \mathbb{R}$ defined by $i(x) = x \forall x \in \mathbb{R}$, is called the identity function. It is a bijective function. Its graph is :



Graph of $i(x) = x$

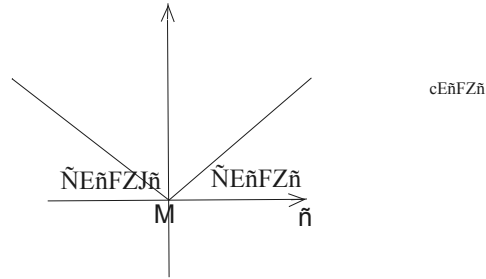
- (ii) **Constant function:** The function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = c \forall x \in \mathbb{R}$, where c is some real number, is called a constant function. It is neither one-one nor onto. Its graph is:



Graph of $f(x) = c$

- (iii) **Absolute value function:** The absolute value function $| \cdot | : \mathbb{R} \rightarrow \mathbb{R}$ is defined by
- $$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

Clearly $|x| \geq 0 \forall x \in \mathbb{R}$. Range = $[0, \infty)$, so the function is not onto. Since $|x| = |-x|$, therefore the function is not one-to-one. Hence the function is neither one-to-one nor onto. Its graph is:

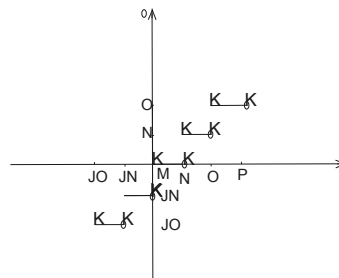


Graph of $|x|$

(iv) **Floor function or greatest integer function**

The floor function $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $\lfloor x \rfloor =$ greatest integer less than or equal to x .

The range is \mathbb{Z} , so the function is not onto. Since $\lfloor 2.3 \rfloor = 2 = \lfloor 2.5 \rfloor$, but $2.3 \neq 2.5$. Therefore the function is not one-to-one. Thus the function is neither one-to-one nor onto. Observe that for any $x \in \mathbb{R}, x - 1 < \lfloor x \rfloor \leq x$, so that $\lfloor 1.8 \rfloor = 1, \lfloor -2.3 \rfloor = -3$. The graph of the function is:



Graph of floor function $\lfloor \cdot \rfloor$

Due to the shape of the function it is also called the step function.

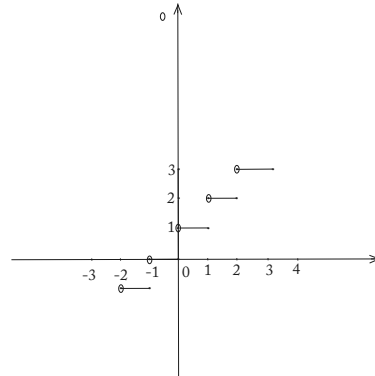
(v) **Ceiling function:** The ceiling function $\lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$\lceil x \rceil =$ least integer greater than or equal to x .

The range is \mathbb{Z} , so the function is not onto. Since $\lceil 2.3 \rceil = 3 = \lceil 2.5 \rceil$ but $2.3 \neq 2.5, \therefore$ the function is not one-to-one. Thus the function is neither one-to-one nor onto.

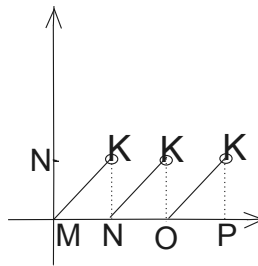
$\lceil 1.9 \rceil = 2, \lceil -1.9 \rceil = -1$

Observe that for any $x \in \mathbb{R}, x \leq \lceil x \rceil < x + 1$. The graph of the function is :



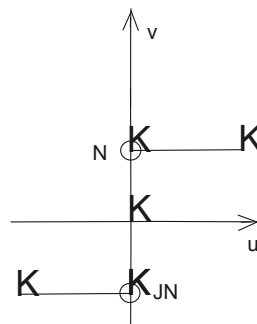
Graph of ceiling function

- (vi) **Grass function:** The grass function $f : \mathbb{R} \rightarrow \mathbb{R}$ is defined as $f(x) = x - \lfloor x \rfloor$.
The graph of the function is:



- (vii) **Signum function:** The signum function sgn is defined by $sgn : \mathbb{R} \rightarrow \mathbb{R}$
- $$sgn(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Clearly signum function is neither one-to-one nor onto. The range is $\{-1, 0, 1\}$.
The graph of the function is:



Graph of signum function $sgn(x)$.

Problem 3.1. Define $f : A \rightarrow B$ by $f(x) = x^2 + 6x - 20$. In the following cases check whether f is one-to-one and/or onto. Give reasons.

- (i) $A = \mathbb{Z}, B = \{b \in \mathbb{Z} | b \geq -29\}$.
(ii) $A = \mathbb{R}, B = \{b \in \mathbb{R} | b \geq -29\}$.

Solution: (i) Suppose $b \in B$. Then $b \in \mathbb{Z}$ and $b \geq -29$. Let, there exists $a \in A$ such that $f(a) = b$.

$$\begin{aligned} \therefore a^2 + 6a - 20 &= b \\ \Rightarrow (a + 3)^2 - 29 &= b \\ \Rightarrow (a + 3)^2 &= b + 29 \quad \dots(1) \end{aligned}$$

This has a solution $a \in \mathbb{Z}$, only when $b + 29$ is a perfect square. Thus it has no solution in \mathbb{Z} for $b = 0$, so f is not onto. Let $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$ the $a_1^2 + 6a_1 - 20 = a_2^2 + 6a_2 - 20$, so that $(a_1 + 3)^2 = (a_2 + 3)^2$

$$\begin{aligned} \Rightarrow a_1 + 3 &= \pm(a_2 + 3) \\ \Rightarrow a_1 + 3 &= a_2 + 3 \text{ or } a_1 + 3 = -(a_2 + 3) \\ \Rightarrow a_1 &= a_2 \text{ or } a_1 + a_2 = -6. \end{aligned}$$

Thus we do not always get $a_1 = a_2$. Let us choose a_1, a_2 such that $a_1 + a_2 = -6$. Let $a_1 = -2, a_2 = -4$. Then $f(a_1) = -28, f(a_2) = -28$, so we get $a_1 \neq a_2$ but $f(a_1) = f(a_2)$ so f is not one-one.

(ii) Let $b \in B$. $\therefore b \in \mathbb{R}$ such that $b \geq -29$. Then as in (i), equation(1) gives $(a + 3)^2 = b + 29$. This equation has a solution for all $b + 29 \geq 0$ and for all $a \in \mathbb{R}$.

$$\begin{aligned} \therefore a + 3 &= \pm\sqrt{(b + 29)} \\ \Rightarrow a &= -3 \pm \sqrt{(b + 29)} \end{aligned}$$

Take $a_1 = -3 + \sqrt{(b + 29)}$. Then $a_1 \in \mathbb{R}$ and $f(a_1) = b$. Hence f is onto.

As in (i) f is not one-to-one.

Problem 3.2. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = \frac{1}{2 - \cos 3x}$. Find the domain and range of f so that f is a bijective function.

Solution: Since f is one-to-one, therefore let $x_1, x_2 \in \mathbb{R}$. Then

$$\begin{aligned} f(x_1) &= f(x_2) \\ \Leftrightarrow \frac{1}{2 - \cos 3x_1} &= \frac{1}{2 - \cos 3x_2} \end{aligned}$$

$$\begin{aligned} \Leftrightarrow \cos 3x_1 &= \cos 3x_2 \\ \Leftrightarrow 3x_1 &= 2k\pi \pm 3x_2 \\ \Leftrightarrow 3(x_1 \pm x_2) &= 2k\pi \end{aligned}$$

$$\Leftrightarrow x_1 \pm x_2 = \frac{2k\pi}{3}$$

Thus choose an interval such that for x_1, x_2 in the interval, $x_1 \neq x_2$ we have $x_1 \pm x_2 \neq \frac{2k\pi}{3}$. Thus we choose the interval as $[0, \frac{\pi}{3}]$.

$$\therefore D(f) = [0, \frac{\pi}{3}].$$

Let $y \in R(f)$. Then there exists $x \in D(f)$ such that $f(x) = y$, as f is onto.

$$\begin{aligned} \therefore \frac{1}{2 - \cos 3x} &= y. \\ x &\in [0, \frac{\pi}{3}] \\ \Rightarrow 3x &\in [0, \pi] \\ \Rightarrow -1 &\leq \cos 3x \leq 1 \\ \Rightarrow 1 &\geq -\cos 3x \geq -1 \\ \Rightarrow 3 &\geq 2 - \cos 3x \geq 1 \\ \Rightarrow \frac{1}{3} &\leq \frac{1}{2 - \cos 3x} \leq 1 \end{aligned}$$

$$\begin{aligned} &\Rightarrow \frac{1}{3} \leq y \leq 1 \\ &\Rightarrow y \in \left[\frac{1}{3}, 1\right] \\ &\therefore R(f) = \left[\frac{1}{3}, 1\right]. \end{aligned}$$

Problem 3.3. Suppose A and B are two finite sets and $f : A \rightarrow B$ be a function. Then

- (i) If $o(A) > o(B)$ then f can not be one-to-one.
(ii) If $o(A) < o(B)$ then f cannot be onto.

Solution: A and B are two finite sets. Let $o(A) = n, o(B) = m$. Let $f : A \rightarrow B$ be any function.

- (i) Suppose $o(A) > o(B)$, i.e $n > m$. If f is one-one then elements of A must have distinct images. Thus $R(f)$ must have $o(A)$ elements.
Then $n = o(A) = o(R(f)) \leq o(B) = m$, which is a contradiction. Thus f cannot be one-one.
- (ii) $o(A) < o(B)$, i.e $n < m$. Since f is a function, therefore $R(f)$ can have at most $o(A)$ elements, i.e $R(f)$ can have at most n elements. Also f is onto
 $\Rightarrow R(f) = B$
 $\Rightarrow o(R(f)) = o(B) = m$
 $\Rightarrow R(f)$ has m elements
 $\Rightarrow m \leq n$
This is a contradiction to the fact that $n < m$.
Hence f cannot be onto.

Problem 3.4. Suppose A and B are two finite sets of the same order. Then any function $f : A \rightarrow B$ is onto if and only if it is one-one. The result may fail to hold if A and B are infinite sets.

Solution: Let $o(A) = o(B) = n$ (say). Let $f : A \rightarrow B$ be any function. Suppose f is onto. Then $R(f) = B$. If f is not one-one then at least two elements of A are mapped to the same element of B .
 $\therefore o(R(f)) < o(A) = o(B) \Rightarrow o(R(f)) < o(B)$
 $\Rightarrow R(f) \subset B$.

This contradicts the fact that $R(f) = B$. Hence f is one-to-one.

Now suppose that f is one-to-one. Then distinct elements of A are mapped to distinct elements of B .

$\therefore o(R(f)) = o(A) = o(B)$.
Now $R(f) \subseteq B$ and $o(R(f)) = o(B)$

$\therefore R(f) = B$.

Hence f is onto.

The result fails to hold if A and B are infinite sets. Let $f : \mathbb{Z} \rightarrow \mathbb{Z}$ be defined by $f(z) = 2z, \forall z \in \mathbb{Z}$. Then f is not onto, as $7 \in \mathbb{Z}$ does not have a preimage. Let $z \in \mathbb{Z}$ be such that $f(z) = 7$. Then $2z = 7$ which does not have a solution in \mathbb{Z} . f is one to one, for if $z_1, z_2 \in \mathbb{Z}$ such that $f(z_1) = f(z_2)$ then $2z_1 = 2z_2$, so that $z_1 = z_2$. Thus f is one-to-one but not onto.

Let $g : \mathbb{Z} \rightarrow N \cup \{0\}$ be defined by
 $g(z) = |z| \quad \forall z \in \mathbb{Z}$.

If $n \in N \cup \{0\}$, then $n \in \mathbb{Z}$ such that $g(n) = n$.

Hence g is onto. g is not one-to-one, because $2, -2 \in \mathbb{Z}$, and
 $g(2) = |2| = 2$

$$g(-2) = |-2| = 2$$

$$\therefore 2 \neq -2, \text{ but } g(2) = g(-2).$$

Thus g is onto but not one-to-one.

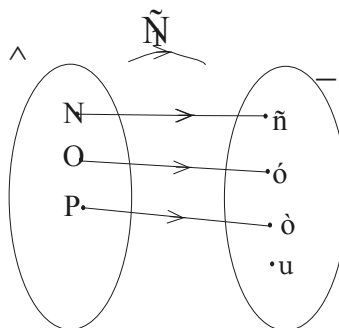
Problem 3.5. Let $A = \{1, 2, 3\}$, $B = \{u, x, y, z\}$

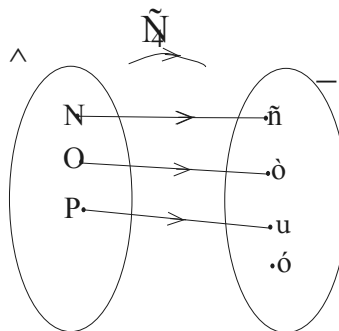
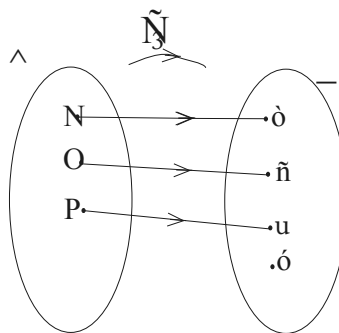
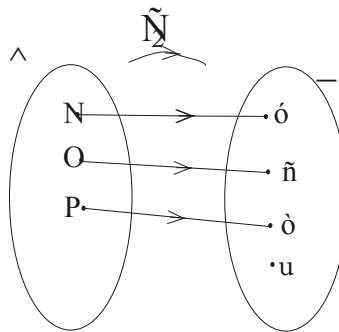
- (i) How many functions are there from A to B ?
- (ii) How many onto functions are there from A to B ?
- (iii) How many one-to-one functions are there from A to B ? List 4 of them.
- (iv) How many bijective functions are there from A to B ?

Solution: $o(A) = 3, o(B) = 4$

- (i) In a function each element of A is mapped to exactly one element of B .
 \therefore There are 4 choices for $f(1)$, 4 choices for $f(2)$ and 4 choices for $f(3)$.
 Total number of functions from A to B
 $= 4 \times 4 \times 4 = 4^3$
 $= 64$.
- (ii) If f is an onto function from A to B , then $R(f)$ has 4 elements. Since $o(A) = 3$, therefore $R(f)$ can have at most three elements, so f can not be onto. Hence there is no onto function from A to B .
- (iii) If f is a one-to-one function, then $f(1)$ has 4 choices. Since $f(2) \neq f(1)$, \therefore $f(2)$ has only 3 choices and consequently $f(3)$ has only 2 choices. Hence number of one-to-one functions from A to B
 $= 4 \times 3 \times 2$
 $= 24$

Four one-to-one functions are





(iv) Since there are no onto functions, hence there are no bijective functions.

3.6 Exercise

1. How many functions can be defined from A to B , if $o(A) = 9$ and $o(B) = 7$?
2. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by

$$f(x) = \cos[\pi^2]x + \cos[-\pi^2]x$$

Find $f(\frac{\pi}{2})$, $f(-\pi)$, $f(\pi)$, $f(\frac{\pi}{4})$.

3. Let $A = \{1, 2, 3, 4\}$ and $B = \{x, y, z\}$ and $f = \{(1, x), (2, y), (3, x), (4, z)\}$
 If $A_1 = \{1, 2\}$, $A_2 = \{2, 3, 4\}$
 Find $f(A_1 \cap A_2)$, $f(A_1) \cap f(A_2)$. What do you conclude?
4. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by
 $f(x) = \lfloor x \rfloor$.
 Find
 (i) $f^{-1}(1)$
 (ii) $f^{-1}(0.5)$, $f^{-1}(\sqrt{2})$, $f(\sqrt{2})$, $f(0.5)$, $f(-e)$
 (iii) Is f one-to-one?
 (iv) Is f onto?
5. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by
 $f(x) = |x|$.
 Find $f^{-1}(A)$ where
 (i) $A = \{1\}$
 (ii) $A = \{-1\}$
 (iii) $A = [-2, 3]$
6. Let $f : \mathbb{Z} \rightarrow \mathbb{Z}$ be defined by
 $f(n) = \text{remainder obtained on divided } n \text{ by } 5$
 Find
 (i) $R(f)$
 (ii) $f(A)$ where A is the set of all multiples of 3
 (iii) $f^{-1}(0)$, $f^{-1}(1)$
 (iv) $f^{-1}(B)$ where $B = \{3, 7\}$
7. Find the range of the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by
 $f(x) = \frac{x-1}{x^2-3x+3}$, $x \in \mathbb{R}$
8. Find the domain and range of the function
 $f(x) = \frac{1}{\sqrt{|x|-x}}$
9. Find the domain of the function
 $f(x) = \frac{1}{x} + 2^{\sin^{-1} x} + \frac{1}{\sqrt{x-2}}$
10. Find the domain and range of the following functions
 (i) $f : \mathbb{R} \rightarrow \mathbb{R}$
 $f(x) = \frac{x}{x^2+1}$
 (ii) $f : \mathbb{R} \rightarrow \mathbb{R}$
 $f(x) = \frac{1}{3-\cos 4x}$
 (iii) $f : \mathbb{R} \rightarrow \mathbb{R}$
 $f(x) = \frac{1}{\sqrt{1-x^2}}$
11. Which of the following functions defined from A to B are one-to-one?
 (i) $A = B = \mathbb{R}$, $f(x) = |x + 1|$
 (ii) $A = (0, \infty)$, $B = \mathbb{R}$
 $g(x) = x + \frac{1}{x}$
 (iii) $A = [-\infty, -4)$, $B = \mathbb{R}$
 $h(x) = x^2 + 4x - 5$
 (iv) $A = B = \mathbb{R}$
 $k(x) = e^{-x}$

12. Let $f : \mathbb{R} \setminus \{-1\} \rightarrow \mathbb{R} \setminus \{-1\}$ defined by
 $f(x) = \frac{1-x}{1+x}$.
 Prove that f is onto and one-to-one.
13. Give an example of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that
- f is one-to-one but not onto.
 - f is onto but not one-to-one.
 - f is both one-to-one and onto.
 - f is neither one-to-one nor onto.
14. If A and B are sets such that $o(A) = 8, o(B) = 10$. Then
- How many functions can be defined from A to B ?
 - How many one-to-one functions can be defined from A to B ?
 - How many onto functions can be defined from A to B ?
15. Prove that the following functions are bijective.
- $f : (-\infty, \infty) \rightarrow (0, \infty) f(x) = 2^x$
 - $f : (0, 1] \rightarrow [1, \infty)$
 $f(x) = \frac{1}{x}$
 - $f : (0, 1] \rightarrow [a, \infty)$
 $f(x) = \frac{1}{x} - 1 + a$
 - $f : (0, 1) \rightarrow (-\infty, \infty)$
 $f(x) = \frac{x - \frac{1}{2}}{x(x-1)}$
 - $f : (-\infty, \infty) \rightarrow (a, \infty)$
 $f(x) = 2^x + a$

3.7 Inverse of a Function

Consider a function defined on the sets A and B , where $A = \{a, b, c, d\}$ and $B = \{p, q, r\}$. Define

$f_1 = \{(a, p), (b, p), (c, q), (d, r)\}$. Then f_1 is onto but not one-to-one.

Also $f_1^{-1} = \{(p, a), (p, b), (q, c), (r, d)\}$.

Since p has two images a and b , therefore f_1^{-1} is not a function from B to A .

This is because f_1 is not one-one.

If we take $A = \{a, b, c, d\}, B_1 = \{p, q, r, s, t\}$

and define $f_2 = \{(a, p), (b, q), (c, r), (d, s)\}$,

then f_2 is one-to-one but not onto. Also $f_2^{-1} = \{(p, a), (q, b), (r, c), (s, d)\}$

The relation f_2^{-1} is not a function from B_1 to A as t does not have any image under f_2^{-1} . This is because f_2 is not onto. Let us now consider

$A = \{a, b, c, d\}, B_2 = \{p, q, r, s\}$ and define

$f_3 = \{(a, p), (b, q), (c, r), (d, s)\}$

Then f_3 is a bijective function from A to B_2 . Now,

$f_3^{-1} = \{(p, a), (q, b), (r, c), (s, d)\}$

f_3^{-1} is also a function from B_2 to A . In fact it is also bijective. The above illustrations lead us to believe that every function may not have an inverse, and that perhaps bijective functions have inverses.

Definition 3.6. (Inverse of a Function):

Let A and B be two sets and $f : A \rightarrow B$ be a function. If there exists a function $g : B \rightarrow A$ such that $(b, a) \in g \Leftrightarrow (a, b) \in f$, then g is called an inverse of f . We denote g by f^{-1} .

If a function has an inverse we say that it is invertible.

If f^{-1} exists then $f^{-1} = \{(b, a) \in B \times A \mid (a, b) \in f\}$ and $b = f(a) \Leftrightarrow a = f^{-1}(b)$.

We now show that $(f^{-1})^{-1} = f$.

Let f be a function defined from A to B .

Then $f^{-1} = \{(b, a) \in B \times A \mid (a, b) \in f\}$. Now, $(f^{-1})^{-1} = \{(a, b) \in A \times B \mid (b, a) \in f^{-1}\}$

$= \{(a, b) \in A \times B \mid (a, b) \in f\}$

$= f$

We would like to know that if f is a function, under what conditions does f^{-1} exist?

In the above examples we saw that f_1 is onto, but not one-to-one and f_1^{-1} does not exist. f_2 is not onto but it is one-to-one and f_2^{-1} does not exist. f_3 is onto and one-to-one and f_3^{-1} exists.

The above examples suggest the following:

Theorem 3.3. A function $f : A \rightarrow B$ has an inverse $f^{-1} : B \rightarrow A$ if and only if f is one-to-one and onto.

Proof: Suppose that f^{-1} exists. We shall prove that f is a bijective function. Since $f^{-1} : B \rightarrow A$ is a function.

$\therefore f^{-1}(b) = a \Leftrightarrow f(a) = b$

f is one-to-one.

Let $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$. Let $f(a_1) = f(a_2) = b$ (say)

Then $f(a_1) = b \Rightarrow f^{-1}(b) = a_1$ and $f(a_2) = b \Rightarrow f^{-1}(b) = a_2$

Thus $a_1 = f^{-1}(b) = a_2$, so that f is one-to-one.

f is onto.

Let $b \in B$. Then $f^{-1}(b) \in A$ (as f^{-1} is a function). Let $f^{-1}(b) = a \in A$. Thus $f(a) = b$, so that f is onto.

Conversely, let f be a bijective function. Let $b \in B$. Since f is onto, there exists $a \in A$ such that $f(a) = b$. Since f is one-to-one, the element $a \in A$ is unique. Hence for each $b \in B$, there exists a unique $a \in A$ such that $f^{-1}(b) = a$. Hence f^{-1} is a function from B to A . \square

Example 3.20. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x + 2 \forall x \in \mathbb{R}$. Then f is a bijective function. Hence f has an inverse. If $f^{-1}(x) = y$, then $f(y) = x$

$\Rightarrow y + 2 = x$

$\Rightarrow y = x - 2$

$\therefore f^{-1}(x) = x - 2$

We now outline the steps involved in finding an inverse of a function if it exists.

Steps involved in finding f^{-1}

Given $f : A \rightarrow B$ is a function

Step 1 Prove that f is a bijective function. Then f^{-1} exists.

Step 2 By step 1

$f^{-1} : B \rightarrow A$

we are required to find $f^{-1}(b)$, for $b \in B$. If $f^{-1}(b) = a$ (1)

Then $b = f(a)$

Solve for a in terms of b . The solution must be in A . Substitute the solution obtained in (1). Thus we get the function f^{-1} .

Example 3.21. Define $f : \mathbb{R}^- \rightarrow \mathbb{R}^+$ by $f(x) = x^2$. To find the inverse function of f .

Step1 We prove that f is a bijective function. Let $x \in \mathbb{R}^+$. Then, $-\sqrt{x} \in \mathbb{R}^-$ such that

$$f(-\sqrt{x}) = (-\sqrt{x})^2 = x$$

Hence f is onto.

Let $x, y \in \mathbb{R}^-$ such that $f(x) = f(y)$. Then, $x^2 = y^2$

$$\Rightarrow x = \pm y$$

Since $x, y \in \mathbb{R}^-$

$\therefore x = -y$ is not possible so that $x = y$. $\therefore f$ is one-to-one.

Thus f is a bijective function. $\therefore f^{-1}$ exists and

$$f^{-1} : \mathbb{R}^+ \rightarrow \mathbb{R}^-$$

Step2 We shall now find the rule for defining f^{-1} .

Let $x \in \mathbb{R}^+$.

Then $f^{-1}(x) \in \mathbb{R}^-$. Let $f^{-1}(x) = y$.

Then $x = f(y)$

$$\Rightarrow x = y^2$$

$$\Rightarrow y = \pm\sqrt{x}$$

The solution of this equation in \mathbb{R}^- is $y = -\sqrt{x}$

$$\therefore f^{-1}(x) = -\sqrt{x}.$$

Example 3.22. Define $f : \mathbb{R} \rightarrow \mathbb{R}^+$ by $f(x) = x^2$. Then f is not one-to-one.

$\because 2, -2 \in \mathbb{R}$ but $f(2) = 4 = f(-2)$. Then $2 \neq -2$, but $f(2) = f(-2)$. Hence f is not bijective and so does not have an inverse.

Example 3.23. Define $f : \mathbb{R}^- \rightarrow \mathbb{R}$ by $f(x) = x^2$.

Then $f(x) \geq 0 \quad \forall x \in \mathbb{R}^-$

$-4 \in \mathbb{R}$ (the codomain) and there does not exist any $x \in \mathbb{R}^-$ such that $f(x) = -4$.

Hence f is not onto, so that f is not a bijective function.

$\therefore f^{-1}$ does not exist.

Theorem 3.4. Let $f : A \rightarrow B$ be any function. Then f is bijective $\Leftrightarrow f^{-1}$ is bijective.

Proof: Let f be bijective. Then f^{-1} exists and $f^{-1} : B \rightarrow A$. Let $b_1, b_2 \in B$ such that $f^{-1}(b_1) = f^{-1}(b_2)$. If $f^{-1}(b_1) = f^{-1}(b_2) = x$, then $f(x) = b_1$ and $f(x) = b_2$, so that $b_1 = b_2$. Hence f^{-1} is one-to-one.

Let $a \in A$ and $f(a) = b$ (as f is a function), so that $f^{-1}(b) = a$. Thus f^{-1} is onto. This proves that f^{-1} is a bijective function.

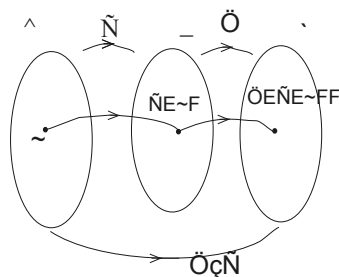
Conversely, let f^{-1} be a bijective function. If $g = f^{-1}$, then g is a bijective function so that g^{-1} exists and is bijective.

However, $g^{-1} = (f^{-1})^{-1} = f$, i.e f is a bijective function. \square

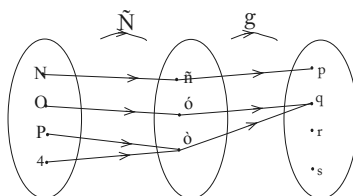
3.8 Composition of Functions

Having defined functions, we would like to combine them by applying one function after the other so as to get another function. Thus we have the following definition.

Definition 3.7. If $f : A \rightarrow B$ and $g : B \rightarrow C$ are two functions, then the composite of f by g is the function $g \circ f : A \rightarrow C$ defined by $(g \circ f)(a) = g(f(a)) \quad \forall a \in A$.



Example 3.24. Let $A = \{1, 2, 3, 4\}$, $B = \{x, y, z\}$, $C = \{p, q, r, s\}$, and



$f = \{(1, x), (2, y), (3, z), (4, z)\}$
 $g = \{(x, p), (y, q), (z, q)\}$
 Then $f : A \rightarrow B, g : B \rightarrow C$, so that
 $g \circ f : A \rightarrow C$ such that $(g \circ f)(a) = g(f(a))$, for $a \in A$.
 Hence $(g \circ f)(1) = g(f(1)) = g(x) = p$.
 $(g \circ f)(2) = q$
 $(g \circ f)(3) = q$
 $(g \circ f)(4) = q$.

Example 3.25. Define $f : \mathbb{R}^* \rightarrow \mathbb{R}$ by $f(x) = 1/x$, and $g : \mathbb{R} \rightarrow \mathbb{R}$ by $g(x) = x^2$. Then $g \circ f : \mathbb{R}^* \rightarrow \mathbb{R}$ and $(g \circ f)(x) = g(f(x))$

$$\begin{aligned}
&= g(1/x) \\
&= \frac{1}{x^2} \\
&\quad \text{Here } R(g) = [0, \infty) \not\subseteq \mathbb{R}^* = D(f) \\
&\quad \therefore fog \text{ is not defined.}
\end{aligned}$$

Example 3.26. Define functions f and g as follows:

$$f : \mathbb{R}^* \rightarrow \mathbb{R}^*$$

$$f(x) = 1/x$$

$$g : \mathbb{R}^* \rightarrow \mathbb{R}^*$$

$$g(x) = x^2 + 1$$

$$\text{Then } fog : \mathbb{R}^* \rightarrow \mathbb{R}^*$$

$$\text{and } (fog)(x) = f(g(x)) = f(x^2 + 1) = \frac{1}{x^2 + 1}$$

$$\text{Also } gof : \mathbb{R}^* \rightarrow \mathbb{R}^*$$

$$\text{and } (gof)(x) = g(f(x)) = g(1/x) = \frac{1}{x^2} + 1$$

Thus we see that

$$(fog)(1) = 1/2$$

$$(gof)(1) = 1 + 1 = 2$$

so that $(fog)(1) \neq (gof)(1)$

Hence $fog \neq gof$.

This shows that the composition of functions is not commutative. In fact, the composition of functions is associative. Let A and B be two sets and $f : A \rightarrow B$ and $g : B \rightarrow A$ be two functions. If $gof = i_A$, the identity function on A , then g is called a left inverse of f . If $fog = i_B$, then g is called a right inverse of f . If g is a right inverse as well as a left inverse of f , then g is called an inverse of f . Though, inverse of a function is unique but a right (left) inverse is not unique. In fact a right (left) inverse may exist but inverse may not exist. The following theorem gives the conditions for a function to have a left (right) inverse.

Theorem 3.5. Let A and B be two sets and $f : A \rightarrow B$ be a function. Then

(i) f is onto if and only if f has a right inverse.

(ii) f is one-to-one if and only if f has a left inverse.

Proof:

(i) Let f be onto. Then for each $b \in B$, there exists $a \in A$ such that $f(a) = b$.

Also $R(f) = B$.

Define

$$g : B \rightarrow A$$

by $g(b) = a$ if $f(a) = b$. Since $R(g) \subseteq A = D(f)$, therefore fog is defined, and

$$fog : B \rightarrow B$$

If $b \in B$, then

$$\begin{aligned}
(fog)(b) &= f(g(b)) \\
&= f(a), \quad \text{where } g(b) = a \text{ if } f(a) = b. \\
&= b.
\end{aligned}$$

$$\therefore (fog)(b) = b.$$

Hence $fog = i_B$, so that g is a right inverse of f .

Conversely, let f have a right inverse h . Then

$$h : B \rightarrow A$$

such that $f \circ h = i_B$. We prove that f is onto. Let $b \in B$. Then

$$\begin{aligned} (f \circ h)(b) &= i_B(b) \\ \Rightarrow f(h(b)) &= b. \\ \Rightarrow f(a) &= b, \text{ where } a = h(b) \in A. \\ \Rightarrow f &\text{ is onto.} \end{aligned}$$

(ii) Let f be one-to-one. We prove that f has left inverse. If $g : B \rightarrow A$ is a left inverse then $g \circ f$ must be defined, so that $R(f) \subseteq D(g)$. We take $D(g) = R(f)$.

Since f is one-to-one, \therefore if $a_1, a_2 \in A$ such that $a_1 \neq a_2$ then $f(a_1) \neq f(a_2)$. So, for every $b \in R(f)$, there exists unique $a \in A$ such that $f(a) = b$. Define

$$\begin{aligned} g : R(f) &\rightarrow A \\ g(b) &= a, \text{ if } f(a) = b. \\ \text{Then } g \circ f : A &\rightarrow A \\ (g \circ f)(a) &= g(f(a)) \text{ for } a \in A \\ &= g(b), \text{ where } f(a) = b. \\ &= a, \text{ by definition of } g. \\ \therefore (g \circ f)(a) &= a \forall a \in A. \\ \text{So that } g \circ f &= i_A. \end{aligned}$$

Conversely, let f have a left inverse say h . Let $h : B \rightarrow A$. For $h \circ f$ to be defined, we must have $R(f) \subseteq B$. Also

$$\begin{aligned} h \circ f &= i_A. \\ \text{Let } a_1, a_2 \in A &\text{ such that } f(a_1) = f(a_2). \text{ Then} \\ h(f(a_1)) &= h(f(a_2)) \\ (h \circ f)(a_1) &= (h \circ f)(a_2) \\ \Rightarrow i_A(a_1) &= i_A(a_2). \\ \Rightarrow a_1 &= a_2 \\ \Rightarrow f &\text{ is one-to-one.} \end{aligned}$$

□

Theorem 3.6. Let A and B be two sets and $f : A \rightarrow B$ be a function. Then f is bijective \Leftrightarrow there exists $g : B \rightarrow A$ such that $f \circ g = i_B$, $g \circ f = i_A$.

Proof: Let f be bijective. By Theorem 3.5, there exists a right inverse g of f and a left inverse h of f . Thus

$$f \circ g = i_B, \quad h \circ f = i_A$$

Since composition of functions is associative,

$$\begin{aligned} \therefore h \circ (f \circ g) &= (h \circ f) \circ g \\ \Rightarrow h \circ i_B &= i_A \circ g \\ \Rightarrow h &= g. \end{aligned}$$

Thus there exists a function $g : B \rightarrow A$ such that $f \circ g = i_B$, $g \circ f = i_A$.

Conversely, let the conditions hold. By Theorem 3.5, f is onto and injective so that f is bijective. □

Theorem 3.7. Let $f : A \rightarrow B$ be any function. Then, a function $g : B \rightarrow A$ is an inverse of f if and only if $f \circ g = i_B$ and $g \circ f = i_A$, where i_A, i_B are identity functions on A and B respectively.

Proof: Suppose g is an inverse of f . Then $g = f^{-1}$. If $a \in A, b \in B$, then

$$f(a) = b$$

$$\Leftrightarrow g(b) = a$$

We see that $f \circ g : B \rightarrow B$ and $g \circ f : A \rightarrow A$. For any $b \in B$

$$\begin{aligned} (f \circ g)(b) &= f(g(b)) \\ &= f(a) \\ &= b \end{aligned}$$

$$\therefore (f \circ g)(b) = b \quad \forall b \in B$$

so that $f \circ g = i_B$. Similarly $g \circ f = i_A$. Hence proved.

Conversely, suppose that the conditions hold. To prove that g is an inverse of f , we must prove that for $a \in A, b \in B$, $f(a) = b \Leftrightarrow g(b) = a$. Let $a \in A, b \in B$ such that $f(a) = b$ then $g(f(a)) = g(b)$

$$\begin{aligned} \Rightarrow (g \circ f)(a) &= g(b) \\ \Rightarrow i_A(a) &= g(b) \\ \Rightarrow a &= g(b) \end{aligned}$$

Conversely let $a \in A, b \in B$ such that $g(b) = a$.

$$\begin{aligned} \text{Then } f(g(b)) &= f(a) \\ \Rightarrow (f \circ g)(b) &= f(a) \\ \Rightarrow i_B(b) &= f(a) \\ \Rightarrow b &= f(a) \end{aligned}$$

Thus we have proved that

$$f(a) = b$$

$$\Leftrightarrow g(b) = a$$

so that g is an inverse of f . \square

We shall now prove that if a function has an inverse, it must be unique. Thus we will say the inverse of a function.

Theorem 3.8. *Let $f : A \rightarrow B$ be a function. Then f is invertible $\Leftrightarrow f$ is bijective.*

Proof: Let f be invertible. By Theorem 3.7, there exists a function $g : B \rightarrow A$ such that $g \circ f = i_A$, $f \circ g = i_B$. By Theorem 3.6, f is bijective.

Conversely, let f be bijective. By Theorem 3.6, there exists a function $g : B \rightarrow A$, such that $g \circ f = i_A$, $f \circ g = i_B$. By Theorem 3.7, f has an inverse, so that f is invertible. \square

Theorem 3.9. *An invertible function has a unique inverse.*

Proof: Let f be an invertible function from A to B and let g and h be two inverses of f . $g : B \rightarrow A, h : B \rightarrow A$ such that

Then $f \circ g = i_B, g \circ f = i_A$ and $f \circ h = i_B, h \circ f = i_A$. If $b \in B$, then $(f \circ h) = i_B$

$$\begin{aligned} \Rightarrow (f \circ h)(b) &= i_B(b) \\ \Rightarrow g \circ (f \circ h)(b) &= (g \circ i_B)(b) \\ \Rightarrow ((g \circ f) \circ h)(b) &= g(b) \\ \Rightarrow (i_A \circ h)(b) &= g(b) \\ \Rightarrow h(b) &= g(b) \end{aligned}$$

Since this holds for all $b \in B$.

$$\therefore h = g.$$

Hence inverse of a function is unique. \square

Theorem 3.10. *The composite of two bijective functions is bijective.*

Proof: Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be two bijective functions. Then $gof : A \rightarrow C$. We prove that gof is bijective.

Step1 To prove that gof is onto. Let $c \in C$. Since g is onto $\therefore \exists b \in B$ such that $g(b) = c$.

Since f is onto $\therefore \exists a \in A$ such that $f(a) = b$

Now $g(f(a)) = g(b) = c$

$\Rightarrow (gof)(a) = c$

$\Rightarrow gof$ is onto.

Step2 To prove that gof is one-to-one.

Let $a_1, a_2 \in A$ such that

$(gof)(a_1) = (gof)(a_2)$

$\therefore g(f(a_1)) = g(f(a_2))$

$\Rightarrow f(a_1) = f(a_2) \quad \because g$ is one-to-one and $f(a_1), f(a_2) \in B$

$\Rightarrow a_1 = a_2 \quad \because f$ is one-to-one

Hence gof is one-to-one.

Step3 Steps 1 and 2 prove that gof is a bijective function. \square

Theorem 3.11. *If f and g are invertible functions, so is gof . Moreover $(gof)^{-1} = f^{-1}og^{-1}$.*

Proof: Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be invertible functions. Then

$gof : A \rightarrow C$

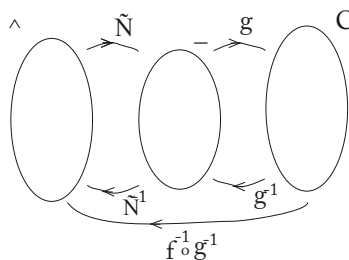
f, g are invertible functions.

$\Rightarrow f, g$ are bijective functions... (By Theorem 3.8)

$\Rightarrow f^{-1}, g^{-1}$ exist and are bijective... (By Theorem 3.9)

Then $f^{-1} : B \rightarrow A$ and $g^{-1} : C \rightarrow B$, so that

$f^{-1}og^{-1} : C \rightarrow A$.



Now f, g are bijective functions.

$\Rightarrow gof$ is a bijective function.

$\Rightarrow (gof)^{-1}$ exists and is bijective function from C to A .

We see that $f^{-1}og^{-1}$ is also a mapping from C to A .

To prove that $f^{-1}og^{-1} = (gof)^{-1}$ we must prove that their action on every element of C is same. Let $c \in C$. Then

$f^{-1}og^{-1}(c) = f^{-1}(g^{-1}(c))$

$= f^{-1}(b)$, where $g^{-1}(c) = b \in B$

$= a$, where $f^{-1}(b) = a \in A$

Now $g^{-1}(c) = b \Leftrightarrow g(b) = c$

$f^{-1}(b) = a \Leftrightarrow f(a) = b.$
 Thus $g(f(a)) = g(b) = c$
 $\Rightarrow (gof)(a) = c$
 $\Rightarrow (gof)^{-1}(c) = a$
 Thus $(gof)^{-1}(c) = (f^{-1}og^{-1})(c) \forall c \in C$
 so that $(gof)^{-1} = f^{-1}og^{-1}.$ □

3.9 Solved Problems

Problem 3.6. Let $A = \{1, 2, 3\}, B = \{a, b\}$

Let $f = \{(1, a), (2, b), (3, a)\}$

and $g = \{(a, 1), (b, 2)\}$

Find gof and fog . Is $g = f^{-1}$?

Solution: Since f is a mapping from A to B and g is a mapping from B to A .

$\therefore fog$ and gof are defined and $fog : B \rightarrow B, gof : A \rightarrow A$. Now, $(fog)(a) =$

$$f(g(a)) = f(1) = a$$

$$(fog)(b) = f(g(b)) = f(2) = b$$

$\therefore fog = i_B$, the identity mapping on B .

$$(gof)(1) = g(f(1)) = 1$$

$$(gof)(2) = 2$$

$$(gof)(3) = 1$$

Thus $gof \neq i_A$

Hence $g \neq f^{-1}$. In fact f does not have an inverse as $f : A \rightarrow B$ is not one-one.

Remark 3.3. If $f : A \rightarrow B$ and $g : B \rightarrow A$, then $fog = i_B$ is not sufficient to ensure that f is invertible. We must also check that $gof = i_A$. Also more than one g can be found such that $fog = i_B$ and $gof \neq i_A$.

So f has two right inverses, h and g . In the above question, let $h = \{(a, 3), (b, 2)\}$. Then $(foh) = i_B$ but $h \neq g$.

Problem 3.7. Define $f : \mathbb{Z} \rightarrow \mathbb{N}$ by $f(x) = \begin{cases} 2|x| & \text{if } x < 0 \\ 2x + 1 & \text{if } x \geq 0 \end{cases}$

Show that f has an inverse and find f^{-1} and hence find $f^{-1}(3686), f^{-1}(231)$.

Solution: To prove that f is one-to-one.

Let $x, y \in \mathbb{Z}$ such that $f(x) = f(y)$. Three cases arise:

Case 1. $x, y \geq 0$

Then $f(x) = f(y)$

$$\Rightarrow 2x + 1 = 2y + 1$$

$$\Rightarrow x = y$$

Case 2. $x, y < 0$

$f(x) = f(y)$

$$\Rightarrow 2|x| = 2|y|$$

$$\Rightarrow -x = -y$$

$$\Rightarrow x = y$$

Case 3. One of them is ≥ 0 and the other is < 0

Without any loss of generality we can take $x \geq 0, y < 0$

Then $f(x) = f(y)$

$\Rightarrow 2x + 1 = 2|y|$
 $\Rightarrow 2x + 1 = -2y$
 $\Rightarrow 2(x + y) = -1$
 $\Rightarrow x + y = -1/2$
 which is not possible in \mathbb{Z} .
 Thus in all cases $f(x) = f(y)$
 $\Rightarrow x = y$

Hence f is one-to-one.

To show that f is onto.

Let $y \in \mathbb{N}$. Then y is either odd or even. Suppose that y is odd.

$\therefore y = 2z + 1$ for some $z \in \mathbb{N} \cup \{0\}$

$f(z) = 2z + 1$
 $= y$, where $z = (y - 1)/2$

Let y be even.

$\therefore y = 2z$ for some $z \in \mathbb{N}$

Then $-z \in \mathbb{Z}$, $-z < 0$ and

$f(-z) = 2|-z| = 2z = y$

Combining the two we get

$f(\frac{-y}{2}) = y$ if y is even, and

$f(\frac{y-1}{2}) = y$ if y is odd

Thus f is onto.

To find f^{-1} . Since f is a one-to-one and onto function, therefore f is invertible so that f^{-1} exists and $f^{-1} : \mathbb{N} \rightarrow \mathbb{Z}$.

We now obtain a rule to define f^{-1} .

For $x \in \mathbb{N}$, let $f^{-1}(x) = y$

Then $x = f(y)$

$= \begin{cases} -2y & \text{if } y < 0 \\ 2y + 1 & \text{if } y \geq 0 \end{cases}$

Solving for y in terms of x , we get

$\therefore f^{-1}(x) = \begin{cases} \frac{-x}{2} & \text{if } x \text{ is even} \\ \frac{x-1}{2} & \text{if } x \text{ is odd} \end{cases}$
 $f^{-1}(3686) = \frac{-3686}{2} \quad \because 3686 \text{ is even.}$

$= -1843$

$f^{-1}(231) = \frac{231-1}{2} \quad \because 231 \text{ is odd.}$

$= 115$

Problem 3.8. Let A be a subset of \mathbb{R} and suppose $f : A \rightarrow A$ is a function with the property that

$f^{-1}(x) = \frac{1}{f(x)} \quad \forall x \in A$. Show that

(i) $0 \notin A$.

(ii) $f^4 = i_A$, the identity function on A .

Solution:

(i) Suppose that $0 \in A$. Since f is onto there exists some $a \in A$ such that

$f(a) = 0$.

$f^{-1}(a) = \frac{1}{f(a)}$, which is not defined.

Hence our assumption is wrong, so that $0 \notin A$.

(ii) Let $x \in A$. Then

$f^{-1}(x) = \frac{1}{f(x)}$

$$f(f^{-1}(x)) = f\left(\frac{1}{f(x)}\right)$$

$$\text{i.e. } (f \circ f^{-1})(x) = f\left(\frac{1}{f(x)}\right)$$

$$\text{i.e. } f\left(\frac{1}{f(x)}\right) = x$$

Applying f on both sides

$$f^2\left(\frac{1}{f(x)}\right) = f(x), \text{ for all } x \in A \quad \dots(1)$$

Let $y \in A$. since f is onto, means that there exists some $z \in A$ such that

$$f(z) = y$$

$$\text{Then } f^2\left(\frac{1}{f(z)}\right) = f(z) \quad \text{using (1)}$$

$$\text{i.e. } f^2\left(\frac{1}{y}\right) = y \quad \forall y \in A \quad \dots (2)$$

$$\text{Now } f^4(x) = f^2(f^2(x))$$

$$= f^2\left(\frac{1}{x}\right) \quad \text{using (2)}$$

$$= x \quad \text{using (2)}$$

$$\therefore f^4(x) = x \quad \forall x \in A$$

Thus f^4 is the identity mapping on A .

Problem 3.9. If $f : A \rightarrow B$ and $g : B \rightarrow C$, are functions, prove that

(i) If f is onto and $g \circ f$ is one-to-one then g is one-to-one.

(ii) If g is one-to-one and $g \circ f$ is onto, then f is onto.

Solution: Clearly $g \circ f : A \rightarrow C$

(i) Let $x_1, x_2 \in B$ such that

$$g(x_1) = g(x_2) \quad (1)$$

Since f is onto

\therefore there exist $x'_1, x'_2 \in A$ such that

$$f(x'_1) = x_1 \text{ and } f(x'_2) = x_2 \quad (2)$$

(1) and (2) \Rightarrow

$$g(f(x'_1)) = g(f(x'_2))$$

$$\Rightarrow (g \circ f)(x'_1) = (g \circ f)(x'_2)$$

$$\Rightarrow x'_1 = x'_2 \quad \because g \circ f \text{ is one-to-one}$$

$$\Rightarrow f(x'_1) = f(x'_2)$$

$$\Rightarrow x_1 = x_2$$

Hence g is one-to-one.

(ii) Let $b \in B$. Then $g(b) \in C$. Since $g \circ f$ is onto C and $g(b) \in C$

$$\therefore \exists \text{ some } a \in A \text{ such that } (g \circ f)(a) = g(b)$$

$$\Rightarrow g(f(a)) = g(b)$$

$$\Rightarrow f(a) = b \quad \text{since } g \text{ is one-to-one}$$

Thus f is onto.

Problem 3.10. If f and g are two functions such that $g \circ f$ is one-to-one then prove that

(i) f must be one-to-one.

(ii) g may not be one-to-one.

Solution: Let $f : A \rightarrow B$, $g : B \rightarrow C$ be two functions. Then $g \circ f : A \rightarrow C$.

Suppose that $g \circ f$ is a one-to-one function .

(i) Let $a_1, a_2 \in A$ such that

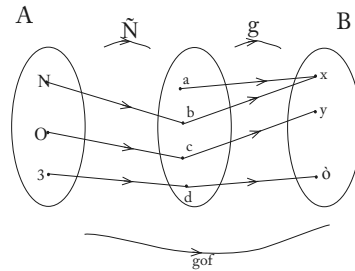
$$f(a_1) = f(a_2) \quad (1)$$

Now, $f(a_1), f(a_2) \in B$ so that

$$g(f(a_1)), g(f(a_2)) \in C$$

$(1) \Rightarrow g(f(a_1)) = g(f(a_2))$
 $\Rightarrow (g \circ f)(a_1) = (g \circ f)(a_2)$
 $\Rightarrow a_1 = a_2$ since $g \circ f$ is one-to-one
 Thus f is one-to-one.

(ii) The functions f and g are defined by the arrow diagram.



Let $g \circ f$ is one-to-one, but g is not one-to-one.

Problem 3.11. A function is defined on the set of real numbers as follows:
 $f: \mathbb{R} \rightarrow (1, \infty)$, $f(x) = 3^{2x} + 1$
 Does f^{-1} exist? If yes, find it.

Solution: To prove that f^{-1} exist, we shall prove that f is a bijective function.
 To prove that f is one-to-one. Let $x_1, x_2 \in \mathbb{R}$ such that
 $f(x_1) = f(x_2)$
 Then $3^{2x_1} + 1 = 3^{2x_2} + 1$
 $\Rightarrow 3^{2x_1} = 3^{2x_2}$
 $\Rightarrow 2x_1 = 2x_2$
 $\Rightarrow x_1 = x_2$

Hence f is one-to-one.

To prove that f is onto

Let $y \in (1, \infty)$. In order to prove that f is onto, we need to solve

$$f(x) = y \quad \dots (1)$$

for $x \in \mathbb{R}$

Thus $3^{2x} + 1 = y$

$\Rightarrow 3^{2x} = y - 1$

Taking logarithm to the base 3, we get

$$x = \frac{1}{2} \log_3(y - 1) \quad \dots (2)$$

Since $y \in (1, \infty) \therefore y > 1$

i.e $y - 1 > 0$ so that $\log_3(y - 1)$ is defined.

Hence $x \in \mathbb{R}$.

Thus f is a bijective function, and therefore f^{-1} exists.

From (1) and (2) it follows that

$$f^{-1}(y) = \frac{1}{2} \log_3(y - 1), \quad \text{where } y \in (1, \infty)$$

Thus $f^{-1} : (1, \infty) \rightarrow \mathbb{R}$ defined by $f^{-1}(x) = \frac{1}{2} \log_3(x - 1)$.

3.10 Exercise

- Define $f : \mathbb{Z} \rightarrow \mathbb{N}$ by

$$f(x) = \begin{cases} 2|x| & x < 0, \\ 2x + 1 & x \geq 0 \end{cases}$$
 Find $f \circ f$.
- Define functions on \mathbb{R} as

$$f(x) = \log\left(\frac{1+x}{1-x}\right)$$

$$g(x) = \frac{3x+x^3}{1+3x^2}$$
 Find the natural domains of f and g . Show that $(f \circ g)(x) = 3f(x)$. Hence deduce that $f \circ g$ and f have the same domain.
- Let $A = \{1, 2, 3, 4, 5\}$, $B = \{1, 3, 5, 7, 9\}$
 Define functions from A to B as follows:
 - $f = \{(1, 9), (2, 7), (3, 5), (4, 3), (5, 1)\}$
 - $g = \{(1, 3), (2, 5), (3, 7), (4, 5), (5, 9)\}$
 Find f^{-1} and g^{-1} , in case they exist.
- Let $f : A \rightarrow B$. Give example of the following:
 - f has a right inverse but not a left inverse.
 - f has a left inverse but not a right inverse.
 - f has neither a right inverse nor a left inverse.
 - f has a right as well as a left inverse.
- Define the following functions

$$f : [1, 3] \rightarrow \mathbb{R}$$
 by $f(x) = 2x$

$$g : \mathbb{R}^* \rightarrow \mathbb{R}^*$$
 by $g(x) = \frac{1}{x}$

$$h : \mathbb{R} \rightarrow \mathbb{R}$$
 by $h(x) = 1 + 3x$
 Find
 - $h \circ g \circ f$.
 - $R(h \circ g \circ f)$.
- The functions f and g are defined as follows:

$$f : (1, 2) \rightarrow \mathbb{R}$$
 by $f(x) = x - [x]$

$$g : \mathbb{R}^* \rightarrow \mathbb{R}^*$$
 by $g(x) = \frac{1}{x}$
 - Find $g \circ f$.
 - Is $g \circ f$ bijective.
 - If answer to (ii) is yes, find $(g \circ f)^{-1}$.
- If f and g are two functions such that $g \circ f$ is onto then prove that
 - g must be onto.
 - f may not be onto.
- Show that the composition of two onto functions is an onto function.

9. Show that the composition of two one-to-one functions is a one-to-one function.
10. Define $f : \mathbb{R}^+ \rightarrow \mathbb{R}$
 by $f(x) = 1 - \frac{1}{x+1}$
 $g : \mathbb{R}^+ \rightarrow \mathbb{R}$
 by $g(x) = \frac{1}{x}$
 $h : \mathbb{R}^+ \rightarrow \mathbb{R}$
 by $h(x) = x + 1$
 (i) Find range of f, g and h .
 (ii) Show that $gof, fog, hogof$, and $fogoh$ are all defined.
 (iii) Show that f, g, gof are all invertible functions.
 (iv) Verify that $(gof)^{-1} = f^{-1}og^{-1}$.
11. Define functions f, g, h as follows:
 $f : \mathbb{R}^* \rightarrow \mathbb{R}$
 by $f(x) = 2x$
 $g : \mathbb{R}^* \rightarrow \mathbb{R}^*$
 by $g(x) = 1/x$
 $h : \mathbb{R} \rightarrow \mathbb{R}$
 by $h(x) = 1 + 3x$
 Find
 (i) Range of $hogof$.
 (ii) If $(hogof)^{-1}$ exists, find it.
12. If $f : (-\infty, 0) \rightarrow \mathbb{R}$ is defined by
 $f(x) = \frac{1}{\sqrt{|x|-x}}$
 Find f^{-1}
13. If $f : [0, 3] \rightarrow \mathbb{R}$ is defined as
 $f(x) = \begin{cases} 1+x & 0 \leq x \leq 2, \\ 3-x & 2 < x \leq 3 \end{cases}$
 Find
 (i) Range f .
 (ii) $f \circ f$.
 (iii) Suitable inverse.
14. Show that the function $f : [2, \infty) \rightarrow [1, \infty)$ defined by
 $f(x) = x^2 - 4x + 5$ is a bijection. Find f^{-1} .
15. Let $f : [0, 1] \rightarrow [0, 1]$ be defined by
 $f(x) = \begin{cases} x & \text{if } x \text{ is rational,} \\ 1-x & \text{if } x \text{ is not rational} \end{cases}$
 Show that
 $f(x) = f^{-1}(x), \forall x \in [0, 1]$.

3.11 Cardinality of a Set

Two finite sets can be compared in size by counting the number of elements they have. But how do we compare the size of two infinite sets? The set of natural numbers, integers, rational numbers, real numbers and complex numbers

are all infinite sets such that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$. Does it mean that \mathbb{Z} is larger in size than \mathbb{N} , \mathbb{Q} is larger in size than \mathbb{Z} and so on? We may also think that since all are infinite sets, so they are of the same size. Certainly not!

We now discuss how the sizes of infinite sets can be compared .

Definition 3.8. (One-to-one correspondence):

Let A and B be two sets. If $f : A \rightarrow B$ is a bijective function then we say that there is a one-to-one correspondence between A and B . Equivalently, we say that A and B are in one-to-one correspondence.

Example 3.27.

- (i) $A = \{a, b, c, d, \dots, z\}$
 $B = \{1, 2, 3, 4, \dots, 26\}$
 Define $f : A \rightarrow B$
 by $f(a) = 1, f(b) = 2, f(c) = 3, \dots, f(z) = 26$
 Then $a \rightarrow 1, b \rightarrow 2, \dots, z \rightarrow 26$
 is a one-to-one correspondence between A and B .
- (ii) Define $f : 2\mathbb{Z} \rightarrow 4\mathbb{Z}$
 by $f(2z) = 4z \forall z \in \mathbb{Z}$
 Then f is a bijective function so that $2\mathbb{Z}$ and $4\mathbb{Z}$ are in one-to-one correspondence.

Definition 3.9. (Finite set):

A set is finite if it is

- (i) either the null set, or
 (ii) in one-to-one correspondence with the set $\{1, 2, 3, \dots, n\}$ for some natural number n .

If a set is not finite, then it is said to be infinite.

Definition 3.10. (Cardinality of a finite set):

If A is a finite, non empty set and it is in one-to-one correspondence with $\{1, 2, 3, \dots, n\}$, we define the cardinality of A to be n . The cardinality of the null set is defined to be zero. The cardinality of A is denoted by $|A|$.

Example 3.28.

- (i) Let $A = \{n \in \mathbb{N} \mid n \leq 15\}$
 Then $|A| = 15$.
- (ii) $B = \{x \in \mathbb{Z} \mid -10 \leq x < 5\}$
 Then $B = \{-10, -9, \dots, 4\}$
 $|B| = 15$.
- (iii) Let $C =$ set of all persons living on a moon of Jupiter
 Then $C = \phi$ so that $|C| = 0$.

Definition 3.11. (Equipollent sets):

Two sets A and B are said to have the same cardinality or are equipollent if and only if there exists a one-to-one correspondence between A and B . We write $|A| = |B|$.

We write $A \sim B$. It can be verified that equipollence relation on the family of all sets is an equivalence relation.

Example 3.29. (i) Let $A =$ set of all alphabets of the English language.

$$B = \{1, 2, \dots, 26\}$$

Then there is a one-to-one correspondence between A and B , so that $|A| = |B| = 26$.

(ii) As seen in Example 3.27 there is one-to-one correspondence between $2\mathbb{Z}$ and $4\mathbb{Z}$, hence $|2\mathbb{Z}| = |4\mathbb{Z}|$.

(iii) $|\mathbb{N}| = |\mathbb{N} \cup \{0\}|$, because

$$f : \mathbb{N} \rightarrow \mathbb{N} \cup \{0\}$$

$$f(n) = n - 1$$

is a bijective function, so that there is a one-to-one correspondence between \mathbb{N} and $\mathbb{N} \cup \{0\}$. Therefore \mathbb{N} and $\mathbb{N} \cup \{0\}$ are equipollent and correspondingly $|\mathbb{N}| = |\mathbb{N} \cup \{0\}|$.

Remark 3.4. If A and B are two sets such that $|A| = |B|$, then there exists a bijective mapping from A to B . This mapping need not be unique. If A and B are infinite sets such that $A \subsetneq B$ it is quite possible that $|A| = |B|$.

Definition 3.12. Let A and B be two sets. If there exists a one-to-one mapping from A to B then we say that $|A| \leq |B|$. If $|A| \leq |B|$ and $|A| \neq |B|$, then $|A| < |B|$.

If A and B are finite sets such that $|A| \leq |B|$ and $|B| \leq |A|$, we conclude immediately that $|A| = |B|$ by the antisymmetry of \leq . But when A and B are infinite sets such a property of antisymmetry also holds between $|A|$ and $|B|$.

This is the following theorem.

Theorem 3.12. (Schröder Bernstein) If A and B are two sets such that $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.

Proof: Beyond the scope of the book. □

3.12 Countable Sets

Definition 3.13. (Countably infinite set): A set A is said to be countably infinite if there is a one-to-one correspondence between A and \mathbb{N} , i.e. $|A| = |\mathbb{N}|$.

Definition 3.14. (Countable set): A set A is said to be countable if it is either finite or countably infinite.

A set which is not countable is said to be uncountable. Traditionality, the cardinality of \mathbb{N} is denoted by the symbol \aleph_0 (pronounced 'alpha naught'). Thus the cardinality of a countably infinite set is \aleph_0 .

Let A be countably infinite set. If f is a bijective mapping from A to \mathbb{N} . Then elements of A can be listed as $\{a_1, a_2, \dots\}$, where $a_k = f^{-1}(k)$, for $k \in \mathbb{N}$.

Example 3.30.

(i) $\mathbb{Z}, 2\mathbb{Z}, \mathbb{Q}^+, \mathbb{Q}$ are examples of countably infinite sets.

(ii) $\mathbb{R}^+, \mathbb{R}, (0, 1), (a, b)$ for $a < b$ are examples of uncountable sets.

Proofs of the above examples are given later.

Since equipollence is an equivalence relation, therefore, to prove that a set A is countable, it is sufficient to prove that it is equipollent to some countable set B , i.e. there is a one-to-one correspondence between A and B . Similarly to prove that a set X is uncountable it must be equipollent to some set Y , which is known to be uncountable. These two facts will be used repeatedly.

Theorem 3.13. *Every subset of a countable set is either finite or countable.*

Proof: Let A be a countable set and B be a subset of A . Since A is countable, therefore it can be listed as $\{a_1, a_2, \dots\}$ (1)

Two cases arise

Case 1. B is finite. Then B is countable by definition.

Case 2. B is infinite. Consider the listing of A as given in (1). From this listing omit those elements of A which are not in B . The list which remains gives a listing of the elements of B . Hence B is countable. \square

Theorem 3.14. *Every infinite set has a countable subset.*

Proof: Let A be an infinite set. Then A is nonempty, so choose $a_1 \in A$. Let $A_1 = A \setminus \{a_1\}$. Since A is infinite, therefore $A_1 \neq \phi$. Choose $a_2 \in A_1$ and let $A_2 = A \setminus \{a_1, a_2\}$. Again, $A_2 \neq \phi$. Choose $a_3 \in A_2$, and let $A_3 = A \setminus \{a_1, a_2, a_3\}$. Continuing in this way, we obtain a countable subset $\{a_1, a_2, a_3, \dots\}$ of A . \square

Corollary 3.15. *If A is any infinite set then $\aleph_0 \leq |A|$.*

The next theorem gives a relation between the cardinalities of a set and its power set.

Theorem 3.16. *If A is any set, then $|A| < |\mathcal{P}(A)|$*

Proof: Define $f : A \rightarrow \mathcal{P}(A)$

by $f(a) = \{a\} \quad \forall a \in A$.

Then f is one-to-one, for if $a, b \in A$ such that

$f(a) = f(b)$

then $\{a\} = \{b\}$

$\Rightarrow a = b$

Hence $|A| \leq |\mathcal{P}(A)|$... (1)

We now prove that $|A| \neq |\mathcal{P}(A)|$

On the contrary, suppose that

$|A| = |\mathcal{P}(A)|$, so that there exists a bijective mapping $g : A \rightarrow \mathcal{P}(A)$

Consider $B = \{a \in A \mid a \notin g(a)\}$

Then $B \subseteq A$ so that $B \in \mathcal{P}(A)$.

Since g is onto, therefore, there exists $x \in A$ such that

$g(x) = B$... (1)

Two cases arise:

Case 1. $x \in B$. Then $x \notin g(x)$ by definition of B

i.e. $x \notin B$ using (1)

which is a contradiction to $x \in B$.

Case 2. $x \notin B$

In this case $x \notin g(x)$ by (1)

Thus $x \in A$ and $x \notin g(x)$

$\therefore x \in B$ by definition of B which contradicts $x \notin B$.

Hence, in either case we reach at a contradiction. So that there does not exist

any bijective mapping g from A to $\mathcal{P}(A)$.

Thus $|A| < |\mathcal{P}(A)|$ □

Corollary 3.17. $\mathcal{P}(\mathbb{N})$ is uncountable.

Taking $A = \mathbb{N}$ in the above theorem, we get $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$

$\Rightarrow |\mathcal{P}(\mathbb{N})| > \aleph_0 \Rightarrow \mathcal{P}(\mathbb{N})$ is uncountable. □

We have proved that $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$ and $\mathcal{P}(\mathbb{N})$ is an uncountable set. It can be shown that $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$

It follows that \mathbb{R} is uncountable as $\mathcal{P}(\mathbb{N})$ is uncountable. The cardinality of \mathbb{R} is denoted by c . Thus

$$|\mathcal{P}(\mathbb{N})| = |\mathbb{R}| = c$$

i.e $2^{\aleph_0} = c$

Thus we get

$$\aleph_0 < c.$$

There is a conjecture: “There does not exist any set A such that $\aleph_0 < |A| < c$.”

This is called Cantor’s *continuum hypothesis*. This hypothesis can be restated as “Every uncountable set of real numbers has cardinality c ”.

It is proved later that $|(0, 1)| = |\mathbb{R}|$

Thus $|(0, 1)| = c$. Let $I = (0, 1)$

Then as above $|\mathcal{P}(I)| > |I|$ and $|\mathcal{P}(I)| = 2^c$

Also $2^c > c$. Thus, given a cardinal number k we can always find a cardinal number bigger than k , namely 2^k . This is because if $k = |A|$, where A is some set, then $|\mathcal{P}(A)| = 2^k$ and $2^k > k$.

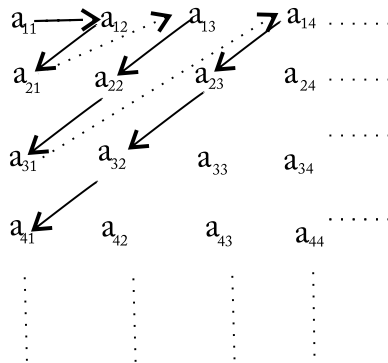
Thus $\aleph_0 < c < 2^c < 2^{2^c} < \dots$, where $c = 2^{\aleph_0}$ and $n < \aleph_0$ for all $n \in \mathbb{N}$. \aleph_0 is the smallest cardinal number. The set of integers can be listed as $0, 1, -1, 2, -2, 3, -3, \dots$ so that we can establish a one-to-one correspondence between \mathbb{N} and \mathbb{Z} . This is done in a problem given later.

Theorem 3.18. *The countable infinite union of countably infinite sets is countable.*

Proof: Let $\{A_n : n \in \mathbb{N}\}$ be a family of sets, where each A_n is a countably infinite set. To prove that $A = \cup_{n \in \mathbb{N}} A_n$ is countable. It is sufficient to give a listing of the elements of A .

Let $A_i = \{a_{i_1}, a_{i_2}, a_{i_3}, \dots\}$

write the elements of the sets A_i as follows:



Traverse this array as indicated by the arrows. Thus every element of A can be labeled as b_1, b_2, \dots . Hence A is countable. \square

Problem 3.12. Show that \mathbb{Z} is countable.

Solution: Define $f : \mathbb{N} \rightarrow \mathbb{Z}$

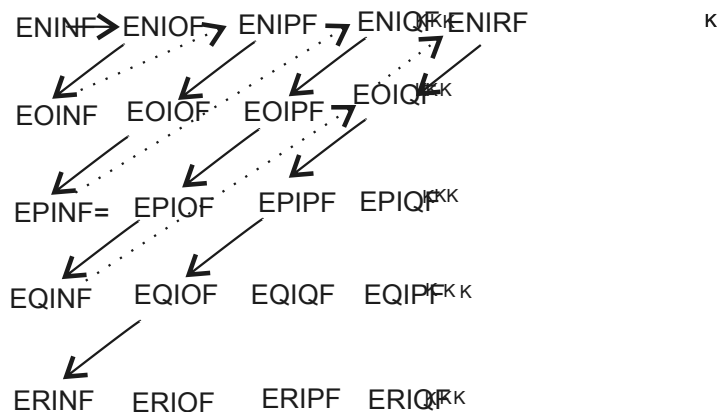
$$\text{by } f(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ -(n-1)/2 & \text{if } n \text{ is odd} \end{cases}$$

Thus f is a bijective function (Verify), so that \mathbb{N} and \mathbb{Z} have same cardinality. Since \mathbb{N} is countable, we conclude that \mathbb{Z} is also countable.

Problem 3.13. Show that $\mathbb{N} \times \mathbb{N}$ and \mathbb{N} have the same cardinality \aleph_0 .

Solution: $\mathbb{N} \times \mathbb{N} = \{(m, n) | m, n \in \mathbb{N}\}$

The elements of $\mathbb{N} \times \mathbb{N}$ can be listed as shown



Listing of $\mathbb{N} \times \mathbb{N}$ K

The order is indicated by the arrow. In case we want to define a mapping, we define

$$f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$$

$$\text{by } f(a, b) = \frac{(a+b-2)(a+b-1)}{2} + a$$

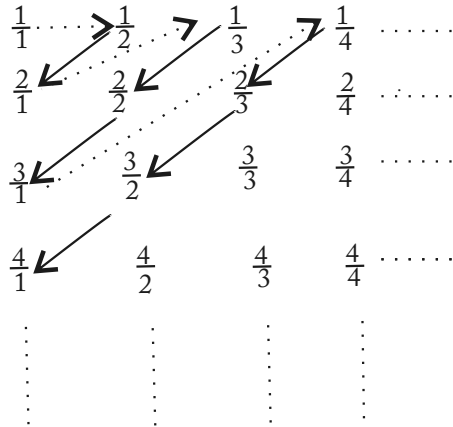
Thus $|\mathbb{N} \times \mathbb{N}| = \aleph_0$.

Problem 3.14. The set of all rational numbers is countable.

Solution:

Step1 We first prove that the set of all positive rational numbers \mathbb{Q}^+ is countable. We arrange these numbers not in order of size, but according to the size of the sum of the numerator and denominator. Begin with all positive rational numbers, $\frac{a}{b}$ such that $a+b = 2$. There is only one such rational number, namely $\frac{1}{1}$. Next list, with increasing numerator, all those numbers $\frac{a}{b}$ for which $a+b = 3$, i.e $\frac{1}{2}$ and $\frac{2}{1}$. Now those for which $a+b = 4$, i.e $\frac{1}{3}, \frac{2}{2}, \frac{3}{1}$, come next in the list. Next are those for which $a+b = 5$, i.e $\frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1} = 4$ and so on. We

now list all these together which are from the beginning, omitting those already listed. Thus we get the sequence $1, \frac{1}{2}, 2, \frac{1}{3}, 3, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, 4, \frac{1}{5}, \dots$ which contains each positive rational number exactly once. Figure below gives a systematic representation of this manner of listing. The first row contains all rational numbers with numerator 1, second row all numbers with numerator 2 and so on. Traverse this array as indicated by the arrows, leaving out the numbers which have been encountered already.



List of positive rationals

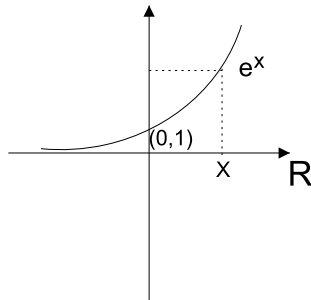
Thus the set of all positive rational numbers can be labelled as $a_1, a_2, a_3, a_4 \dots$

Step 2 We now prove that the set of all rational numbers is countable. All the rational numbers can be listed as $0, a_1, -a_1, a_2, -a_2, \dots$ which proves that \mathbb{Q} is countable.

Problem 3.15. Prove that \mathbb{R} and \mathbb{R}^+ have the same cardinality.

Solution: Define $f: \mathbb{R} \rightarrow \mathbb{R}^+$ by $f(x) = e^x$

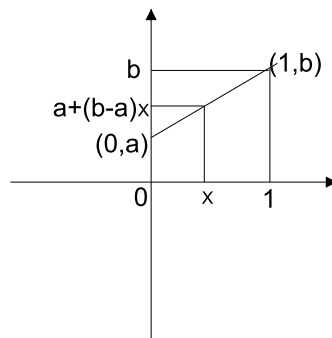
It is easy to verify that f is a bijective mapping. Thus \mathbb{R} and \mathbb{R}^+ have the same cardinality.



Problem 3.16. Show that if $a, b \in \mathbb{R}$ such that $a < b$, then $(0, 1)$ and (a, b) are equipollent.

Solution: Define $f : (0, 1) \rightarrow (a, b)$ by $f(x) = a + (b - a)x$

It can be easily seen that f is a bijective mapping. Thus $(0, 1)$ and (a, b) are equipollent.

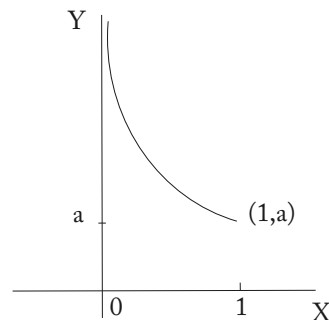


Problem 3.17. Prove that $(0, 1]$ and $[a, \infty)$ where $a \in \mathbb{R}$ have the same cardinality.

Solution: Define $f : (0, 1] \rightarrow [a, \infty)$

by $f(x) = \frac{1}{x} - 1 + a$

Then verify that f is a bijective mapping, so that $(0, 1]$ and $[a, \infty)$ have the same cardinality.

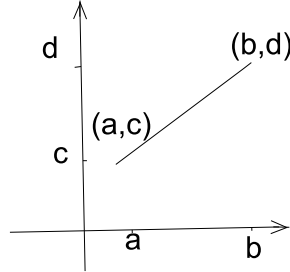


Problem 3.18. Any two closed intervals have the same cardinality.

Solution: Let $[a, b]$ and $[c, d]$ be two intervals. Define $f : [a, b] \rightarrow [c, d]$

by $f(x) = c + \frac{d-c}{b-a}(x - a)$

Then f is a bijective mapping (Verify). Hence $[a, b]$ and $[c, d]$ have the same cardinality.



Problem 3.19. Prove that $(0, 1)$ is uncountable.

Solution: We prove the result by contradiction. Let, if possible $(0, 1)$ be countable. List all the numbers in $(0, 1)$ as x_1, x_2, x_3, \dots . Write each x_i in the decimal form. We also need to agree to write $0.499\dots$ as 0.5 etc, so that there is no repetition. Thus we have

$$x_1 = 0.x_{11}x_{12}x_{13}x_{14}\dots$$

$$x_2 = 0.x_{21}x_{22}x_{23}x_{24}\dots$$

$$x_3 = 0.x_{31}x_{32}x_{33}x_{34}\dots$$

$$x_4 = 0.x_{41}x_{42}x_{43}x_{44}\dots$$

\vdots
 \vdots
 \vdots

We will find a number y in $(0, 1)$ different from the x_i 's. To do this we proceed as follows:

If $x_{11} = 5$, define $y_1 = 6$

If $x_{11} \neq 5$, define $y_1 = 5$

Similarly, if $x_{22} = 5$, define $y_2 = 6$

If $x_{22} \neq 5$, define $y_2 = 5$

In general, for $i = 1, 2, 3, \dots$

$$\text{define } y_i = \begin{cases} 6 & \text{if } x_{ii} = 5 \\ 5 & \text{if } x_{ii} \neq 5 \end{cases}$$

Let $y = 0.y_1y_2y_3\dots$. Thus y differs from x_i in the i^{th} place, for all $i = 1, 2, 3, \dots$

i.e $y \neq x_i$ for any $i = 1, 2, 3, \dots$

Also $y \neq 0$, $y \neq 0.99\dots$

so that $y \in (0, 1)$, which is a contradiction. Hence our assumption that $(0, 1)$ is countable, is wrong, so that $(0, 1)$ is not countable.

Problem 3.20. Show that $(0, 1)$ and \mathbb{R} have same cardinality.

Solution: Define $f : (0, 1) \rightarrow \mathbb{R}$ by $f(x) = \frac{x-1/2}{x(x-1)}$

We prove that f is a bijective mapping.

Step 1. To prove that f is one-to-one.

Let $x_1, x_2 \in (0, 1)$ such that $f(x_1) = f(x_2)$

$$\text{Then } \frac{x_1-1/2}{x_1(x_1-1)} = \frac{x_2-1/2}{x_2(x_2-1)}$$

Simplifying, we get

$$(x_1 - x_2)[(1 - x_1)(1 - x_2) + x_1x_2] = 0$$

Since $0 < x_1, x_2 < 1$

$$\therefore (1 - x_1)(1 - x_2) + x_1x_2 > 0$$

so we must have $x_1 - x_2 = 0$, i.e $x_1 = x_2$. Hence f is one-to-one.

Step 2. To prove that f is onto.

Let $y \in \mathbb{R}$. To prove that f is onto, we have to find some $x \in (0, 1)$ such that $f(x) = y$. Clearly $f(\frac{1}{2}) = 0$. Choose $y \neq 0$. Suppose $f(x) = y$.

We will prove that $x \in (0, 1)$.

$$\begin{aligned} f(x) &= y \\ \Rightarrow y &= \frac{x-1/2}{x(x-1)} \\ \Rightarrow x^2y - x(y+1) + 1/2 &= 0 \\ \Rightarrow x &= \frac{(y+1) \pm \sqrt{(y+1)^2 - 2y}}{2y} \end{aligned}$$

$$\text{Take } x = \frac{(y+1) - \sqrt{(y+1)^2 - 2y}}{2y}$$

When $y > 0$

$$\text{Now } (y+1)^2 = y^2 + 1 + 2y$$

$$\therefore (y+1)^2 > y^2 + 1$$

$$\Rightarrow y+1 > \sqrt{y^2 + 1} \quad (\text{taking the positive square root})$$

$$\Rightarrow y+1 - \sqrt{(y+1)^2 - 2y} > 0$$

$$\Rightarrow \frac{(y+1) - \sqrt{(y+1)^2 - 2y}}{2y} > 0, \quad \because y > 0$$

$$\text{Also } \sqrt{y^2 + 1} > 1$$

$$\therefore -\sqrt{(y+1)^2 - 2y} < -1$$

$$\Rightarrow y - \sqrt{(y+1)^2 - 2y} < y - 1$$

$$\Rightarrow 1 + y - \sqrt{(y+1)^2 - 2y} < y$$

$$\Rightarrow \frac{1+y - \sqrt{(y+1)^2 - 2y}}{2y} < \frac{1}{2}$$

$$\text{Thus } 0 < \frac{1+y - \sqrt{(y+1)^2 - 2y}}{2y} < \frac{1}{2}$$

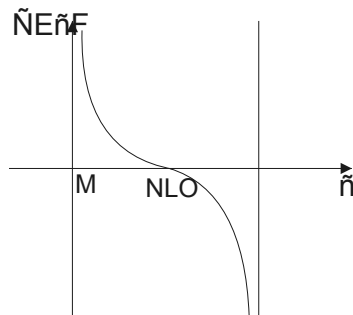
i.e $0 < x < \frac{1}{2}$, so that $x \in (0, \frac{1}{2}) \subseteq (0, 1)$

When $y < 0$, we can similarly prove that $\frac{1}{2} < x < 1$

Thus, there exists $x \in (0, 1)$ such that $f(x) = y$

Hence f is onto.

Hence f is a bijective mapping so that $(0, 1)$ and \mathbb{R} have the same cardinality.



Graph of $y = f(x)$

Aliter: An elegant proof of the above result is given below.

The function f

$$f : (0, 1) \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

defined by $f(x) = \pi(x - \frac{1}{2})$ is a bijective function.

Also $g : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R}$

$g(x) = \tan x$ is a bijective function.

Therefore $g \circ f : (0, 1) \rightarrow \mathbb{R}$ is a bijective function.

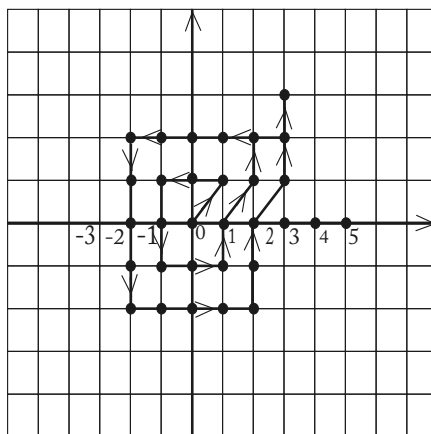
Hence $|(0, 1)| = |\mathbb{R}|$.

The beauty of mathematics lies in not only proving the result, but proving it elegantly. So, you can choose the proof you like.

Problem 3.21. *The elements of $Z \times Z$ are shown in the figure.*

Show that $Z \times Z$ is countable by indicating a systematic way of listing the elements. Also list the first 20 elements.

Solution: The arrow show how the elements are listed. First 20 elements are $(0, 0), (1, 1), (0, 1), (-1, 1), (-1, 0), (-1, -1), (0, -1), (1, -1), (1, 0), (2, 1), (2, 2), (1, 2), (0, 2), (-1, 2), (-2, -2), (-2, 1), (-2, 0), (-2, -1), (-2, -2), (-1, -2)$.



1

Problem 3.22. *Prove that $(0, 1]$ is uncountable by proving $|(0, 1]| = |(0, 1)|$.*

Solution: Let $A = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right\}$

$= \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$

Then A is countable subset of $(0, 1]$. Define

$f : (0, 1] \rightarrow (0, 1)$

by $f(1) = \frac{1}{2}$

$f\left(\frac{1}{k}\right) = \frac{1}{k+1}, \quad k \geq 2, \quad k \in \mathbb{N}$

$f(x) = x$ if $x \notin A$

Then f is a bijective mapping so that $|(0, 1]| = |(0, 1)|$. Since $(0, 1)$ is uncountable, therefore $(0, 1]$ is also uncountable.

Problem 3.23. *Let S be an infinite set and $x \notin S$. Then prove that S and $S \cup \{x\}$ have the same cardinality.*

Solution: Since S is infinite, therefore it has a countable infinite set S_1 .

Let $S_2 = S \setminus S_1$, so that $S = S_1 \cup S_2$

$S \cup \{x\} = S_1 \cup S_2 \cup \{x\}$

$= (S_1 \cup \{x\}) \cup S_2$

Since S_1 is countably infinite, list its elements as

$\{s_1, s_2, s_3, \dots\}$

Define $f : S \rightarrow S \cup \{x\}$

by $f(s_1) = x$

$f(s_{k+1}) = s_k, \quad k \geq 1$

$f(s) = s, \quad s \notin S_1$

Then f is a bijective mapping. (Prove it!)

Hence $|S| = |S \cup \{x\}|$.

Corollary 3.19. F is a finite set and S is any infinite set. Then $|S| = |S \cup F|$.

3.13 Exercise

1. Find a one-to-one correspondence between the sets A and B where

(i) $A = \{5, 11, 31, 18\}$
 $B = \{\phi, \{b, c\}, \{d, e\}, \{1, 2, x\}\}$

(ii) $A = [0, \infty), B = (-\infty, 0]$

(iii) $\mathbb{Z} \times \mathbb{Z}$ and $\{a + ib \in \mathbb{C} | a, b \in \mathbb{Z}\}$.

2. Prove that the following sets have the same cardinality

(i) $2\mathbb{N}$ and $3\mathbb{N}$

(ii) N and $N \cup \{0\}$

(iii) \mathbb{Z} and $2\mathbb{Z}$

(iv) $4\mathbb{Z}$ and $31\mathbb{Z}$

(v) $A \times B$ and $B \times A$, where A and B are any two sets.

(vi) $(A \times B) \times C$ and $A \times (B \times C)$ where A, B, C are any 3 sets.

3. On the family of all sets J , define a relation \sim as follows:

For $A, B \in J$, $A \sim B$ if and only if $|A| = |B|$.

Prove that \sim is an equivalence relation on J .

4. Prove that the union of two countable sets is countable.

5. Prove that the set of all irrational numbers is uncountable.

6. Prove that the sets A and B have the same cardinality, where

(i) $A = [5, \infty), B = [3, \infty)$

(ii) $A = [5, \infty), B = [-3, \infty)$

(iii) $A = (6, \infty), B = (7, \infty)$

(iv) $A = (-6, \infty), B = (7, \infty)$

(v) $A = (a, \infty), B = (b, \infty)$

(vi) $A = (a, \infty), B = (-\infty, -b)$.

7. Prove that $(0, 1)$ and $(0, \infty)$ have the same cardinality.

8. Prove that $(a, b), (c, d)$ have the same cardinality.

9. Show that the following sets have the same cardinality as \mathbb{R} ,

(i) $[2, 8]$

(ii) $[a, b]$

(iii) (a, ∞)

(iv) $(-\infty, -5)$

(v) $(-\infty, a)$.

10. Show that any two circles have the same number of points on their circumference. What is the cardinality of this set?

11. Prove that the number of points on the circumference of a semicircle is the same as those on \mathbb{R} .
12. By listing the elements in a systematic way, prove that the following sets are countable. Also list the first 10 elements.
- All the positive integral powers of 5.
 - All integral powers of 3.
 - $\{a, b, c\} \times \mathbb{N}$
 - $\mathbb{N} \times \mathbb{Z}$
 - $\mathbb{Z} \times (\mathbb{N} \cup \{0\})$
 - $(\mathbb{N} \times \mathbb{N}) \cup ((-\mathbb{N}) \times (-\mathbb{N}))$
13. Determine whether the following sets are finite, countably infinite or uncountable. Justify your answer.
- The set of all sentences in the English language that contain five words and each word of length at most ten.
 - $\{(a, b) \in \mathbb{Q} \times \mathbb{Q} \mid a + b = 50\}$
 - $\{(a, b) \in \mathbb{R} \times \mathbb{R} \mid a + b = 50\}$
 - $\{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a + b = 50\}$
 - $\{(a, b) \in \mathbb{Q} \times \mathbb{Q} \mid b = \sqrt{4 - a^2}\}$
 - $\{(a, b) \in \mathbb{R} \times \mathbb{R} \mid b = \sqrt{4 - a^2}\}$
 - The set of all grains of bajra in a gunny bag.

3.14 Solved Problems

Problem 3.24. Show that the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = \frac{x}{\sqrt{(x^2+2)}}$ is one-to-one. Find range of f . Is it onto? Find a suitable inverse.

Solution: Let $x_1, x_2 \in \mathbb{R}$ such that $f(x_1) = f(x_2)$.

$$\text{Then } \frac{x_1}{\sqrt{(x_1^2+1)}} = \frac{x_2}{\sqrt{(x_2^2+2)}}$$

Squaring both sides, we get

$$\frac{x_1^2}{x_1^2+2} = \frac{x_2^2}{x_2^2+2}$$

$$\Rightarrow x_1^2 = x_2^2$$

$$\Rightarrow x_1 = \pm x_2$$

Since $f(x_1) = f(x_2)$, therefore x_1, x_2 both have the same sign, so that

$$x_1 = x_2.$$

To find range f

Let $x \in \mathbb{R}$ and $f(x) = y$

$\therefore \frac{x}{\sqrt{2+x^2}} = y$. Thus x, y have the same sign.

Squaring we get

$$y^2 = \frac{x^2}{2+x^2} = 1 - \frac{2}{2+x^2}$$

$$\Rightarrow x^2 = \frac{2y^2}{1-y^2}$$

$$\Rightarrow x = \frac{\sqrt{2y}}{\sqrt{1-y^2}} \quad (\because x, y \text{ have the same sign})$$

Thus x is real when $y^2 < 1$

i.e $|y| < 1$

Hence range $f = (-1, 1)$

Thus f is not onto.

To find f^{-1}

Clearly f^{-1} is a function defined on $(-1, 1)$.

If $f^{-1}(x) = y$, then

$$f(y) = x$$

Solving for y in terms of x (as above) we get

$$y = \frac{\sqrt{2x}}{\sqrt{1-x^2}},$$

$$\text{Hence } f^{-1}(x) = \frac{\sqrt{2x}}{\sqrt{1-x^2}}, x \in (-1, 1).$$

Problem 3.25. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by $f(x) = 2[x] - x$.

(i) Prove that f is bijective.

(ii) Find a formula for $f^{-1}(x)$.

(iii) Draw the graph of f . Can you decide from the graph that f has an inverse?

Solution: If n is an integer, and $n \leq x < (n+1)$, then $[x] = n$.

$\therefore f(x) = 2n - x$, if $n \leq x < (n+1)$.

f is onto

Let $y \in \mathbb{R}$, the codomain. Then there exists $n \in \mathbb{Z}$ such that $n \leq y < n+1$

Two cases arise:

Case 1. $y = n$. Take $x = y$. Then $x \in \mathbb{R}$ and

$$f(x) = 2n - x$$

$$= n$$

$$= y$$

Case 2. $n < y < n+1$

In this case $[y] = n$

Now $n < y < n+1$

$$\Rightarrow n > 2n - y > n - 1$$

$$\Rightarrow n + 2 > 2n - y + 2 > n + 1$$

Let $x = 2n - y + 2$, Then $x \in \mathbb{R}$

$$[x] = n + 1$$

$$f(x) = 2[x] - x$$

$$= 2(n+1) - (2n - y + 2)$$

$$= y$$

Also $x = 2n - y + 2$

$$= 2([y] + 1) - y$$

Thus there exists $x \in \mathbb{R}$ such that $f(x) = y$.

Hence in both cases, we get for $y \in \mathbb{R}$, there exists $x \in \mathbb{R}$ such that $f(x) = y$, so that f is onto.

Now, show that f is one-to-one.

Let $x_1, x_2 \in \mathbb{R}$ such that

$$f(x_1) = f(x_2)$$

Let $x_1 = n_1 + r_1$, where n_1 is an integer and $0 \leq r_1 < 1$.

$x_2 = n_2 + r_2$, where n_2 is an integer and $0 \leq r_2 < 1$.

Then $[x_1] = n_1, [x_2] = n_2$

Now $f(x_1) = f(x_2)$

$$\Rightarrow 2[x_1] - x_1 = 2[x_2] - x_2$$

$$\Rightarrow 2n_1 - (n_1 + r_1) = 2n_2 - (n_2 + r_2)$$

$$\Rightarrow n_1 - r_1 = n_2 - r_2$$

$$\Rightarrow n_1 - n_2 = r_1 - r_2$$

But $n_1 - n_2 \in \mathbb{Z}$ and $0 \leq |r_1 - r_2| < 1$ so that above result holds when both

sides are zero.

$$\text{i.e } r_1 - r_2 = 0, n_1 - n_2 = 0$$

$$\Rightarrow r_1 = r_2, n_1 = n_2$$

$$\Rightarrow x_1 = x_2$$

Hence f is one-to-one.

(ii) Clearly,

$$f(x) = -x, 0 \leq x < 1$$

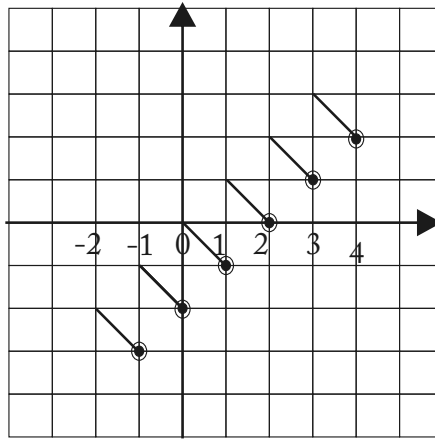
$$f(x) = 2 - x, 1 \leq x < 2$$

$$f(x) = 4 - x, 2 \leq x < 3$$

$$f(x) = -2 - x, -1 \leq x < 0$$

$$f(x) = -4 - x, -2 \leq x < -1$$

etc.



Graph of $f(x)$

Problem 3.26. A is a set such that $o(A) = 6$. Find $o(\mathcal{P}(A))$.

Solution: Let $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$. Let C = set of six digits numbers with digits 0 or 1

We now define a mapping from $\mathcal{P}(A)$ to C and prove that the mapping is bijective. If $B \in \mathcal{P}(A)$, i.e $B \subseteq A$, then to B assign the six digits number $b_1b_2b_3b_4b_5b_6$ such that

$$b_i = \begin{cases} 0 & \text{if } a_i \notin B \\ 1 & \text{if } a_i \in B \end{cases}$$

Thus $f : \mathcal{P}(A) \rightarrow C$ such that $f(B) = b_1b_2\dots b_6$ as defined above.

$$\text{Clearly } f(\phi) = 000000$$

$$f(A) = 111111$$

Now we show that f is onto. Let $c = d_1d_2d_3d_4d_5d_6 \in C$.

The d_i 's are 0 or 1.

$$f^{-1}(c) = \begin{cases} \phi & \text{if } d_i = 0 \quad \forall i = 1, 2, 3, \dots, 6 \\ \{a_i \in A \mid d_i = 1, 1 \leq i \leq 6\} & \end{cases}$$

$$= K \text{ (say)}$$

Clearly, $K \in \mathcal{P}(A)$ and $f(K) = c$. Hence f is onto.

Now we show that f is one-to-one. Let $X, Y \in \mathcal{P}(A)$ such that $f(X) = f(Y)$.

Then $x_1x_2x_3x_4x_5x_6 = y_1y_2y_3y_4y_5y_6$. so that $x_i = y_i \quad i = 1, 2, \dots, 6$, proving that $X = Y$.

Hence f is one-to-one.

Thus f is a bijective mapping so that f is one-to-one correspondence between $\mathcal{P}(A)$ and C . But $o(C) = 2^6$. (since each digit can be chosen in 2 ways).

Hence $o(\mathcal{P}(A)) = 2^6$.

The above method can be generalized to find the cardinality of the power set of a finite set A .

3.15 Supplementary Exercise

1. State whether the following statements are true or false. Justify the false ones
 - (i) Every relation is a function.
 - (ii) Every function is a relation.
 - (iii) The smallest equivalence relation on a set of n elements has n elements.
 - (iv) The smallest equivalence relation on a set is the identity relation.
 - (v) Reflexivity is redundant in the definition of an equivalence relation R because if $(a, b) \in R$, then $(b, a) \in R$, by symmetry. By transitivity $(a, b), (b, a) \in R \Rightarrow (a, a) \in R$.
 - (vi) Every symmetric relation and anti-symmetric relation is reflexive.
 - (vii) $R = \{(1, 2), (1, 3)\}$ is a transitive relation on $A = \{1, 2, 3\}$.
 - (viii) A binary operation associates at least one element of A to every of $A \times A$.
 - (ix) If a binary operation $*$ is commutative than parenthesis are not needed in $a * b * c$.
 - (x) A binary operation is always commutative and associative.
 - (xi) The number of bijective functions from A to A is n^n , where $n = o(A)$
 - (xii) Every function is invertible.
 - (xiii) If $f : A \rightarrow B$ and $g : B \rightarrow A$ are functions such that $gof = i_A$ then f is invertible and $g = f^{-1}$.
 - (xiv) If $A = \{a, b, c\}, B = \{x, y\}$ and f is a function from A to B , then we can define two functions g_1 and g_2 from B to A such that $fog_1 = i_B, fog_2 = i_B$.
 - (xv) If $f : A \rightarrow B$ and $g : B \rightarrow A$ are such that $fog = i_B$, then it is always true that $gof = i_A$.
 - (xvi) If $f : A \rightarrow B$ is a bijection and $g : B \rightarrow A$ is inverse of f , then $gof = fog =$ identity function.
 - (xvii) The function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = (x + 1)(x - 2)(x + 5)$ is a bijective function.
 - (xviii) The function $f(x) = 4x + 5$ is a bijective function from \mathbb{Z} to itself.
 - (xix) If $f : A \rightarrow B$ be any function, $X \subseteq A$ then $f^{-1}(f(X)) = X$.
 - (xx) If $f : A \rightarrow B$ be any function and Y be any subset of B , then $f(f^{-1}(Y)) = Y$.

- (xxi) The floor function from \mathbb{R} to \mathbb{Z} is onto.
- (xxii) If \sim is not an equivalence relation on a set A then \sim is neither reflexive, nor symmetric nor transitive.
- (xxiii) A countable set must be infinite.
- (xxiv) If A and B have the same cardinality then their power sets also have the same cardinality.
- (xxv) If A is a proper subset of B , then $|A| < |B|$.
- (xxvi) If A is a subset of B , then $|B \sim A| = |B| - |A|$.
- (xxvii) A subset of an infinite set may be finite.
- (xxviii) Every superset of an countable set is uncountable.
- (xxix) The set of irrational numbers is countable.
- (xxx) The set of rational numbers is uncountable.
- (xxxi) The power set of a countable set is countable.
2. Let $A = \{1, 2, 3, 4, 5\}$, $B = \{x, y, z, t\}$, $C = \{x, y, t\}$, $D = \{2, 4\}$. Define a function f from A to B and g from B to A as follows: $f = \{(2, x), (3, z), (4, t), (1, t), (5, x)\}$, $g = \{(x, 1), (y, 3), (z, 4), (t, 1)\}$ Find
 (i) fog ,
 (ii) gof (iii) $(fog)of$
 (iv) $f^{-1}(C)$, (v) $f(D)$ (vi) $f^{-1}(f(D))$
 (vii) $f(f^{-1}(C))$
 Also find $R(f)$ and $f^{-1}R(f)$.
3. Let $f : \mathbb{Z} \rightarrow \mathbb{N}$ be defined by $f(x) = \begin{cases} |x|, & \text{if } x < 0 \\ 2|x|, & \text{if } x \geq 0 \end{cases}$
 Is f an invertible function? If yes, find f^{-1} .
4. If f and g are functions defined on \mathbb{R} by $f(x) = ax + b$, $g(x) = cx + d$. Prove that $fog = gof$ if and only if $f(d) = g(b)$.
5. Find the inverse of the function $f : A \rightarrow A$ defined by $f(x) = \frac{1-x}{1+x}$, where $A = \mathbb{R} \sim \{-1\}$.
6. Can you construct an example of a function which has a right inverse and a left inverse, but the two are not equal? Justify your answer.
7. Prove that the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = x^3 + 3ax^2 + 3bx + c$ is a bijection if $a^2 < b$.
8. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = \begin{cases} x & \text{if } x \text{ is rational,} \\ 1-x & \text{if } x \text{ is irrational} \end{cases}$
 Prove that f is invertible and find f^{-1} .
9. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $f(x) = x + \frac{1}{x}$. Find suitable domain and range of f , so that f has an inverse. Also find f^{-1} .
10. Prove that the number of points on a sphere is the same as those on a plane.
11. Prove that the number of points on the surface of any two spheres is the same.

12. If A is a set with n elements find $|\mathcal{P}(A)|$.
13. Prove that any two open intervals have the same cardinality.
14. Prove by induction that a finite union of countably infinite sets is countably infinite.
15. Prove that $|\mathbb{Q}^+| = \aleph_0$.
16. Prove that the countable union of finite sets is countable.
17. Prove that the following mappings f are bijective mappings on the given intervals:
 - (i) $f : [a, \infty) \rightarrow [b, \infty)$ defined by
 $f(x) = x - a + b$
 - (ii) $f : [a, b] \rightarrow [c, d]$ defined by
 $f(x) = b + \frac{d-b}{c-a}(x - a)$
 - (iii) $f : (0, 1) \rightarrow (0, \infty)$ defined by
 $f(x) = \frac{1-x}{x}$
18. For some $c \in \mathbb{R}$, prove that $(0, 1)$ and (c, ∞) have the same cardinality.

3.16 Answers to Exercises

Exercise - 3.6

1. 7^9
2. $f(\frac{\pi}{2}) = -1$
 $f(-\pi) = 0 = f(\pi)$
 $f(\frac{\pi}{4}) = \frac{1}{\sqrt{2}}$
3. $\{y\}, \{x, y\}, f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$. They may not be equal.
4. (i) $[1, 2)$
(ii) $\phi, \phi, 1, 0, -3$
(iii) No, No
5. (i) $\{-1, 1\}$
(ii) ϕ
(iii) $[-3, 3]$
6. (i) $\{0, 1, 2, 3, 4\}$
(ii) $\{0, 1, 2, 3, 4\}$
(iii) All multiples of 5, set of integers of the form $5k + 1, k \in \mathbb{Z}$
(iv) Set of integers of the form $5k + 3$
7. $[\frac{-1}{3}, 1]$
8. *Hint* $\sqrt{|x| - x}$ is defined when $|x| \geq x$ i.e when $x \leq 0$
 $D(f) = (-\infty, 0), R(f) = (0, \infty)$
9. ϕ

10. (i) $D(f) = \mathbb{R}, R(f) = [-\frac{1}{2}, \frac{1}{2}]$
 (ii) $D(f) = \mathbb{R}, R(f) = [\frac{1}{4}, \frac{1}{2}]$
 (iii) $D(f) =]-1, 1[, R(f) = [1, \infty)$
11. (iii) and (iv)
14. (i) 10^8
 (ii) $10P_8$
 (iii) None

Exercise - 3.10

1. $f \circ f(x) = \begin{cases} 2 - 4x & \text{if } x < 0, \\ 4x + 5 & \text{if } x \geq 0 \end{cases}$
2. $(-1, 1)$
3. (i) f^{-1} exists and is
 $f^{-1} = \{(9, 1), (7, 2), (5, 3), (3, 4), (1, 5)\}$
 (ii) g^{-1} does not exist as g is not bijective.
5. (i) $(h \circ g \circ f)(x) = 1 + \frac{3}{2}x$
 (ii) $[\frac{3}{2}, \frac{5}{2}]$
6. (i) $(g \circ f)(x) = \frac{1}{x - [x]}, x \in (1, 2)$
 (ii) Yes
 (iii) $(g \circ f)^{-1}(x) = \frac{x+1}{x}, x \in (1, \infty)$
10. (i) $R(f) = (1, \infty), R(g) = (0, \infty), R(h) = (1, \infty)$.
 (ii) Since range of the first function is contained in the domain of second function, \therefore all composition are defined.
 (iii) Since f and g are bijective mappings, so f^{-1}, g^{-1} and $(g \circ f)^{-1}$ exist.
11. (i) $(-\infty, 1) \cup (1, \infty)$
 (ii) $(h \circ g \circ f)^{-1} = \frac{3}{2(x-1)}$
12. $f^{-1}(x) = \frac{-1}{2x^2}, x \in (0, \infty)$
13. (i) $[0, 3]$
 (ii) $(f \circ f)(x) = \begin{cases} 2 + x, & 0 \leq x \leq 2, \\ 4 - x, & 2 < x \leq 3 \end{cases}$
 (iii) $f^{-1}(x) = \begin{cases} 3 - x, & 0 \leq x < 1, \\ x - 1, & 1 \leq x \leq 3 \end{cases}$
14. $f^{-1}(x) = 2 + \sqrt{x-1}$.

Exercise - 3.13

6. (i) – (iv) are particular cases of this.

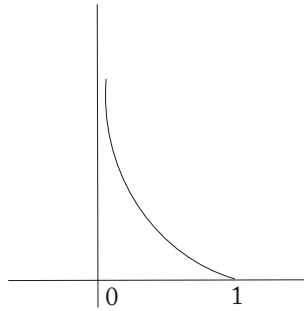
$$(v) f : A \rightarrow B$$

$$f(x) = x - a + b$$

$$(vi) f : A \rightarrow B$$

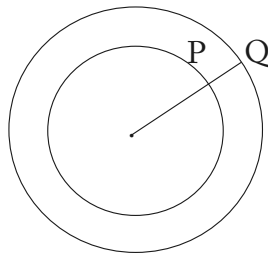
$$f(x) = -x + a + b$$

7. *Hint:* $x \rightarrow \frac{1}{x} - 1$



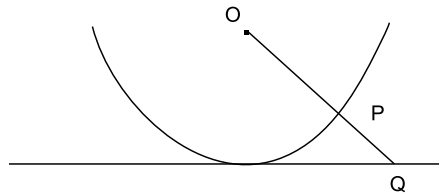
10. *Hint:* w.l.o.g. take circles to be concentric of radii r_1, r_2 .

$$P \leftrightarrow Q$$



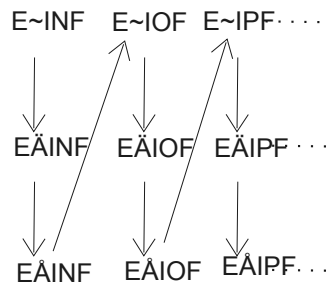
$$(r_1 \cos \theta, r_1 \sin \theta) \leftrightarrow (r_2 \cos \theta, r_2 \sin \theta)$$

11.

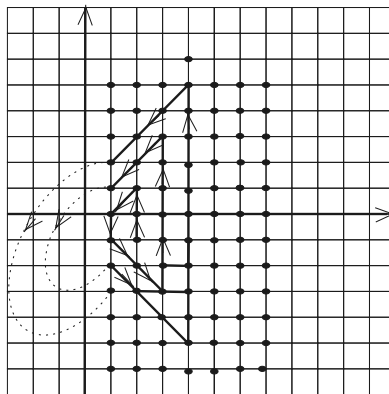


O —centre of a circle
 $P \leftrightarrow Q$

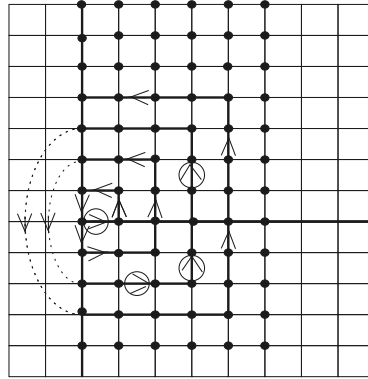
12. (i) $5^1, 5^2, 5^3, 5^4, 5^5, \dots$
 $\{5^n : n \in \mathbb{N}\}$
 (ii) $3^0, 3^1, 3^{-1}, 3^2, 3^{-2}, \dots$
 (iii) $(a, 1), (b, 1), (c, 1), (a, 2), (b, 2), (c, 2), (a, 3), (b, 3), (c, 3), (a, 4) \dots$



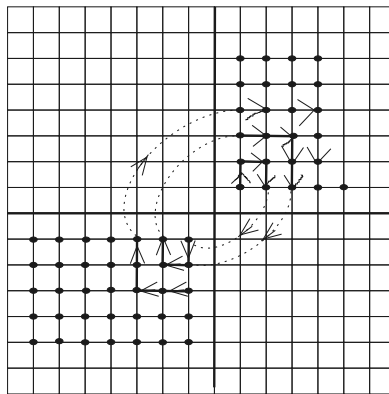
(iv) One route is shown and others are possible.



(v)



(vi)



13. (i) countable.
 (ii) countable.
 (iii) uncountable.
 (iv) countable.
 (v) countable.
 (vi) uncountable.
 (vii) finite.

Answers to Supplementary Exercises

1.
 (i) F
 (ii) T
 (iii) T
 (iv) T ~
 (v) F
 (vi) F
 (vii) T
 (viii) F

- (xi) F
- (xii) F
- (xiii) F
- (xiv) T
- (xv) F
- (xvi) F
- (xvii) F
- (xviii) F
- (xix) F
- (xx) F
- (xxi) T
- (xxii) F
- (xxiii) F
- (xxiv) T
- (xxv) F
- (xxvi) F
- (xxvii) T
- (xxviii) F
- (xxix) F
- (xxx) F
- (xxxi) F

2. (i) $\{(x, t), (y, z), (z, t), (t, t)\}$
 (ii) $\{(1, 1), (2, 1), (3, 4), (4, 1), (5, 1)\}$
 (iii) $\{(1, t), (2, t), (3, t), (4, t), (5, t)\}$
 (iv) $\{1, 2, 4, 5\}$
 (v) $\{x, t\}$
 (vi) $\{1, 2, 4, 5\}$
 (vii) $\{x, t\}$
 $R(f) = \{x, z, t\}$, $f^{-1}(R(f)) = \{1, 2, 3, 4, 5\}$.

3. No, f is not one-to-one.

5. $f^{-1} = f$

6. No. Existence of right inverse \Rightarrow function is onto.
 Existence of left inverse \Rightarrow function is one-to-one.
 \therefore function is bijective and therefore invertible.

7. *Hint.* $\Rightarrow f$ is not one-to-one.
 \Rightarrow For $x_1 \neq x_2$, $f(x_1) = f(x_2)$.
 $\Rightarrow f'(k) = 0$ for some k between x_1 and x_2 . $\Rightarrow x^2 + 2ax + b = 0$ has a real root $x = k$. \Rightarrow no real root for $a^2 < b$. $\Rightarrow f$ is one-to-one.
 $f(x) \rightarrow -\infty$ as $x \rightarrow -\infty$ and $f(x) \rightarrow \infty$ as $x \rightarrow \infty$.
 Since $f(x)$ is continuous, $\therefore f(x)$ assumes all values between $-\infty$ and ∞ .
 $\therefore f$ is onto. Hence f is bijective.

8. $f^{-1} = f$

9. $D(f) = (-1, 1) \sim \{0\}$, $R(f) = \mathbb{R} \sim (-2, 2)$

$$f(x) = \begin{cases} \frac{x + \sqrt{x^2 - 4}}{2}, & \text{if } x \in (0, 1] \\ \frac{x - \sqrt{x^2 - 4}}{2}, & \text{if } x \in [-1, 0) \end{cases}$$
$$f^{-1}(x) = \begin{cases} \frac{x + \sqrt{x^2 - 4}}{2}, & \text{if } x \in [2, \infty] \\ \frac{x - \sqrt{x^2 - 4}}{2}, & \text{if } x \in (-\infty, -2] \end{cases}$$

Chapter 4

Number System

The natural members have been extended to integers, rational numbers, real number etc... This chapter is devoted to the study of natural numbers and integers. The principle of mathematical induction, well ordering principle and their applications are given the greatest common divisor and the congruence relation in integers have also been discussed.

4.1 Number Systems

To count objects we use the numbers 1, 2, 3, ... in our daily life. These numbers are called counting numbers. Mathematicians call them natural numbers. An axiomatic approach to study these numbers was given in 1899 by the Italian mathematician Giuseppe Peano. He gave axioms for the study of natural numbers called Peano's Axioms. They are:

Axiom 1. There exists a natural number 1.

Axiom 2. There exists a one-to-one mapping

$$f : \mathbb{N} \rightarrow \mathbb{N}$$

If $n \in \mathbb{N}$, then $f(n)$ is called the successor of n .

Axiom 3. The mapping f is not bijective. In fact $1 \notin f(\mathbb{N})$.

Axiom 4. If $K \subseteq \mathbb{N}$ such that $1 \in K$ and $n \in K \Rightarrow f(n) \in K$, then $K = \mathbb{N}$.

Axiom 1 gives us that \mathbb{N} is a non-empty set.

Axiom 2 gives us that if $m, n \in \mathbb{N}$ such that $f(m) = f(n)$ then $m = n$.

Axiom 3 tells us that 1 is not the successor of any natural number.

Axiom 4 gives us a test to determine when a subset K of \mathbb{N} is identical with \mathbb{N} .

It is also called the **axiom of induction**.

Starting with these axioms, we build the system of natural numbers and define addition and multiplication in \mathbb{N} . How we go about doing this, is not the purpose of this book. Interested reader may refer to any book on number systems.

Algebraic Properties of Natural Numbers

For all $a, b, c \in \mathbb{N}$,

1. $a + (b + c) = (a + b) + c$ (Associative law of addition)

2. $a + b = b + a$ (Commutative law of addition)
3. $a + c = b + c \Rightarrow a = b$ (Cancellation law of addition)
4. $(ab)c = a(bc)$ (Associative law of multiplication)
5. $ab = ba$ (Commutative law of multiplication)
6. $ac = bc \Rightarrow a = b$ (Cancellation law of multiplication)
7. There exists a natural number 1 such that $a \cdot 1 = 1 \cdot a = a$ (Existence of identity for multiplication)
8. $a(b + c) = ab + ac$ (Distributivity of multiplication over addition)

It can be shown that if m and n are any two natural numbers, then exactly one of the following holds:

- (i) $m = n$
- (ii) $m = n + u$ for some $u \in \mathbb{N}$
- (iii) $n = m + v$ for some $v \in \mathbb{N}$

This helps us in defining an order relation in \mathbb{N} . We say that ' m is greater than n ' (denoted by $m > n$) if $m = n + u$ for some $u \in \mathbb{N}$. We can also define other order relations in \mathbb{N} in terms of the relation $>$.

Let $m, n \in \mathbb{N}$. Then we define

- (i) m is less than n ($m < n$) if $n > m$.
- (ii) m is less than or equal to n ($m \leq n$) if either $m = n$ or $m < n$.
- (iii) m is greater than or equal to n ($m \geq n$) if either $m = n$ or $m > n$.

Example 4.1. $4 > 2$, $\because 4 = 2 + 2$
 $5 < 9$, $\because 9 > 5$. In fact $9 = 5 + 4$.

If K is a non empty subset of \mathbb{N} , then $l \in K$ is said to be a least element of K if $x \in K \Rightarrow x = l$ or $x > l$.

Order Properties of Natural Numbers

The relation 'greater than' i.e. $>$ satisfies the following properties:

1. Law of trichotomy
 For $m, n \in \mathbb{N}$ exactly one of the following holds:
 $m = n$, $m > n$, $n > m$.
2. Transitivity
 For $m, n, p \in \mathbb{N}$
 $m > n$ and $n > p \Rightarrow m > p$.
3. Monotone property for addition
 For $m, n, p \in \mathbb{N}$
 $m > n \Rightarrow m + p > n + p$.
4. Monotone property for multiplication
 For $m, n, p \in \mathbb{N}$
 $m > n \Rightarrow mp > np$.
5. Well ordering principle(WOP)
 Every non-empty subset of the set of the natural numbers has a least element.

The properties stated above can be formulated in terms of the relation $<$ also. To solve the equation $x + 11 = 15$, in the set of natural numbers, we get $x = 4$. But if we want to solve the equation $x + 15 = 11$, we cannot find any solution in the set of natural numbers. Thus, the set \mathbb{N} lacks many properties. If $m, n \in \mathbb{N}$, then the equation $x + m = n$ may or may not have a solution in \mathbb{N} . Naturally we would like a number system such that for every pair of elements m, n of this system, the equation $x + m = n$ has a solution in this system. Moreover, we would like this number system to share all the properties of \mathbb{N} , if possible. It would be nice if this number system is an extension of \mathbb{N} . Such a system is the set of integers, denoted by \mathbb{Z} .

Algebraic Properties of Integers

For all $a, b, c \in \mathbb{Z}$,

1. $a + (b + c) = (a + b) + c$ (Associative law for addition)
2. $a + b = b + a$ (Commutative law for addition)
3. There exists an element $0 \in \mathbb{Z}$ such that $a + 0 = a$ (Existence of zero element)
4. For each $a \in \mathbb{Z}$, there exists $-a \in \mathbb{Z}$ such that $a + (-a) = 0$ (Existence of negative)
5. $(ab)c = a(bc)$ (Associative law for multiplication)
6. $ab = ba$ (Commutative law for multiplication)
7. There exists an element $1 \in \mathbb{Z}$ such that $a1 = a$ (Existence of unity)
8. $a + c = b + c \Rightarrow a = b$ (Cancellation law for addition)
9. $ac = bc, c \neq 0 \Rightarrow a = b$ (Cancellation law for multiplication)
10. $a(b + c) = ab + ac$ (Distributivity of multiplication over addition)

Order Properties of Integers

There exists a subset \mathbb{Z}^+ of \mathbb{Z} , called the set of positive integers such that

- (i) For each $a \in \mathbb{Z}$, exactly one of the following holds
 $a \in \mathbb{Z}^+, a=0, -a \in \mathbb{Z}^+$
- (ii) $a, b \in \mathbb{Z}^+ \Rightarrow a+b \in \mathbb{Z}^+, ab \in \mathbb{Z}^+$

In \mathbb{Z} we define $a > b$ if $a + (-b) = a - b \in \mathbb{Z}^+$.

In terms of the relation $>$, the above properties can be rewritten as:

- (a) If $a, b \in \mathbb{Z}$, then exactly one of the following holds:
 $a > b, a = b, b > a$ (Law of trichotomy)
- (b) If $a, b, c \in \mathbb{Z}$, such that
 $a > b, b > c \Rightarrow a > c$ (Transitivity)
- (c) If $a, b, c \in \mathbb{Z}$, such that $a > b \Rightarrow a + c > b + c$ (Monotone property for addition)
- (d) If $a, b, c \in \mathbb{Z}$ such that $a > b, c > 0 \Rightarrow ac > bc$ (Monotone property for multiplication)

The well ordering principle does not hold in \mathbb{Z} . In this respect all the properties of \mathbb{N} are not carried over to \mathbb{Z} . In fact, a modified version of this principle holds. It is “Well Ordering Principle for Integers”.

“Every non-empty subset of the set of non-negative integers has a least element.”

The principle of mathematical induction is important in every area of mathematics. It is one of the most basic method which is used to prove results. This is a way which establishes the truth of a statement about all natural numbers or sometimes about all sufficiently large natural numbers. A formal statement of the principle of induction is as follows:

Theorem 4.1. (*First principle of induction*)(FPI) Let $\{P(n)|n \in \mathbb{N}\}$ be a set of statements. If

(i) $P(1)$ is true,

(ii) If $k \in \mathbb{N}$, such that

if $P(k)$ is true, then $P(k+1)$ is also true,

then $P(n)$ is true for all natural numbers n .

Proof: Let $K = \{n \in \mathbb{N} | P(n) \text{ is true}\}$.

Step1 Since $P(1)$ is true.

$\therefore 1 \in K$. Hence $K \neq \phi$.

Step2 Let $k \in K$.

Then $P(k)$ is true

$\Rightarrow P(k+1)$ is true by (ii)

$\Rightarrow k+1 \in K$.

Hence, by the axiom of induction $K = \mathbb{N}$, i.e. $P(n)$ is true for all $n \in \mathbb{N}$. \square

The method of induction is one of the most powerful tools for proving theorems. A proof by induction is like climbing a staircase with infinite number of steps. The first step has to be climbed, and having climbed any particular step, we can climb the next step. Then the whole staircase can be climbed. This is similar to the two steps which have been described in the proof of the above theorem. While using induction we shall always use these two steps.

Example 4.2. The number of subsets of a set containing n elements is 2^n . Let $P(S)$ denote the power set of S . Here the statement $T(n)$ is:

If S is a set containing n elements then $P(S)$ has 2^n elements.

Step1 If $S = \{a\}$, then $P(S) = \{\{a\}, \phi\}$.

Thus $P(S)$ has 2 elements. Hence the result holds for $n = 1$, So $T(1)$ is true.

Step2 Suppose that $T(k)$ is true, i.e. if S is a set containing k elements, then $P(S)$ contains 2^k elements.

Consider a set S with $k+1$ elements.

Let $S = \{a_1, a_2, \dots, a_{k+1}\}$

For each subset A of S , either

$a_{k+1} \in A$ or $a_{k+1} \notin A$. The collection of all those subsets of S which do not contain a_{k+1} is $P(B)$, where $B = \{a_1, a_2, \dots, a_k\}$.

Since B contains k elements, by hypothesis $P(B)$ contains 2^k elements. Thus the number of subsets of S not containing a_{k+1} is 2^k .

Each subset A of S containing a_{k+1} can be obtained from a subset G of B by adding a_{k+1} to G . Thus there are precisely 2^k subsets of S each of which contains a_{k+1} .

Thus the total number of subsets of S is $2^k + 2^k = 2^{k+1}$, so that $T(k+1)$ is true.

Thus, by first principle of induction, $T(n)$ is true for all $n \in \mathbb{N}$.

The two conditions in the first principle of induction are equally important. In case any one of them fails to hold, the result need not hold. This is shown by the following examples.

Example 4.3. Let $P(n)$ be the statement:

$1 + 2 + \cdots + n = n(n+1)/2 + 5$ for each natural number n .

Is $P(n)$ true for all $n \in \mathbb{N}$?

Suppose $k \in \mathbb{N}$ such that $P(k)$ is true

$$\text{i.e. } 1 + 2 + \cdots + k = \frac{k(k+1)}{2} + 5 \quad (1)$$

Now

$$\begin{aligned} 1 + 2 + \cdots + k + (k+1) &= \frac{k(k+1)}{2} + 5 + (k+1), && \dots \text{using (1)} \\ &= \frac{(k+1)(k+2)}{2} + 5, \end{aligned}$$

Hence $P(k+1)$ is true.

Does this mean that $P(n)$ is true for all $n \in \mathbb{N}$?

For $n = 3$, $1 + 2 + 3 = 6$

whereas $P(3)$, $1 + 2 + 3 = 11$.

$\therefore P(3)$ is not true. This is because $P(1)$ is not true, as $P(1)$ gives $1 = 6$. Thus, step 1 of the FPI fails to hold and so FPI can not be applied.

Example 4.4. For each $n \in \mathbb{N}$, let $P(n)$ be the statement:

$$1 + 2 + \cdots + n = n$$

Clearly $P(1)$ is true.

Suppose $P(k)$ is true, for $k \in \mathbb{N}$. Thus $1 + 2 + \cdots + k = k$ (1)

$$\begin{aligned} 1 + 2 + \cdots + k + (k+1) &= k + (k+1), && \text{using (1)} \\ &= 2k + 1 \end{aligned}$$

Hence $P(k+1)$ is not true. So FPI cannot be applied.

Sometimes it happens that the statement $P(n)$ does not hold for a finitely many of natural numbers. In such cases the FPI may be modified. The starting point of the induction, instead of being 1, is some natural number $m > 1$. This is precisely the second principle of induction.

Theorem 4.2. (Second principle of induction) Let $P(n)$ be a statement for each natural number n . suppose

(a) $P(m)$ is true for some $m \in \mathbb{N}$

(b) If $k \in \mathbb{N}$ such that $k \geq m$ and $P(k)$ is true $\Rightarrow P(k+1)$ is true

then $P(n)$ is true for all $n \geq m$.

Proof: Define $A = \begin{cases} \phi, & \text{if } m=1 \\ \{1, 2, \dots, m-1\} & \text{if } m>1 \end{cases}$

$T = \{n \in \mathbb{N} : n \geq m \text{ and } P(n) \text{ is true}\}$ is the truth set of $P(n)$ and $K = T \cup A$

Step 1 Clearly $1 \in K$.

Step 2 Let $k \in K$. Then there are three possibilities

(i) $k \in \{1, 2, \dots, m-2\}$, (ii) $k = m-1$, (iii) $k > m-1$

We shall take these cases one by one.

Case 1. When (i) holds

$k+1 \in \{2, 3, \dots, m-1\} \subseteq A$

so that $k+1 \in K$.

Case 2. When (ii) holds

$k = m-1 \Rightarrow k+1 = m$

Since $P(m)$ is true, $\therefore P(k+1)$ is true.

So $k+1 \in T \subseteq K$

Hence $k+1 \in K$.

Case 3. When (iii) holds

$k > m-1 \Rightarrow k+1 > m$

Since $k \geq m$ and $k \in K$. Also $k \notin A$, $\therefore k \in T$

so that $P(k)$ is true. By (b) $P(k+1)$ is also true i.e. $k+1 \in T \subseteq K$.

Hence in either of the three cases $k+1 \in K$, so that by the axiom of induction $K = \mathbb{N}$.

Thus $P(n)$ is true for all $n \geq m$. \square

Observe that for $m = 1$, it is the first principle of induction. This can be considered as a generalization of the first principle of induction, in the sense that the starting point is not necessarily 1 but some other natural number m . This can be compared to a child climbing a staircase with infinite number of steps, and the child starts from some particular step (say m th) and not necessarily the first step.

Theorem 4.3. (*Third principle of induction*) Let $\{P(n) : n \in \mathbb{N}\}$ be a set of statements, one for each natural number n . If

(a) $P(1)$ is true, and

(b) If for each $k \in \mathbb{N}$, $P(m)$ is true $\forall m < k \Rightarrow P(k)$ is true,

then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof: Let $F = \{p \in \mathbb{N} : P(p) \text{ is false}\}$

We assert that $F = \phi$.

If $F \neq \phi$, Then F is a non empty subset of \mathbb{N} and so by the well ordering principle, F has a least element, say l . Thus if $m < l$, then $m \notin F$. In other words $P(m)$ is true for all $m < l$. By hypothesis (b) $P(l)$ is true which implies that $l \notin F$.

This contradicts that $l \in F$.

Hence our assumption is wrong, so that $F = \phi$.

$\therefore P(n)$ is true for all $n \in \mathbb{N}$. \square

The FPI is nothing but a characterization of the WOP, as is proved in the following theorem.

Theorem 4.4. *The well ordering principle and the first principle of induction are equivalent.*

Proof: Suppose that WOP holds. We shall prove FPI holds.

Let $P(n)$ be a statement, one for each $n \in \mathbb{N}$, and $K = \{n \in \mathbb{N} : P(n) \text{ is true}\}$ is such that

- (i) $1 \in K$
- (ii) $k \in K \Rightarrow k + 1 \in K$.

We prove that $P(n)$ is true, for all $n \in \mathbb{N}$. For this we show that $K = \mathbb{N}$. Clearly $K \subseteq \mathbb{N}$. Suppose that $K \neq \mathbb{N}$, so that K is a proper subset of \mathbb{N} . Let $F = \mathbb{N} - K$. Then $F \neq \emptyset$. Also $F \subseteq \mathbb{N}$. Thus F is a non-empty subset of \mathbb{N} so that by WOP, F has a least element say l . Thus $l \in F$.

Now $1 \in K \Rightarrow 1 \notin F \Rightarrow l > 1 \Rightarrow l \geq 2$.

Thus $l = m + 1$ for some $m \in \mathbb{N}$,

Now $m < m + 1 = l$ i.e. $m < l$

$\therefore m \notin F$ as l is the least element of F .

$\therefore m \in K \Rightarrow m + 1 \in K$

$\Rightarrow l \in K$

$\Rightarrow l \notin F$ which is a contradiction.

Hence our assumption is wrong.

$\therefore F = \emptyset \Rightarrow K = \mathbb{N}$.

Hence proved.

Conversely, let the FPI holds. We prove that WOP holds. Let S be any non-empty subset of the set of natural numbers.

Let $K = \{x \in \mathbb{N} \mid x \leq s, \forall s \in S\}$.

Clearly $K \subseteq \mathbb{N}$.

Since $1 \leq s \forall s \in S$

$\therefore 1 \in K$ so that $K \neq \emptyset$.

Let $m \in S$. Then $m + 1 \not\leq m$ so that $m + 1 \notin K$.

$\therefore K \subseteq \mathbb{N}$ and $K \neq \mathbb{N}$.

Thus K is a non-empty, proper subset of \mathbb{N} so that $\exists l \in K$ such that $l + 1 \notin K$.

We assert that l is the least element of S .

Since $l \in K$

$\therefore l \leq s \forall s \in S$.

If $l \notin S$, then $l < s \forall s \in S$.

so that $l + 1 \leq s \forall s \in S$

$\Rightarrow l + 1 \in K$, which is a contradiction,

as $l + 1 \notin K$.

Hence $l \in S$.

Thus $l \in S$ is such that $l \leq s \forall s \in S$,

so that l is the least element of S . Thus S has a least element, is proved.

\therefore WOP holds. □

Divisibility

We now give a formal definition of divisibility, though you have been doing this since childhood.

Definition 4.1. Let $a, b \in \mathbb{Z}$. We say that 'a divides b' if there exists an element $c \in \mathbb{Z}$ such that $b = ac$.

We write it as $a \mid b$ and read it as 'a divides b'.

When 'a divides b' we may also say that 'a is a divisor of b' or 'a is a factor of b' or 'b is a multiple of a'. If 'a does not divide b' we write $a \nmid b$.

The following results about divisibility though trivial and obvious, will be used over and over again. We list them here for the sake of completeness.

Theorem 4.5. *Let $a, b, c, m, n \in \mathbb{Z}$. Then*

- (i) $a \mid a$ (*Reflexive property*)
- (ii) if $a \mid b$ and $b \mid c \Rightarrow a \mid c$ (*Transitive property*)
- (iii) if $a \mid b$ and $a \mid c \Rightarrow a \mid mb + nc$ (*Linear property*)
- (iv) if $a \mid b \Rightarrow ac \mid bc$ (*Multiplication property*)
- (v) if $ac \mid bc, c \neq 0 \Rightarrow a \mid b$ (*Cancellation property*)
- (vi) $1 \mid a$ (*Property of unity*)
- (vii) $a \mid 0$ (*Property of zero*)
- (viii) if $0 \mid a \Rightarrow a=0$ (*Zero divides only zero*)
- (ix) if $a \mid b \Rightarrow a \mid |b|$
- (x) if $a \mid b$ and $a \neq 0 \Rightarrow (b/a) \mid b$
- (xi) if $a \mid b$ and $b \neq 0 \Rightarrow |a| \leq |b|$.
- (xii) if $a \mid b$ and $b \mid a \Rightarrow |a| = |b|$

Proof: Left to the reader. □

For each $a \in \mathbb{Z}$, the integers $a, -a$ are such that each divides the other. Moreover, because of (xii), these are the only two elements which divides each other. So, $a, -a$ are called **associates**.

Definition 4.2. (Prime Number): *An integer $p > 1$ is said to be a prime number if its only divisors are $\pm 1, \pm p$.*

A number $p > 1$ is said to be composite if it is not a prime number. The number 1 is neither prime nor composite. It is called a unit. The set of integers can be divided into 4 disjoint classes, namely

- (i) primes and their associates
- (ii) composites and their associates
- (iii) units, i.e. 1 and -1
- (iv) zero.

4.2 Division Algorithm

A fundamental property of the integers is the division algorithm which can be proved by using the well ordering principle.

Theorem 4.6. *If $a, b \in \mathbb{Z}, b \neq 0$, then there exist unique $q, r \in \mathbb{Z}$ such that $a = bq + r, 0 \leq r < |b|$.*

Proof: Two cases arise

Case 1. $b > 0$

Existence of q and r

Let $A = \{a - bx : x \in \mathbb{Z}, a - bx \geq 0\}$.

We first prove that $A \neq \emptyset$. Two cases arise:

Case (i). $a \geq 0$

Then $a = a - b \cdot 0 \geq 0$ also $0 \in \mathbb{Z}$

$\therefore a \in A$. Hence $A \neq \emptyset$.

Case (ii). $a < 0$

$\therefore -a > 0$

Also $b > 0 \Rightarrow b \geq 1$
 $\Rightarrow -ab \geq -a$
 $\Rightarrow a - ab \geq 0$
 $\Rightarrow a \in A$
 $\Rightarrow A \neq \phi$

Hence in both the cases $A \neq \phi$.

Thus A is a non-empty subset of the set of non-negative integers so that by the well ordering principle A must have a least element, say r .

$\therefore \exists q \in \mathbb{Z}$ such that

$a - bq = r$ is the least element of A .

We assert that $r < b$. Let, if possible, $r \geq b$. Then $r = b + c$, for some $c \in \mathbb{Z}$ s.t. $0 \leq c < r$. Then

$$\begin{aligned} c &= r - b \\ &= (a - bq) - b \\ &= a - (b + 1)q \end{aligned}$$

Since $c \geq 0$, $\therefore c \in A$. Also $c < r$, which contradicts the fact that r is the least element of A . Hence our assumption is wrong, so that $r < b$.

Hence $r < b = |b|$.

Uniqueness

We have proved that $a = bq + r$, for some $q, r \in \mathbb{Z}$, $0 \leq r < b \dots (1)$

Let, there exists $q_1, r_1 \in \mathbb{Z}$ such that

$a = bq_1 + r_1$, $0 \leq r_1 < b \dots (2)$

Suppose $q > q_1$. Then

$(1) - (2) \implies$

$0 = b(q - q_1) + r - r_1$

$\implies b(q - q_1) = r_1 - r$

But $b(q - q_1) > b \because q > q_1$

$\therefore r_1 - r > b \dots (3)$

But $0 \leq r$, $r_1 < b$

So that $r_1 - r < b \dots (4)$

Thus (3) contradicts (4).

Hence $q > q_1$ is not possible, so $q \leq q_1$.

Similarly we prove that $q_1 \leq q$ so that $q = q_1$.

Now $bq + r = a = bq + r_1$

$\implies r = r_1$.

Hence the uniqueness of q and r is proved.

Case 2. $b < 0$

$b < 0 \implies -b > 0$

Applying case 1 to $-b$, there exists unique $q, r \in \mathbb{Z}$

such that

$a = (-b)q + r$, $0 \leq r < (-b)$

$\therefore a = b(-q) + r$, $0 \leq r < |b|$.

Hence in both cases, there exists unique $q, r \in \mathbb{Z}$ such that

$a = bq + r$,

$0 \leq r < |b|$

□

The integers q and r in the above theorem are called the quotient and remainder respectively, when a is divided by b . Note that the remainder is always non-negative.

Problem 4.1. Prove that the sum of the cubes of 3 consecutive positive integers is divisible by 9.

Solution: The problem amounts to proving that $n^3 + (n+1)^3 + (n+2)^3$ is divisible by 9, for all $n \in \mathbb{N}$.

Let $P(n) : n^3 + (n+1)^3 + (n+2)^3$ is divisible by 9.

$P(1) : 1^3 + 2^3 + 3^3$ is divisible by 9.

i.e. 36 is divisible by 9, which is true.

For $k \in \mathbb{N}$, let $P(k)$ be true, i.e.

$k^3 + (k+1)^3 + (k+2)^3$ is divisible by 9

i.e. $k^3 + (k+1)^3 + (k+2)^3 = 9m$ for some $m \in \mathbb{N}$

Now

$$\begin{aligned} (k+1)^3 + (k+2)^3 + (k+3)^3 &= (k+1)^3 + (k+2)^3 + k^3 + 27 + 9k(k+3) \\ &= 9m + 27 + 9k(k+3) \\ &= 9[m + 3 + k(k+3)] \end{aligned}$$

Hence $(k+1)^3 + (k+2)^3 + (k+3)^3$ is divisible by 9.

$\therefore P(k+1)$ is true.

By the principle of induction $P(n)$ is true for all $n \in \mathbb{N}$.

Problem 4.2. A rubber costs Rs 5 and a ball pen costs Rs 9. Show by using induction that any amount, in exact rupees, exceeding Rs 31 can be spent in buying rubbers and ball pens.

Solution: Let m be the number of rubbers and n be the number of pens. Then for k rupees, the problem is equivalent to finding non-negative integral solutions of

$$5m + 9n = k, \quad \text{for } k \geq 32$$

When $k = 32$, $m = 1, n = 3$ is a solution.

Suppose that for $k = t > 32$ a solution exists. Thus for some non-negative integers m_1, n_1 we have

$$\begin{aligned} t &= 5m_1 + 9n_1 \\ \therefore t+1 &= 5m_1 + 9n_1 + 1 \\ &= 9(n_1 - 1) + 5(m_1 + 2) \\ &= 9n_2 + 5m_2 \quad (\text{say}) \end{aligned}$$

where $m_2 = m_1 + 2, n_2 = n_1 - 1$.

Thus, m_2, n_2 is a solution, provided $n_2 \geq 0$ i.e. $n_1 \geq 1$.

If $n_1 = 0$, then $t = 5m_1$ and

$$\begin{aligned} t+1 &= 5m_1 + 1 \\ &= 9 \times 4 + 5(m_1 - 7) \quad \text{Since } t > 32, \therefore 5m_1 > 32. \end{aligned}$$

So for integral value of m_1 , the least value of $m_1 = 7$.

$$\therefore m_1 - 7 \geq 0. \therefore t+1 = 9 \times 4 + 5(m_1 - 7), \text{ where } m_1 - 7 \geq 0.$$

Hence there exists non-negative integral solution for $k = t + 1$.

Thus by the first principle of induction the result follows.

Problem 4.3. Use induction to prove that $5^{2n} - 2^{5n}$ is divisible by 7 for all $n \in \mathbb{N}$.

Solution: Let $P(n) : 5^{2n} - 2^{5n}$ is divisible by 7.

$5^2 - 2^5 = 25 - 32 = -7$ which is divisible by 7.

Hence $P(1)$ is true.

For some $k \geq 1$, let $P(k)$ be true.

i.e. $5^{2k} - 2^{5k}$ is divisible by 7.

$\therefore 5^{2k} - 2^{5k} = 7m$ for some $m \in \mathbb{Z} \quad \dots (1)$

Now

$$\begin{aligned} 5^{2(k+1)} - 2^{5(k+1)} &= 5^2 5^{2k} - 2^5 2^{5k} \\ &= 25 \cdot 5^{2k} - 32 \cdot 2^{5k} \\ &= 25(7m + 2^{5k}) - 32 \cdot 2^{5k} && \text{using (1)} \\ &= 25 \times 7m - 7 \times 2^{5k} \\ &= 7(25m - 2^{5k}) \end{aligned}$$

Thus $5^{2(k+1)} - 2^{5(k+1)}$ is divisible by 7, so that $P(k+1)$ is true.

Hence, by the principle of induction $P(n)$ is true for all $n \in \mathbb{N}$.

Problem 4.4. The fibonacci sequence $\langle a_n \rangle$ is given by $a_1 = a_2 = 1$, $a_n = a_{n-2} + a_{n-1}$, $n \geq 3$. Using the principle of induction, prove that

(i) $a_1 + a_4 + a_7 + \dots + a_{3(n+1)} = \frac{a_{3(n+1)}}{2}$

(ii) For $n \geq 2$,

$$a_n^2 + 1 = \begin{cases} a_{n-2}a_{n+2} & \text{if } n \text{ is odd} \\ a_{n-1}a_{n+1} & \text{if } n \text{ is even} \end{cases}$$

Solution:

(i) For $n = 1$,

$$\begin{aligned} \text{L.H.S.} &= a_1 + a_4 \\ \text{R.H.S.} &= \frac{a_6}{2} \\ &= \frac{1}{2}(a_5 + a_4) \\ &= \frac{1}{2}(a_4 + a_3 + a_4) \\ &= \frac{1}{2}a_3 + a_4 \\ &= \frac{1}{2}(a_1 + a_2) + a_4 \\ &= \frac{1}{2}(2a_1) + a_4 \\ &= a_1 + a_4 \end{aligned}$$

$$\therefore \text{L.H.S.} = \text{R.H.S.}$$

So the results holds for $n = 1$.

Let the result hold for $n = k$, i.e. $a_1 + a_4 + \dots + a_{3(k+1)} = \frac{a_{3(k+1)}}{2}$

$$\begin{aligned}
\text{Now} \\
a_1 + a_4 + \dots + a_{3(k+1)} + a_{3(k+1)+1} &= \frac{1}{2}a_{3k+3} + a_{3k+4} \\
&= \frac{1}{2}(a_{3k+3} + 2a_{3k+4}) \\
&= \frac{1}{2}(a_{3k+3} + a_{3k+4} + a_{3k+4}) \\
&= \frac{1}{2}(a_{3k+5} + a_{3k+4}) \\
&= \frac{1}{2}a_{3k+6} \\
&= \frac{a_{3(k+1)+3}}{2}
\end{aligned}$$

Hence the results holds for $n = k + 1$. Thus by the principle of induction the result holds for all $n \in \mathbb{N}$.

$$\begin{aligned}
\text{(ii) For } k = 2, \\
a_2^2 + 1 &= a_2a_1 + a_1^2 \quad \because a_1 = a_2 = 1 \\
&= a_1(a_2 + a_1) \\
&= a_1a_3
\end{aligned}$$

So the results holds for $k = 2$.

$$\begin{aligned}
\text{For } k = 3, \\
a_3^2 + 1 &= (a_2 + a_1)^2 + a_2 \quad \because a_2 = 1 \\
&= (a_1 + a_2)(a_2 + a_1) + a_2 \\
&= 2a_3 + a_2 \\
&= a_3 + a_4 \quad \text{using } a_4 = a_3 + a_2 \\
&= a_5 \\
&= a_1a_5 \\
\therefore a_3^2 + 1 &= a_1a_5
\end{aligned}$$

So that the result holds for $k = 3$.

Let the result hold for $n = k$.

We shall prove for $n = k + 1$. Two cases arise:

Case 1. k is odd.

Then $k + 1$ is even.

$$\begin{aligned}
a_{k+1}^2 + 1 &= a_{k+1}^2 + 1 + a_k^2 - a_k^2 \\
&= a_k^2 + 1 + a_{k+1}^2 - a_k^2 \\
&= a_{k-2}a_{k+2} + (a_{k+1} - a_k)(a_{k+1} + a_k) \\
&\quad \text{using induction hypothesis} \\
&= a_{k-2}a_{k+2} + a_{k-1}a_{k+2} \\
&= (a_{k-2} + a_{k-1})a_{k+2} \\
&= a_k a_{k+2} \\
&= a_{(k+1)-1}a_{(k+1)+1}
\end{aligned}$$

Hence the result holds for $n = k + 1$.

Case 2. k is even.

Then $k + 1$ is odd.

Also $a_k^2 + 1 = a_{k-1}a_{k+1}$. As in *Case 1*. it can be proved that

$$a_{k+1}^2 + 1 = a_{(k+1)-2}a_{(k+1)+2}$$

So that the result holds for $n = k + 1$. Thus by the principle of induction the result holds for all $n \in \mathbb{N}$.

Problem 4.5. Using induction prove that

$$(5 + \sqrt{13})^n + (5 - \sqrt{13})^n \text{ is divisible by } 2^n \text{ for all } n \in \mathbb{N}.$$

Solution: Let $a = 5 + \sqrt{13}$, $b = 5 - \sqrt{13}$, and for each $n \in \mathbb{N}$,

$P(n) : a^n + b^n$ is divisible by 2^n .

For $n = 1$, $a + b = 10$ which is divisible by $2 = 2^1$.

Hence the result holds for $n = 1$.

Let the result hold for all natural numbers $n \leq m$.

$$\begin{aligned} \text{Now } a^{m+1} + b^{m+1} &= (a^m + b^m)(a + b) - a^m b - ab^m \\ &= 10(a^m + b^m) - ab(a^{m-1} + b^{m-1}) \\ &= 10(a^m + b^m) - 12(a^{m-1} + b^{m-1}) \end{aligned}$$

Now 2^m divides $(a^m + b^m)$, so that 2^{m+1} divides $10(a^m + b^m)$.

Also 2^{m-1} divides $(a^{m-1} + b^{m-1})$, so that 2^{m+1} divides $12(a^{m-1} + b^{m-1})$.

Hence 2^{m+1} divides $10(a^m + b^m) - 12(a^{m-1} + b^{m-1})$ i.e. 2^{m+1} divides $(a^{m+1} + b^{m+1})$.

So that the result holds for $n = m + 1$. Thus by the third principle of induction the result holds for all $n \in \mathbb{N}$.

Problem 4.6. Show that $n! > 2^n$ for $n \in \mathbb{N}$, $n \geq 4$.

Solution: $4! = 24$, $2^4 = 16$.

Since $24 > 16$

$\therefore 4! > 2^4$

Hence the result holds for $n = 4$.

Let the result hold for some $k \geq 4$, i.e. $k! > 2^k$.

Now $(k+1)! = (k+1)k!$

Since $k \geq 4$

$\therefore k + 1 \geq 5 > 2$

Now $k! > 2^k$

$k + 1 > 2$

so that $(k + 1)k! > 2^k \cdot 2$

i.e. $(k + 1)! > 2^{k+1}$.

Hence the result holds for $n = k + 1$.

Thus, by the principle of induction, the result holds for $n \geq 4$.

Problem 4.7. If every non-empty finite set of natural numbers has a least element, prove that every non-empty subset of natural numbers has a least element.

Solution: Let S be a non-empty subset of natural numbers. Let $n \in S$, be any element of S . Define $T = \{x \in S \mid 1 \leq x \leq n\}$. Thus $n \in T$, so that $T \neq \phi$. Then T is a non-empty finite subset of \mathbb{N} , so by the given condition, T has a least element say t .

Then $t \in S$ and $t \leq n$.

For any $s \in S$ there are two possibilities: $s > n$ or $s \leq n$.

If $s > n$ then $t \leq n < s$, so that $t < s$.

If $s \leq n$ then $s \in T$ so that $t \leq s$ as t is the least element of T .

In either case $t \leq s \forall s \in S$.

Hence t is the least element of S .

Problem 4.8. If a_1, a_2, \dots, a_n are n positive numbers, then prove that

$$\frac{(a_1 + a_2 + \dots + a_n)}{n} \geq \sqrt[n]{a_1 a_2 \dots a_n}$$

The equality sign holds if and only if

$$a_1 = a_2 = \dots = a_n.$$

Solution: We shall use the principle of induction to prove the result.

Step1

We first prove the result for those integers n which are powers of 2, i.e. for $n = 2^m$.

Let $a_1, a_2 > 0$ be two numbers.

Then

$$\begin{aligned} & (\sqrt{a_1} - \sqrt{a_2})^2 \geq 0 \\ \therefore a_1 + a_2 - 2\sqrt{a_1a_2} & \geq 0 \\ \Rightarrow \frac{a_1 + a_2}{2} & \geq \sqrt{a_1a_2} \end{aligned} \quad (4.1)$$

Also equality in (4.1) holds if and only if $\sqrt{a_1} = \sqrt{a_2}$ i.e. $a_1 = a_2$.

Thus $\frac{a_1+a_2}{2} \geq \sqrt{a_1a_2}$ and equality holds if and only if $a_1 = a_2$. Thus the results hold when $n = 2$ i.e. when $m = 1$.

Let us assume that the result holds when $m = k$ i.e. $n = 2^k$.

Let $n = 2^{k+1}$ and a_1, a_2, \dots, a_n be n positive numbers. Then

$$\left(\begin{array}{l} \frac{a_1 + a_2}{2} \geq \sqrt{a_1a_2} \\ \frac{a_3 + a_4}{2} \geq \sqrt{a_3a_4} \\ \vdots \\ \vdots \\ \vdots \\ \frac{a_{n-1} + a_n}{2} \geq \sqrt{a_{n-1}a_n} \end{array} \right) \quad (a)$$

These are 2^k relations.

Adding the above inequalities, we get

$$\frac{(a_1 + a_2 + \dots + a_n)}{2^k} \geq (\sqrt{a_1a_2} + \dots + \sqrt{a_{n-1}a_n}) \quad (4.2)$$

$(\sqrt{a_1a_2}, \sqrt{a_3a_4}, \dots, \sqrt{a_{n-1}a_n})$ are 2^k numbers so that by the induction hypothesis, we have

$$\frac{(\sqrt{a_1a_2} + \dots + \sqrt{a_{n-1}a_n})}{2^k} \geq (\sqrt{a_1a_2 \dots a_{n-1}a_n})^{1/2^k} \quad (4.3)$$

(a) and (4.2) \Rightarrow

$$\begin{aligned} & \frac{a_1 + a_2 + \dots + a_n}{2^{k+1}} \geq (\sqrt{a_1a_2 \dots a_{n-1}a_n})^{1/2^k} \\ \Rightarrow \frac{a_1 + a_2 + \dots + a_n}{2^{k+1}} & \geq (a_1a_2 \dots a_{n-1}a_n)^{1/2^{k+1}} \\ \Rightarrow \frac{a_1 + a_2 + \dots + a_n}{n} & \geq (a_1a_2 \dots a_{n-1}a_n)^{1/n} \end{aligned} \quad (4.4)$$

Thus the result holds for $n = 2^{k+1}$.

Also equality holds in (4.4) if and only if it holds in (4.3) and (4.2), i.e. if and only if it holds in each of the inequalities (a). Thus we must have

$a_1 = a_2, a_3 = a_4, \dots, a_{n-1} = a_n$ and $\sqrt{a_1 a_2} = \sqrt{a_3 a_4} = \dots = \sqrt{a_{n-1} a_n}$.

This is so if and only if $a_1 = a_2 = \dots = a_n$.

Thus, by the principle of induction, the result holds when $n = 2^m$ for any $m \in \mathbb{N}$.

Step 2

We now prove the result for any integer n . Let a_1, a_2, \dots, a_n be n positive numbers, n being any natural number. Then there exists $m \in \mathbb{N}$ such that $2^m > n$. Let $A = (a_1 + a_2 + \dots + a_n)/n$, $G = (a_1 a_2 \dots a_n)^{1/n}$. Applying step 1 to the 2^m numbers $a_1, a_2, \dots, a_n, A, A, \dots, A$ (the number of A 's is $2^m - n$), we obtain

$$\frac{(a_1 + a_2 + \dots + a_n + A + A + \dots + A)}{2^m} \geq (a_1 a_2 \dots a_n A \dots A)^{1/2^m}$$

$$\text{i.e.} \quad \frac{nA + (2^m - n)A}{2^m} \geq (G^n A^{2^m - n})^{1/2^m}$$

$$\text{i.e.} \quad A^{2^m} \geq G^n A^{2^m - n}$$

$$\text{i.e.} \quad A^n \geq G^n$$

$$\text{i.e.} \quad A \geq G$$

$$\text{i.e.} \quad \frac{(a_1 + a_2 + \dots + a_n)}{n} \geq (a_1 a_2 \dots a_n)^{1/n}$$

The equality holds if and only if $a_1 = a_2 = \dots = a_n = A$.

Hence the result holds for all $n \in \mathbb{N}$.

Remark 4.1. *The above result is very important and it is called inequality of means. It can be stated as*
Arithmetic mean \geq Geometric mean.

4.3 Exercise

- Use mathematical induction to prove the truth of the following assertions for all $n \in \mathbb{N}$:
 - $3 \cdot 5^{2n+1} + 2^{3n+1}$ is divisible by 17.
 - $n(n+1)(n+2)$ is divisible by 6.
 - $10^n + 3 \cdot 4^{n+2} + 5$ is divisible by 9.
 - $2^n + 3^n - 5^n$ is divisible by 6.
 - $8^n - 3^n$ is divisible by 5.
 - $(2n)!$ is divisible by 2^n .
 - $2^{2^n} - 1$ is divisible by 3.
 - $3^{2^n} - 1$ is divisible by 8.
- Prove the following statements using the principle of induction, for all $n \in \mathbb{N}$.

- (i) $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.
- (ii) $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$
- (iii) $1^3 + 2^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}$
- (iv) $1 + 3 + 5 + \cdots + (2n-1) = n^2$
- (v) $1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} = (2 - \frac{1}{2^n})$
- (vi) $1.2.3 + 2.3.4 + 3.4.5 + \cdots + n(n+1)(n+2) = \frac{n(n+1)(n+2)(n+3)}{4}$
- (vii) $\frac{1}{1.4} + \frac{1}{4.7} + \cdots + \frac{1}{(3n-2)(3n+1)} = \frac{n}{3n+1}$
- (viii) $1^2 + 3^2 + \cdots + (2n-1)^2 = \frac{n(2n-1)(2n+1)}{3}$
- (ix) $\cos \theta + \cos 3\theta + \cdots + \cos(2n-1)\theta = \frac{1}{2} \sin 2n\theta \operatorname{cosec} \theta$

3. The Fibonacci sequence $\langle a_n \rangle$ is given by $a_1 = a_2 = 1$, $a_n = a_{n-1} + a_{n-2}$, $n \geq 3$. Using the principle of induction, prove the following

- (i) $a_1 + a_2 + \cdots + a_n = a_{n+2} - 1$
- (ii) $a_1 + a_3 + \cdots + a_{2n-1} = a_{2n}$
- (iii) $a_2 + a_4 + \cdots + a_{2n} = a_{2n+1} - 1$
- (iv) For any $r \in \mathbb{N}$, $a_r + a_{r+1} + \cdots + a_n = a_{n+2} - a_{r+1}$
- (v) $a_2 + a_5 + \cdots + a_{3n-1} = \frac{(a_{3n+1} - 1)}{2}$
- (vi) $a_3 + a_6 + \cdots + a_{3n} = \frac{(a_{3n+2} - 1)}{2}$
- (vii) $a_1^2 + a_2^2 + \cdots + a_n^2 = a_n a_{n+1}$
- (viii) $\sum_{k=1}^n (-1)^k a_k = (-1)^n a_{n-1} - 1$
- (ix) $a_{2n} + (-1)^n = (a_{n+2} + a_n) a_{n-1}$
- (x) $a_{2n+1} - (-1)^n = (a_{n+2} + a_n) a_n$.

4. Prove or disprove the following statements

- (i) $\sum_{k=0}^{n-1} (k+1) = \frac{1}{2} n(n-1)$
- (ii) $5^n + n + 1$ is divisible by 7 for all $n \geq 1$
- (iii) $3^n \geq n^3$ for all $n \in \mathbb{N}$.

5. If $x \in \mathbb{R} - \{1, -1\}$, prove that

$$(1+x^2)(1+x^4) \cdots (1+x^{2^n}) = \frac{1-x^{2^{n+1}}}{1-x^2}$$

using the principle of induction, for $n \in \mathbb{N}$.

Hence prove that

$$\lim_{x \rightarrow 1} \frac{1-x^{2^{n+1}}}{1-x^2} \text{ and } \lim_{x \rightarrow -1} \frac{1-x^{2^{n+1}}}{1-x^2} \text{ both exist and are equal to } 2^n.$$

6. Using induction prove that $(3 + \sqrt{5})^n + (3 - \sqrt{5})^n$ is divisible by 2^n for all $n \in \mathbb{N}$.

7. If n is an odd positive integer, use induction to prove that $n(n^2 - 1)$ is divisible by 24.

8. Find the flaw in the following argument which shows that

$$3 + 5 + \cdots + (2n - 1) = n^2 \text{ for all } n \in \mathbb{N}.$$

Assume that

$$3 + 5 + \cdots + (2k - 1) = k^2 \tag{4.5}$$

for some $k \in \mathbb{N}$. Then

$$\begin{aligned} 3 + 5 + \cdots + (2k - 1) + (2k + 1) &= k^2 + 2k + 1 \text{ using()} \\ &= (k + 1)^2 \end{aligned}$$

Hence the statement holds for $n = k + 1$. Thus, by the principle of induction the statement holds for all $n \in \mathbb{N}$.

9. Find the flaw in the following proof by induction of the statement
“All numbers in a set of n natural numbers are equal.”

Clearly the statement is true for $n = 1$. Suppose that the result holds for $n = k$. Let $\{a_1, a_2, \dots, a_{k+1}\}$ be any set consisting of $k + 1$ natural numbers. By hypothesis, all the members of the set $\{a_1, a_2, \dots, a_k\}$ consisting of k elements are equal, i.e. $a_1 = a_2 = \cdots = a_k$. Similarly, all members of the set $\{a_2, a_3, \dots, a_{k+1}\}$ consisting of k elements are equal, i.e. $a_2 = a_3 = \cdots = a_{k+1}$.

Hence $a_1 = a_2 = \cdots = a_{k+1}$.

Thus result holds for $n = k + 1$.

By induction, the result holds for all natural numbers n .

10. A chocolate costs Rs 7 and a toffee costs Rs 3. Show by using the principle of induction that any amount, in exact rupees exceeding Rs 11 can be spent in buying chocolates and sweets.

4.4 Greatest Common Divisor

If a, b are two integers and d is an integer which divides a as well as b , then d is said to be a common divisor of a and b . The greatest of all the common divisors of a and b is said to be the greatest common divisor of a and b . For example consider $a = 24, b = 36$.

Common divisors of a and b are $\pm 1, \pm 2, \pm 3, \pm 4, \pm 6$ and ± 12 . Moreover, $-6 < -4 < -3 < -2 < -1 < 2 < 3 < 4 < 6 < 12$. Thus 12 is the greatest common divisor of 24 and 36. Formally we have the following:

Definition 4.3. (Greatest common divisor):

Let $a, b \in \mathbb{Z}$ such that not both are zero. If $d \in \mathbb{Z}$ is the largest common divisor of a and b , then we say that d is the greatest common divisor of a and b .

Symbolically, $d \in \mathbb{Z}$ is called the greatest common divisor of a and b if

(i) $d|a$ and $d|b$.

(ii) if $c \in \mathbb{Z}$ such that $c|a$ and $c|b$ then $c \leq d$.

We write d as $\gcd(a, b)$ or (a, b) .

Since 1 divides every integer,

\therefore by definition of the greatest common divisor,

$$\gcd(a, b) \geq 1$$

If $a, b \in \mathbb{Z}$, not both zero, then $\gcd(a, b)$ is unique, for if d_1, d_2 are two \gcd 's of a and b , then $d_1 \leq d_2$ as d_1 is a common divisor and d_2 is a \gcd .

Similarly $d_2 \leq d_1$ so that $d_1 = d_2$.

Definition 4.4. (Relatively prime): Two integers are said to be relatively prime if their greatest common divisor is 1.

Relatively prime integers are also called co-prime.

Theorem 4.7. Let $d = \gcd(a, b)$. If $a = da'$ and $b = db'$ for some $a', b' \in \mathbb{Z}$, then a' and b' are relatively prime.

Proof: Let $\gcd(a', b') = k \therefore k > 0$

Then $k \mid a'$ and $k \mid b'$.

Let $a' = ka''$ and $b' = kb''$ for some $a'', b'' \in \mathbb{Z}$

$a = da' = dka'' \therefore dk \mid a$

Similarly $dk \mid b$

By definition of \gcd ,

$$dk \leq d$$

$$\Rightarrow k \leq 1$$

Since $k > 0 \therefore k=1$.

Hence $\gcd(a', b') = 1$ □

The following theorem gives a characterization of the \gcd .

Theorem 4.8. Let a, b be integers, not both zero. Then the following statements are equivalent:

(i) $d = \gcd(a, b)$.

(ii) $d \mid a$ and $d \mid b$. If $d' \in \mathbb{Z}$ such that $d' \mid a$ and $d' \mid b$ then $d' \mid d$.

Proof: (i) \Rightarrow (ii)

Let $d = \gcd(a, b)$

Then $d \mid a$ and $d \mid b$ by definition of \gcd .

Let $d' \in \mathbb{Z}$ such that $d' \mid a$ and $d' \mid b$.

$d \mid a$ and $d \mid b$

$$\Rightarrow a = da', b = db', \gcd(a', b') = 1$$

for some $a', b' \in \mathbb{Z}$.

Now $d' \mid a, d' \mid b$

$$\Rightarrow a = d'm, b = d'n \text{ for some } m, n \in \mathbb{Z}$$

Thus $da' = d'm, db' = d'n$

$$\therefore da'b' = d'b'm. \text{ Also } da'b' = d'na' \dots (1)$$

so that $d'b'm = d'na', \therefore b'm = na'$

$$b' \mid mb' \Rightarrow b' \mid na' \Rightarrow b' \mid n \therefore \gcd(a', b') = 1$$

$\therefore n = b'x$ for some $x \in \mathbb{Z}$.

Substituting in (1),

$$da'b' = d'b'xa'$$

$$\Rightarrow d = d'x$$

$$\Rightarrow d' \mid d.$$

(ii) \Rightarrow (i) obvious, since $d' \mid d \Rightarrow d' \leq d$. □

Theorem 4.9. Let $a, b \in \mathbb{Z}$, not both zero. Then $\gcd(a, b)$ exists and is unique. Also, there exists integers m and n such that

$$am + bn = \gcd(a, b)$$

Proof: Existence

Let $A = \{ax + by \mid x, y \in \mathbb{Z}, ax + by > 0\}$.

Since $aa + b0 \in A \therefore A \neq \phi$ if $a \neq 0$ otherwise $a.0 + b.b \in A$.

Thus A is a non-empty subset of the set of positive integers, so by the well ordering principle A has a least element, say d . By definition of A , $\exists m, n \in \mathbb{Z}$ such that

$$am + bn = d \quad (4.6)$$

We prove that $d = \gcd(a, b)$.

By division algorithm, applied to a and d , $\exists q, r \in \mathbb{Z}$ such that $a = dq + r$, $0 \leq r < d$.

$$\begin{aligned} r &= a - dq \\ &= a - (am + bn)q \\ &= a(1 - mq) + b(-nq) \end{aligned}$$

Also $r \geq 0$. If $r \neq 0$, then $r > 0$ and $r = a(1-mq)+b(-nq)$ so that $r \in A$. But $r < d$, which contradicts the fact that d is the smallest element of A . Hence $r = 0$.

$\therefore a = dq$ so that $d|a$.

Similarly $d|b$.

Thus d is a common divisor of a and b . If $d' \in \mathbb{Z}$ such that $d'|a$ and $d'|b$,

then $d'|(ma + nb)$

$\Rightarrow d'|d$

$\Rightarrow d' \leq d$.

Hence $d = \gcd(a, b)$

and $am + bn = \gcd(a, b)$.

Uniqueness

It has already been proved that if $\gcd(a, b)$ exists, it is unique.

Theorem 4.9 tells us that two integers always have a unique greatest common divisor. However, it does not give any method to determine it. We shall give an algorithm to determine it. \square

Corollary 4.10. *Given integers a, b and c , the equation $ax + by = c$ has integral solutions if and only if $\gcd(a, b)|c$.*

Proof: Let $d = \gcd(a, b)$

Suppose the given equation has integral solution x_0, y_0 . Then

$$ax_0 + by_0 = c$$

Also $d|a$ and $d|b$.

$\therefore d|(ax_0 + by_0) \Rightarrow d|c$.

Conversely, let $d|c$.

$\therefore c = kd$ for some $k \in \mathbb{Z}$.

Since $d = \gcd(a, b)$.

\therefore By the above theorem there exists integers m and n such that

$$am + bn = d$$

Hence $akm + bkn = kd$ or $ax_0 + by_0 = c$
 where $x_0 = km$, $y_0 = kn$. Clearly $x_0, y_0 \in \mathbb{Z}$.

Thus the given equation has a solution. \square

Remark 4.2. *The integers m and n obtained above need not be unique, for if*

$$\begin{aligned} d &= am + bn \\ &= am + bn + kab - kab, k \in \mathbb{Z} \\ &= a(m + kb) + b(n - ka) \end{aligned}$$

Giving different values to k , other values of m and n are obtained.

The gcd of any two integers have the following properties.

Theorem 4.11. *If $a, b, c \in \mathbb{Z}$, then*

1. $gcd(a, b) = gcd(b, a)$
2. $gcd(a, 0) = |a|$
3. $gcd(a, 1) = 1$
4. $gcd(a, b) = gcd(|a|, |b|)$
5. $gcd(a, gcd(b, c)) = gcd(gcd(a, b), c)$
6. $gcd(ac, bc) = |c|gcd(a, b)$

Proof: Left to the reader. \square

Theorem 4.12. *If p is a prime and $a, b \in \mathbb{Z}$ such that if $p|ab$, then $p|a$ or $p|b$.*

Proof: The result holds trivially when $a = 0$ or $b = 0$.

Suppose that $a \neq 0$ and $b \neq 0$. If $p|a$ result is proved.

Suppose that $p \nmid a$. We prove that $p|b$.

We assert that $gcd(p, a) = 1$.

Let $d = gcd(p, a)$.

Then $d|p$ and $d|a$

$d|p$, p is prime $\Rightarrow d = 1, p$

If $d = p$ then $d|a \Rightarrow p|a$ which contradicts the assumption that $p \nmid a$.

Hence $d=1$.

Since $gcd(p, a) = 1$, therefore, by Theorem 4.9, there exists integers m and n such that

$$\begin{aligned} am + pn &= 1 \\ \Rightarrow am + pn &= b \quad (\text{multiplying by } b) \end{aligned}$$

Now $p|ab$, $p|p$

$\Rightarrow p|(am + pn)$

$\Rightarrow p|b$

Hence proved. \square

This result can be extended to product of n integers.

Theorem 4.13. Let p be a prime and a_1, a_2, \dots, a_n be integers such that $p|a_1a_2 \cdots a_n$. Then $p|a_i$ for some $i = 1, 2, \dots, n$.

Proof: Prove by using induction on n , using Theorem 4.12.

Theorem 4.14. If $a, b, c \in \mathbb{Z}$ such that $a|bc$ and $\gcd(a, b) = 1$, then $a|c$.

Proof: $\gcd(a, b) = 1$
 $\Rightarrow \exists m, n \in \mathbb{Z}$ such that

$$am + bn = 1$$

Multiplying by c , we get

$$acm + bcn = c$$

Now $a|a, a|bc$
 $\Rightarrow a|(acm + bcn)$
 $\Rightarrow a|c$. □

In Theorem 4.14, it is essential for $\gcd(a, b) = 1$. For example, if $a = 6$, $b = 3$, $c = 4$. Then $a|bc$, but $a \nmid c$.
 Note that $\gcd(a, b) = 3 \neq 1$.

Lemma 4.15. If $a, b \in \mathbb{Z}$ are not both zero, and $q, r \in \mathbb{Z}$ such that $a = bq + r$, then $\gcd(a, b) = \gcd(b, r)$.

Proof: Let $d_1 = \gcd(a, b)$, $d_2 = \gcd(b, r)$

$d_1 = \gcd(a, b)$
 $\Rightarrow d_1|a$ and $d_1|b$
 $\Rightarrow d_1|(a - bq)$
 $\Rightarrow d_1|r$

Thus $d_1|r$ and $d_1|b$.
 $\Rightarrow d_1 \leq d_2 \quad \dots (1)$

Now $d_2 = \gcd(b, r)$
 $\Rightarrow d_2|b$ and $d_2|r$
 $\Rightarrow d_2|(bq + r)$
 $\Rightarrow d_2|a$

Thus $d_2|a$ and $d_2|r$
 $\Rightarrow d_2 \leq d_1 \quad \dots (2)$

Using (1) and (2), we get

$$d_1 = d_2. \quad \square$$

We now give an algorithm to determine the gcd of two integers.

Euclidean Algorithm

Let a and b be integers, not both zero.

Since $\gcd(a, b) = \gcd(|a|, |b|)$, so, without any loss of generality, we may assume $a > 0$ and $b > 0$.

By division algorithm, $\exists q_1, r_1 \in \mathbb{Z}$
 such that

$$a = bq_1 + r_1, 0 \leq r_1 < b$$

and $\gcd(a, b) = \gcd(b, r_1), \dots$ (by the above lemma)

If $r_1 \neq 0$, then $\exists q_2, r_2 \in \mathbb{Z}$ such that

$$b = q_2r_1 + r_2, 0 \leq r_2 < r_1$$

and $\gcd(b, r_1) = \gcd(r_1, r_2)$

If $r_2 \neq 0$, then $\exists q_3, r_3 \in \mathbb{Z}$ such that

$$r_1 = q_3 r_2 + r_3, 0 \leq r_3 < r_2$$

and $\gcd(r_1, r_2) = \gcd(r_2, r_3)$

Continue this process. As $r_1 > r_2 > r_3 > \dots$

After a finite number of steps, the remainder $r_{k+1} = 0$ for some integer $k \geq 0$.

Then $r_{k-1} = q_{k+1} r_k + 0$, and $\gcd(r_{k-1}, r_k) = r_k$.

Also $\gcd(a, b) = \gcd(b, r_1) = \gcd(r_1, r_2) = \gcd(r_2, r_3) = \dots = \gcd(r_{k-1}, r_k) = r_k$

Thus $\gcd(a, b) = r_k$, the last non-zero remainder.

Example 4.5. Find the gcd of 595 and 205. Also find integers m and n such that $\gcd(595, 205) = 595m + 205n$.

Solution:

Let $a = 595$, $b = 205$, Dividing a by b , we get

$$\underline{595} = 2 \times \underline{205} + 185 \quad (4.7)$$

Now divide 205 by 185 (= r_1 say)

$$\underline{205} = 1 \times \underline{185} + 20 \quad (4.8)$$

Now divide 185 by 20 (= r_2 say)

$$\underline{185} = 9 \times \underline{20} + 5 \quad (4.9)$$

Now divide 20 by 5 (= r_3 say)

$$\underline{20} = 4 \times \underline{5} \quad (4.10)$$

The gcd is the last nonzero remainder, namely r_3

$$\gcd(595, 205) = 5 = d \text{ (say)}$$

Now, we express d as a linear combination of a and b . To do this, we can do back substitution from (4.9), backwards to equation 4.7 and get the desired result.

But the calculations can be done in another way, in which the backward substitution is not required.

Write the given numbers a and b as a linear combination of a and b , as the first two equations. At every step, the remainder r_i is expressed as a linear combination of a and b .

$$595 = 1a + 0b \quad \dots (1)$$

$$205 = 0a + 1b \quad \dots (2)$$

Apply operations on (1) and (2) such that L.H.S. is the remainder obtained on dividing 595 by 205.

Thus (1) $-2 \times$ (2) \Rightarrow

$$185 = a - 2b \quad \dots (3)$$

This process is repeated till the remainder obtained is zero.

Applying (2)-(3), we get

$$20 = -a + 3b \quad \dots (4)$$

(3) $-9 \times$ (4) \Rightarrow

$$5 = 10a - 29b \quad \dots (5)$$

$$(4) -4 \times (5) \Rightarrow$$

$$0 = -41a + 119b \dots (6)$$

The least non-zero remainder is the greatest common divisor. It is given by equation (5). Thus

$$\gcd=5 \text{ and } 10a - 29b = 5 \text{ or } ma + nb = 5.$$

where $m = 10, n = -29$.

This expression for the gcd is not unique. If $d = \gcd(a, b)$ and m, n are integers such that

$$d = ma + nb \dots (*)$$

Then $d = ma + nb + kab - kab$ where $k \in \mathbb{Z}$.

$$= a(m + bk) + b(n - ka).$$

Thus $m_1 = m + bk, n_1 = n - ka$, are also integers which satisfy

$$d = m_1a + n_1b$$

For each integer k , we get values for m_1 and n_1 . Hence there are infinitely many values for m_1 and n_1 .

Working Rule

To obtain \gcd of two numbers 'a' and 'b' and to express the \gcd in terms of 'a' and 'b'.

Step 1: Without any loss of generality we can assume $a, b > 0$ and $a > b$

Step 2: Express a and b in terms of 'a' and 'b'

i.e

$$a = 1a + 0b \tag{4.11}$$

$$b = 0a + 1b \tag{4.12}$$

choose q_1 such that $a - bq_1$ is the remainder obtained on dividing 'a' by 'b'. Thus apply (4.11) - q_1 (4.12), to get

$$r_1 = a - bq_1 \tag{4.13}$$

Clearly

$$0 \leq r_1 < b.$$

Choose q_2 such that $b - q_2r_1$ is the remainder obtained on dividing b by r_1 .

then applying

$$\begin{aligned} (4.12) - q_2(4.13) & \text{ gives} \\ r_2 = b - q_2r_1 & = -q_2a + b + bq_1q_2 \\ r_2 = m_2a + n_2b & \quad \text{(say)} \end{aligned}$$

$$\therefore r_2 = m_2a + n_2b. \tag{4.14}$$

Thus the remainder at each step is expressed in terms of a and b . Continue this process. The last non-zero remainder gives an expression of the gcd in terms of a and b .

Step 3: Adjust the signs of m and n so to get the actual signs of numbers.

Example 4.6. Find the gcd of 154 and 260.

Also express the gcd as a combination of 154 and 260. Is this expression unique? If not, obtain two such expressions.

Let $a = 260, b = 154.$

Now

$$a = 260 = 1a + 0b \quad (4.15)$$

$$b = 154 = 0a + 1b \quad (4.16)$$

Apply

$$(4.15) - (4.16),$$

$$r_1 = 106 = a - b \quad (4.17)$$

Apply

$$(4.16) - (4.17),$$

$$r_2 = 48 = -a + 2b \quad (4.18)$$

Apply

$$(4.17) - 2 \times (4.18),$$

$$r_3 = 10 = 3a - 5b \quad (4.19)$$

Apply

$$(4.18) - 4 \times (4.19),$$

$$r_4 = 8 = -13a - 22b \quad (4.20)$$

Apply

$$(4.19) - (4.20),$$

$$r_5 = 2 = 16a - 27b \quad (4.21)$$

$$\text{Apply } (4.20) - 4 \times (4.21),$$

$$r_6 = 0 = -61a + 130b$$

Since $r_6 = 0$, therefore the last non-zero values of r_i , namely r_5 is the *gcd*. Thus

$$r_5 = 2 = \gcd(a, b) = 16a - 27b = 16 \times 260 - 27 \times 154$$

$$\therefore r_5 = 16 \times 260 + (-27) \times (-154)$$

So $m = 16, n = -27$. This expression of the *gcd* as a combination of the numbers a and b is not unique.

In fact $2 = 16a - 27b + ab - ab$

$$= (16 + b)a + (-27 - a)b$$

Thus we have

$$2 = m_1a + n_1b$$

$$2 = m_2a + n_2b$$

where $m_1 = 16, n_1 = -27$

$$m_2 = 16 + b = 170$$

$$n_2 = (-27 - a) = -287.$$

Problem 4.9.

(i) Find the *gcd* of 3719 and 8146.

Express the *gcd* in the form $3719m + 8146n$, for $m, n \in \mathbb{Z}$.

Are the values of m and n unique?

If not, can you find 30 sets of values?

(ii) Also express the *gcd* of -3719 and 8146 as $m(-3719) + n(8146)$, for $m, n \in \mathbb{Z}$.

Solution:(i) Let $b = 3719$, $a = 8146$

$$a = 8146 = 1a + 0b \quad (4.22)$$

$$b = 3719 = 0a + 1b \quad (4.23)$$

Apply (4.22) $- 2 \times$ (4.23)

$$r_1 = 708 = a - 2b \quad (4.24)$$

Apply (4.23) $- 4 \times$ (4.24)

$$r_2 = 179 = -5a + 11b \quad (4.25)$$

Apply (4.24) $- 3 \times$ (4.25)

$$r_3 = 171 = 16a - 35b \quad (4.26)$$

Apply (4.25) $- \times$ (4.26)

$$r_4 = 8 = -21a + 46b \quad (4.27)$$

Apply (4.26) $- 21 \times$ (4.27)

$$r_5 = 3 = 457a - 1001b \quad (4.28)$$

Apply (4.27) $- 2 \times$ (4.28)

$$r_6 = 2 = -935a + 2048b \quad (4.29)$$

Apply (4.28) $-$ (4.29)

$$r_7 = 1 = 1392a - 3049b \quad (4.30)$$

Thus $\gcd(a, b) = 1$, and $ma + nb = 1$, where $m = 1392$, $n = -3049$.
Values of m and n are not unique.

$$1 = ma + nb$$

$$\begin{aligned} &= ma + nb + kab - kab, \text{ for all } k \in \mathbb{Z} \\ &= (m + kb)a + (n - ka)b \\ &= m_k a + n_k b \end{aligned}$$

where $m_k = m + kb$, $n_k = n - ka$.

Giving different values to k , we get different values of m_k and n_k .

Thirty sets of values can be obtained by giving 30 values to k .

(ii) In (i) we have proved

$$1 = ma + nb$$

$$\begin{aligned} &= m \times 8146 + n \times 3719 \\ &= m \times 8146 - n \times (-3719) \\ &= m \times 8146 + (-n) \times (-3719) \\ &= m_1 \times 8146 + n_1 \times (-3719) \end{aligned}$$

Here $m_1 = m = 1392$, $n_1 = -n = 3049$.

4.5 Least Common Multiple

If a and b are integers and l is an integer which is multiple of a as well as b then l is called a common multiple of a and b . The least of all the positive multiples of a and b is called the least common multiple of a and b .

Consider $a = 9$, $b = 15$

Multiples of a are $\pm 9, \pm 18, \pm 27, \pm 36, \pm 45, \dots$

Multiples of b are $\pm 15, \pm 30, \pm 45, \pm 60, \pm 75, \pm 90, \dots$

Common multiples of a and b are $\pm 45, \pm 90, \dots$

The smallest of all the positive multiple of ' a ' and ' b ' is 45.

Thus, the least common multiple of 9 and 15 is 45.

Formally, we have the following definition.

Definition 4.5. If a, b are non-zero integers, then a positive integer l is called the least common multiple (lcm) of a and b if

(i) $a|l$ and $b|l$

(ii) If m is a positive integer such that $a|m$ and $b|m$ then $l \leq m$.

We write $l = lcm(a, b)$.

Note that just like the gcd , lcm is always positive by definition.

Example 4.7.

(1) The least common multiple of 18 and 24 is 72.

(2) $lcm(8, -12) = 24$

(3) $lcm(-15, -12) = 60$

Since $|ab|$ is a common multiple of a and b , therefore least common multiple always exists and is less than or equal to $|ab|$.

The following theorem gives a characterization of the lcm .

Theorem 4.16. Let a, b be two integers then the following statements are equivalent

(1) $l = lcm(a, b)$

(2) If l is a positive integer such that $a|l$ and $b|l$. If l' is any positive integer such that $a|l'$ and $b|l'$ then $l|l'$.

Proof: Left as an exercise. □

Theorem 4.17. The gcd of any two integers always divides their lcm .

Proof: Let $a, b \in \mathbb{Z}$ and let $d = gcd(a, b)$ and $l = lcm(a, b)$.

Then $d|a$ and $d|b$. Also $a|l$ and $b|l$.

Now $d|a$ and $a|l$, so that $d|l$.

Hence proved. □

The next theorem gives a relationship between the gcd and lcm of two numbers.

Theorem 4.18. If $a, b \in \mathbb{Z}$, not both zero, then $gcd(a, b) \times lcm(a, b) = |ab|$.

Proof: Let $d = gcd(a, b)$, $l = lcm(a, b)$.

Then $d|a$ and $d|b$. Also by Theorem 4.9 there exist $m, n \in \mathbb{Z}$ such that

$$ma + nb = d \tag{4.31}$$

$$l = \text{lcm}(a, b) \Rightarrow l = ax, l = by \quad (4.32)$$

for some integers $x, y \in \mathbb{Z}$.

Multiplying (4.31) by l , we get

$$\begin{aligned} dl &= mal + nbl \\ &= maby + nbax \quad \text{using (4.32)} \\ &= ab(my + nx) \end{aligned}$$

Hence

$$ab|dl \quad (4.33)$$

Also $d|a$ and $d|b$

$$\Rightarrow a = dx_1, \quad b = dy_1 \quad (4.34)$$

for some $x_1, y_1 \in \mathbb{Z}$. Thus

$$ab = d^2x_1y_1 = (dx_1y_1)d \quad (4.35)$$

Since $ay_1 = dx_1y_1 = bx_1 \quad \dots$ using (4.34).

Thus $a|dx_1y_1$ and $b|dx_1y_1$.

So, by definition of $\text{lcm} \quad l|dx_1y_1$

Let $dx_1y_1 = lk$ for some $k \in \mathbb{Z}$

$$\Rightarrow d^2x_1y_1 = dlk$$

$$\Rightarrow ab = dlk \quad \text{using (4.35)}.$$

$$\Rightarrow dl|ab. \quad (4.36)$$

$$(4.33) \text{ and } (4.36) \Rightarrow dl = \pm ab = |ab| \quad \because l > 0, d > 0.$$

□

Euclid had proved that the number of primes is infinite. In order to prove this result, we shall prove a lemma first.

Lemma 4.19. *Given any natural number $n > 1$, there exists a prime p such that $p|n$.*

Proof: We shall prove this result by contradiction. Let, if possible, the result does not hold.

Let $S = \{n \in \mathbb{N} \mid n > 1 \text{ and } n \text{ is not divisible by any prime}\}$.

Then $S \neq \emptyset$ by assumption. By the Well Ordering Principle, S being a non-empty subset of \mathbb{N} , it has a least element, say s . Also s is not divisible by any prime, Since $s|s$, therefore s can not be prime. Thus there exists some $k \in \mathbb{N}$ such that $1 < k < s$ and $k|s$. Since s is the least element of S , therefore k must be divisible by some prime p .

Now, $p|k$ and $k|s$ so that $p|s$ which contradicts that $s \in S$.

Hence our assumption is wrong, so that there exists a prime p which divides n . □

Though the prime numbers are very special, but they are infinite in number. This result, due to Euclid, will be proved by using the fundamental theorem of arithmetic.

Theorem 4.20. *(Euclid's theorem)*
There exists infinitely many primes.

Proof: Let, if possible, there be only a finite number of primes, say m . Let them be $p_1, p_2, p_3, \dots, p_m$.

Let $n = p_1 p_2 p_3 \cdots p_m + 1$. Then $n \in \mathbb{N}$ and $n > 1$.

But the above lemma there is a prime p such that $p|n$.

But $p_i \nmid n$ for any $i = 1, 2, 3, \dots, m$. So that none of the primes divides n , which is a contradiction to the fact that there is a prime dividing n .

Hence our assumption is wrong, so that the number of primes is infinite. \square

One of the basic result of numbers is that every positive integer greater than 1 can be expressed as a product of prime numbers in essentially one way.

This also brings out the importance of prime numbers as building blocks of the system of integers from the point of view of factorization. Thus we have the following theorem.

Theorem 4.21. (*Fundamental theorem of arithmetic*)

Every integer $n \geq 2$ is either prime or is expressible as a product of finitely many prime numbers. Moreover, such an expression is unique except for the order of the factors.

Proof: *Existence*

Let $S = \{m \in \mathbb{Z} \mid m > 1, m \text{ is not prime, } m \text{ is not expressible as a product of primes}\}$.

If, $S = \emptyset$, then the proof is complete.

If $S \neq \emptyset$, then S is a non-empty subset of the set of natural numbers, so that by the Well Ordering Principle, S has a least element, say s .

Since $s \in S$ is not prime and $s > 1$, therefore it must have a divisor other than 1 and s . Let s_1 be a positive divisor of s such that $1 < s_1 < s$. Then there exists a positive integer s_2 such that $1 < s_2 < s$ and $s = s_1 s_2$. Since s is the least element of S , therefore $s_1 \notin S$ and $s_2 \notin S$. Therefore, either s_1, s_2 are primes or they can be expressed as a product of primes. In either case, s is expressible as a product of primes.

This contradicts the fact that $s \in S$. Hence $S = \emptyset$.

This proves that every integer $n > 1$ is either prime or is a product of primes.

Uniqueness

We shall use induction on n .

If $n = 2$, then trivially, expression is unique.

Assume the uniqueness for all integers m such that $2 \leq m < k$.

Either k is prime, in which case result holds.

If k is not a prime, then let it have two expression as a product of primes, say

$$k = p_1 p_2 \cdots p_s \quad (1)$$

$$k = q_1 q_2 \cdots q_t \quad (2)$$

where $p_i, q_j, 1 \leq i \leq s, 1 \leq j \leq t$ are all prime numbers. Also k is not prime $\Rightarrow s \geq 2$ and $t \geq 2$.

Then we shall prove that $s = t$ and $q'_i s$ are a rearrangement of the $p'_j s$.

$$(1) \text{ and } (2) \Rightarrow p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t \quad (3)$$

Now $p_1 \mid (p_1 p_2 \cdots p_s)$

$$\Rightarrow p_1 \mid q_1 q_2 \cdots q_t$$

$$\Rightarrow p_1 \mid q_j \text{ for some } j, 1 \leq j \leq t \text{ using Theorem 4.13}$$

$\Rightarrow p_1 = q_j \quad \because q_j$ is prime.

Since $p_1 \neq 0$, therefore by cancellation law in (3), we get

$$p_2 p_3 \cdots p_s = q_1 \cdots q_{j-1} q_{j+1} \cdots q_t = k_1 \quad (\text{say}) \quad (4).$$

Since $1 < k_1 < k$, therefore by the induction hypothesis, the two expressions of k_1 in (4) are identical, except for the order of the prime factors. Hence $s-1 = t-1$ and $q_1, \dots, q_{j-1}, q_{j+1}, \dots, q_t$ is just a rearrangement of p_1, p_2, \dots, p_s .

Thus $s = t$ and q_1, q_2, \dots, q_t is a rearrangement of p_1, p_2, \dots, p_s . By the principle of induction the result holds for all $n \geq 2$. \square

Corollary 4.22. *Let $n \in \mathbb{Z}$ such that $|n| \geq 2$. Then n is either prime or is expressible as a product of a unit and finitely many prime numbers. Moreover such an expression is unique except as to the order in which the factors occur.*

Proof: Two cases arise.

Case 1. $n \geq 2$.

Result follows from the above theorem.

Case 2. $n < 0$, $|n| \geq 2$.

Let $n = -m$, where $m > 0$

then $|m| \geq 2$. Also $n = (-1)m$.

Applying the above theorem to m , we get the result, as (-1) is a unit. \square

Problem 4.10. *If $a, b, c \in \mathbb{Z}$ such that a, b are relatively prime and $a|c$ and $b|c$, then $ab|c$. What happens when a and b are not relatively prime?*

Solution: We know that $\gcd(a, b) \operatorname{lcm}(a, b) = |ab|$

$$\Rightarrow \operatorname{lcm}(a, b) = |ab| \quad \because \gcd(a, b) = 1.$$

Let $l = \operatorname{lcm}(a, b)$

$$\therefore l = |ab|$$

Since $a|c$ and $b|c$

$\therefore l|c$ by definition of lcm .

$$\Rightarrow |ab| |c$$

$$\Rightarrow ab |c$$

Take $a = 8$, $b = 12$, $c = 24$.

Then $a|c, b|c$, but $ab \nmid c$. Note that $\gcd(a, b) = 4$.

4.6 Exercise

1. For any natural number n , in any set of n consecutive integers, one of the integers is always divisible by n .
2. If $a, b \in \mathbb{Z}$ such that $a | b$ then prove that $\gcd(a, b) = |a|$.
3. For any two integers a and b prove that $\gcd(a, b) = \gcd(|a|, |b|)$.
4. Given integers d, a, b , suppose there exists integers m and n such that

$$ma + nb = d$$

then prove that

- (i) $\gcd(a, b)$ divides d .
- (ii) $\gcd(m, n)$ divides d .
- (iii) $\gcd(a, n)$ divides d .
- (iv) $\gcd(m, b)$ divides d .

5. Let $a, b, m, n \in \mathbb{Z}$ such that $am + bn = 1$, then $\gcd(a, b) = \gcd(m, n) = \gcd(a, n) = \gcd(b, m) = 1$
6. If $a, b \in \mathbb{Z}$ show that $\gcd(a, a + b) = \gcd(a, b)$.
7. If a, b are integers then $\gcd(\gcd(a, b), a) = \gcd(a, b)$.
8. Prove that any two consecutive integers are always relatively prime.
9. If a, b are relatively prime integers, then
- (i) $\gcd(a, a + b) = 1$.
 - (ii) $\gcd(a + b, a - b) = 1$ (or) 2 .
 - (iii) $\gcd(a + b, a^2 + b^2) = 1$ (or) 2 .
 - (iv) $\gcd(a^n, b) = 1, n \in \mathbb{N}$.
10. If $d, a, b \in \mathbb{Z}$ and d is an odd integer such that $d \mid (a + b)$ and $d \mid (a - b)$ then $d \mid \gcd(a, b)$.
11. Find the \gcd of a and b and express it in the form $ma + nb$ for $m, n \in \mathbb{Z}$.
- (i) $a = 143, b = 247$
 - (ii) $a = -143, b = 247$
 - (iii) $a = 314, b = 159$
 - (iv) $a = -314, b = -159$
 - (v) $a = 4144, b = 7696$
 - (vi) $a = 4144, b = -7696$
 - (vii) $a = 394, b = -562$
12. Find the \gcd of a and b . If $d = \gcd(a, b)$, find three solutions in integers of $d = ma + nb$.
- (i) $a = 243, b = 189$
 - (ii) $a = 741, b = 1079$
 - (iii) $a = 4453, b = 1314$.
13. Find integers m and n such that
- $$159m + 314n = 7.$$
- Are m and n unique? If not, find another pair also.
14. Find integers m and n such that
- $$9m + 11n = 4.$$
- Show that $m = 11k - 2, n = 2 - 9k$ for some integers k .
15. Show that there do not exist any integers m and n such that
- $$219m + 153n = 5.$$
16. Let $m, n \in \mathbb{Z}$. Prove that $10m + n$ is divisible by 7 if and only if $m + 5n$ is divisible by 7.
17. For any integer m , prove that

$$\gcd(m, m + 2) = \begin{cases} 1 & \text{if } m \text{ is odd} \\ 2 & \text{if } m \text{ is even.} \end{cases}$$

18. For $m \in \mathbb{Z}$, are the following pairs co-prime

- (i) $7m + 1, 6m + 1$
- (ii) $5m + 3, 3m + 2$
- (iii) $9m + 4, 11m + 5$
- (iv) $7m + 4, 5m + 2$.

4.7 Congruence Relation

Modular Arithmetic

If a 12-hour clock shows 10, then after 6 hours it should show $10 + 6 = 16$.

But the time shown by it is $16 - 12 = 4$. This is because multiples of 12 are subtracted to get the actual time.

If it is a thursday on 16 July then after 30 days it is 15 August. since the days of week repeat after every 7 days, and $30 = 7 \times 4 + 2$, therefore on 15 August, the day will be the one which is two days after thursday, that is, Saturday. Similarly if your birthday falls on Saturday in 2009 then in 2010 it will be on Sunday as $365 = 7 \times 52 + 1$.

Let $n > 1$ be any integer and $a \in \mathbb{Z}$. By division algorithm, there exists unique $q, r \in \mathbb{Z}$ such that $a = nq + r$, $0 \leq r < n$.

If multiples of n redundant, we reduce a by multiples of n and we say that $r = a$ mod n . Thus, we define $a \text{ mod } n$ as follows:

Definition 4.6. If $n > 1$ is any integer and $a \in \mathbb{Z}$, we define $a \text{ mod } n$ as the remainder r obtained on dividing a by n and we write $a \text{ mod } n = r$. Clearly $0 \leq r < n$, so that $a \text{ mod } n \geq 0$.

Example 4.8.

1. $54 \text{ mod } 8 = 6$, because on dividing 54 by 8, the remainder is 6.
2. $-54 \text{ mod } 8 = 2$, because on dividing -54 by 8, the remainder is 2. We note that the remainder is positive.
3. $59 \text{ mod } 9 = 5$, $77 \text{ mod } 9 = 5$. Also observe that $77 - 59 = 18$ and $9 \mid 18$.
4. $a \text{ mod } 5 = 0$, whenever $5 \mid a$.

Remark 4.3. If $a, b \in \mathbb{Z}$, then $a \text{ mod } n = b \text{ mod } n \Leftrightarrow n \mid (a - b)$ Suppose $a \text{ mod } n = b \text{ mod } n = r$, then $a = nq_1 + r$, $b = nq_2 + r$ for some $q_1, q_2 \in \mathbb{Z}$. Therefore $a - b = n(q_1 - q_2) \Rightarrow n \mid (a - b)$.

Conversely, let $n \mid (a - b)$.

$$a = nq_1 + r_1, 0 \leq r_1 < n$$

$$b = nq_2 + r_2, 0 \leq r_2 < n$$

$$\text{Hence } a - b = n(q_1 - q_2) + r_1 - r_2.$$

Since $n \mid (a - b)$ and $n \mid n(q_1 - q_2)$, therefore $n \mid (r_1 - r_2) \Rightarrow n \mid |r_1 - r_2|$

But $0 \leq |r_1 - r_2| < n$ so that $|r_1 - r_2| = 0 \Rightarrow r_1 - r_2 = 0 \Rightarrow r_1 = r_2$, therefore $a \text{ mod } n = r_1 = r_2 = b \text{ mod } n$.

Example 4.9.

1. Calculate $283 \bmod 13$, $729 \bmod 13$, $(283 \bmod 13 + 729 \bmod 13) \bmod 13$, $(283 + 729) \bmod 13$.
 Since $283 = 21 \times 13 + 10$
 $\therefore 283 \bmod 13 = 10$.
 Similarly $729 \bmod 13 = 1$
 $(283 \bmod 13 + 729 \bmod 13) \bmod 13$
 $= (10 + 1) \bmod 13$
 $= 11$
 $(283 + 729) \bmod 13$
 $= 1012 \bmod 13$
 $= 11$
2. Calculate $(283 \bmod 13)(729 \bmod 13) \bmod 13$ and $(283 \times 729) \bmod 13$.
 $(283 \bmod 13)(729 \bmod 13) \bmod 13$
 $= (10 \times 1) \bmod 13$
 $= 10$
 $(283 \times 729) \bmod 13 = (206307) \bmod 13 = 10$.

In the above illustration, we observe that $(283+729) \bmod 13 = (283 \bmod 13 + 729 \bmod 13) \bmod 13$, and $(283 \times 729) \bmod 13 = ((283 \bmod 13)(729 \bmod 13)) \bmod 13$.

In general, we have the following result.

Theorem 4.23. If $n > 1$ is the integer then for $a, b \in \mathbb{Z}$,

1. $(a + b) \bmod n = ((a \bmod n) + (b \bmod n)) \bmod n$
2. $ab \bmod n = ((a \bmod n)(b \bmod n)) \bmod n$.

Proof: Let $a \bmod n = r$, $b \bmod n = s$.
 Then there exists $q_1, q_2 \in \mathbb{Z}$ such that

$$a = nq_1 + r$$

$$b = nq_2 + s.$$

1. $(a + b) = n(q_1 + q_2) + r + s$.
 $\therefore (a + b) \bmod n = (r + s) \bmod n = (a \bmod n + b \bmod n) \bmod n$.
2. $ab = n(nq_1q_2 + q_1s + q_2r) + rs$.
 $\therefore (ab) \bmod n = (rs) \bmod n = (a \bmod n)(b \bmod n) \bmod n$. □

This theorem helps us to simplify calculations in modulo n .

Problem 4.11. A college XYZ assigns its students roll numbers. The last three digits of the roll number of a female student born in month m on date b is $69m+2b+1$ and that of a male student is $69m + 2b$. Find the date of birth and sex corresponding to the numbers.

(i) 194

(ii) 074

(iii) 683

Solution: We will express the numbers in the form $69m + 2b + 1$ or $69m + 2b$.

(i) $194 = 69 \times 2 + 56 = 69 \times 2 + 2 \times 28$ Hence $m = 2$, $b = 28$ and the person is male.

Therefore date of birth is 28th Feb and the student is male.

(ii) $074 = 69 \times 1 + 5 = 69 \times 1 + 2 \times 2 + 1$ Hence $m = 1$, $b = 2$ and the person is female.

Therefore date of birth is 2nd Jan and the student is female.

(iii) $683 = 69 \times 9 + 62 = 69 \times 9 + 2 \times 31$ Thus $m = 9$, $b = 31$ and the person is male.

It is incorrect, Since the date of birth is 31 September, which is not possible.

Definition 4.7. Let $n > 1$ be a fixed natural number. If $a, b \in \mathbb{Z}$ we say that a is congruent to b modulo n if and only if $n \mid (a - b)$. We write it as $a \equiv b \pmod{n}$. We read it as ' a is congruent to $b \pmod{n}$ ' and n is called the modulus of the congruence.

Example 4.10.

1. $77 \equiv 59 \pmod{9}$, $\because 9 \mid (77 - 59)$.2. $125 \equiv 136 \pmod{11}$, as $11 \mid (125 - 136)$.

Theorem 4.24. For a fixed integer $m > 0$, the relation $a \equiv b \pmod{m}$ on \mathbb{Z} is an equivalence relation.

Proof: Left to the reader. □

The above relation partitions \mathbb{Z} into mutually disjoint equivalence classes. The class to which an integer ' a ' belongs is called the equivalence class of ' a ' and is denoted by $[a]$ or \bar{a} or $cl(a)$.

Thus $[a] = \{x \in \mathbb{Z} \mid x \equiv a \pmod{m}\}$
 $= \{a + km \mid k \in \mathbb{Z}\}$.

Example 4.11. Find the distinct equivalence classes of the relation congruence modulo 6 on \mathbb{Z} .

Clearly $1 \equiv 7 \pmod{6}$, $1 \equiv 13 \pmod{6}$, etc. . . .

$\therefore [1] = \{1, 7, 13, 19, \dots\}$
 $= \{1 + 6k \mid k \in \mathbb{Z}\}$

$[2] = \{2 + 6k \mid k \in \mathbb{Z}\}$

$[3] = \{3 + 6k \mid k \in \mathbb{Z}\}$

$[4] = \{4 + 6k \mid k \in \mathbb{Z}\}$

$[5] = \{5 + 6k \mid k \in \mathbb{Z}\}$

$[6] = \{6k \mid k \in \mathbb{Z}\} = [0]$

Thus, the distinct equivalence classes are $[0], [1], [2], [3], [4], [5]$.

The following theorems gives some results regarding congruences, which are very useful in manipulations.

Theorem 4.25. *Let $a \equiv b \pmod{m}$ and $x \in \mathbb{Z}$ then*

1. $(a + x) \equiv (b + x) \pmod{m}$
2. $(a - x) \equiv (b - x) \pmod{m}$
3. $ax \equiv bx \pmod{m}$.

Proof:

1. $a \equiv b \pmod{m}$
 $\Rightarrow m \mid (a - b)$
 $\Rightarrow m \mid ((a + x) - (b + x)) \forall x \in \mathbb{Z}$
 $\Rightarrow (a + x) \equiv (b + x) \pmod{m}$.

2. and 3. can be proved similarly.

□

Theorem 4.26. *Let $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$. Then*

1. $(a + c) \equiv (b + d) \pmod{m}$
2. $(a - c) \equiv (b - d) \pmod{m}$
3. $ac \equiv bd \pmod{m}$
4. $(pa + qc) \equiv (pb + qd) \pmod{m}$, for all integers p and q .
5. $a^n \equiv b^n \pmod{m}$ for all $n \in \mathbb{N}$
6. $f(a) \equiv f(b) \pmod{m}$, for every polynomial $f(x)$ with integer coefficients.

Proof:

The proofs of 1., 2., 4. are left to the reader.

3. Since $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$
 $\therefore a = b + mx, c = d + my$ for some $x, y \in \mathbb{Z}$.
 $ac = (b + mx)(d + my)$
 $= bd + m(xd + by + mxy)$
 $\therefore ac \equiv bd \pmod{m}$.

5. We prove the result by induction on n .

Since $a \equiv b \pmod{m} \therefore a^1 \equiv b^1 \pmod{m}$.

Hence the result is true for $n = 1$.

Let the result be true for $n = k$.

ie. $a^k \equiv b^k \pmod{m}$

also $a \equiv b \pmod{m}$

\therefore By (3), we get

$a^k a \equiv b^k b \pmod{m}$

i.e. $a^{k+1} \equiv b^{k+1} \pmod{m}$

Hence the result holds for $n = k + 1$.

Thus, by the principle of induction, the result holds for all $n \in \mathbb{N}$.

6. Let $f(x) = p_0 + p_1x + \dots + p_nx^n$, $p_i \in \mathbb{Z}$, $n \in \mathbb{N} \cup \{0\}$.
 Since $a \equiv b \pmod{m}$
 $\therefore a^k \equiv b^k \pmod{m}$ for all $k \in \mathbb{N}$ by (5.).
 $\Rightarrow p_k a^k \equiv p_k b^k \pmod{m}$, using Theorem 4.25
 $\Rightarrow p_1 a + \dots + p_n a^n \equiv p_1 b + \dots + p_n b^n \pmod{m}$, using (1.)
 $\Rightarrow p_0 + p_1 a + \dots + p_n a^n \equiv p_0 + p_1 b + \dots + p_n b^n \pmod{m}$, using Theorem 4.25
 $\Rightarrow f(a) \equiv f(b) \pmod{m}$ \square

The following two theorems give the equivalent of cancellation laws in congruence modulo m .

Theorem 4.27. *If $ac \equiv bc \pmod{m}$ and $(c, m) = 1$ then $a \equiv b \pmod{m}$.*

Proof: Since $(c, m) = 1$, therefore $c \neq 0$, for if $c = 0$, then $(c, m) = m$, but $m > 1$

$$\begin{aligned} ca &\equiv cb \pmod{m} \\ \Rightarrow m &\mid (ca - cb) \\ \Rightarrow m &\mid c(a - b) \\ \Rightarrow m &\mid (a - b) && \because (c, m) = 1 \\ \Rightarrow a &\equiv b \pmod{m} \end{aligned} \quad \square$$

The next theorem gives the cancellation law when $(c, m) = d$. In fact, the above theorem becomes a special case, for $d = 1$.

Theorem 4.28. *Let $ac \equiv bc \pmod{m}$, and $c \neq 0 \pmod{m}$. If $(c, m) = d$ and $m = m_1 d$, then $a \equiv b \pmod{m_1}$.*

Proof: $d = (c, m) \Rightarrow d \neq 0$ and $d \mid c, d \mid m$
 $\Rightarrow c = c_1 d, m = m_1 d$ for some $c_1, m_1 \in \mathbb{Z}$.
 Such that $(c_1, m_1) = 1$

$$\begin{aligned} ac &\equiv bc \pmod{m} \\ \Rightarrow ac - bc &= mt \quad \text{for some } t \in \mathbb{Z} \\ \Rightarrow ac_1 d - bc_1 d &= m_1 dt \\ \Rightarrow (ac_1 - bc_1)d &= m_1 dt \\ \Rightarrow ac_1 - bc_1 &= m_1 t, \text{ as } d \neq 0 \\ \Rightarrow ac_1 &\equiv bc_1 \pmod{m_1} \end{aligned}$$

Since $(c_1, m_1) = 1$, therefore by the above theorem
 $a \equiv b \pmod{m_1}$. \square

We now find solution of congruences.

Definition 4.8. *An integer x_0 is a solution of the linear congruence $ax \equiv b \pmod{m}$ if $ax_0 \equiv b \pmod{m}$.*

Example 4.12. *Consider the congruence*

$$3x \equiv 2 \pmod{5} \quad (4.37)$$

$x_0 = 4$ is a solution, because $3x_0 - 2 = 10 = 2 \times 5$

$$\therefore 3x_0 \equiv 2 \pmod{5}$$

Similarly $x_0 = 9, 14, -1, -6$ are all solution of (4.37)
 Observe that $9 = 4 + 5$
 $14 = 4 + 2 \times 5$
 $-1 = 4 + (-1)5$
 $-6 = 4 + (-2)5$

So $9, 14, -1, -6 \in [4]$.

Does it mean that every member of $[4]$ is a solution of equation (4.37). This is precisely the case, as is proved in the next theorem.

Theorem 4.29. *If $(a, m) = 1$, then the linear congruence $ax \equiv b \pmod{m}$ has a solution. Moreover, if x_0 is a solution then set of all solutions is $[x_0]$, the equivalence class of x_0 .*

Proof: *Existence of a solution*

$$(a, m) = 1$$

\Rightarrow there exist $r, s \in \mathbb{Z}$ such that $ar + ms = 1$.

$$\begin{aligned} \Rightarrow arb + msb &= b, && \text{(multiplying by } b\text{).} \\ \Rightarrow arb - b &= m(-sb) \\ \Rightarrow ax_0 - b &= m(-sb), && \text{where } x_0 = rb \in \mathbb{Z} \\ \Rightarrow ax_0 &\equiv b \pmod{m}. \end{aligned}$$

Thus the given congruence has a solution, namely x_0 .

We now find the set of all solutions.

Now x_0 is solution, then

$$ax_0 \equiv b \pmod{m} \tag{4.38}$$

Let y be any solution of $ax \equiv b \pmod{m}$. Then

$$ay \equiv b \pmod{m} \tag{4.39}$$

$$\begin{aligned} (4.38) \text{ and } (4.39) &\Rightarrow ax_0 \equiv ay \pmod{m} \\ &\Rightarrow x_0 \equiv y \pmod{m}, && \because (a, m) = 1 \\ &\Rightarrow y \in [x_0]. \end{aligned}$$

Let $z \in [x_0]$

$$\therefore z \equiv x_0 \pmod{m}$$

$$\Rightarrow az \equiv ax_0 \pmod{m}$$

$$\Rightarrow az \equiv b \pmod{m}$$

Using (4.38).

$\Rightarrow z$ is a solution of (4.38). Thus the solution set is $[x_0]$. \square

Remark 4.4. *Since any two solutions are congruent modulo m , therefore we say that the solution is unique modulo m . In case there is no chance of confusion, we simply say that a congruence has unique solution.*

Corollary 4.30. *The linear congruence $ax \equiv b \pmod{p}$ where p is a prime such that $p \nmid a$ has a unique solution modulo p .*

Proof: p is a prime and $p \nmid a$, $\therefore (a, p) = 1$.

The result now follows from Theorem 4.29. \square

Example 4.13. Solve the linear congruence $4x \equiv 3 \pmod{5}$.

Here $a = 4$, $b = 3$, $m = 5$

since $(a, m) = 1$, therefore the given congruence has a unique solution modulo 5.

Step1 Find r, s such that

$$ar + ms = 1 \quad (4.40)$$

In this case $s = 1$, $r = (-1)$. $\therefore 4 \times (-1) + 5 \times 1 = 1$

Multiplying (4.40) by b .

$$\therefore arb + msb = b$$

$$\Rightarrow a(rb) - b = m(-sb)$$

$$\Rightarrow ax_0 \equiv b \pmod{m}, \quad \text{where } x_0 \equiv rb$$

Thus x_0 is a solution and $[x_0]$ is the set of all solutions.

$$x_0 = rb = (-1)3 = (-3) \equiv 2 \pmod{5}$$

(We generally take the smallest positive value of the solution).

Thus $x_0 = 2$ is the solution which is unique $\pmod{5}$ of the given linear congruence.

Solution set is $[2]$.

We have seen that the linear congruence $ax \equiv b \pmod{m}$ always has a unique solution, if $(a, m) = 1$.

If $(a, m) = d \neq 1$, then a solution may or may not exist.

Example 4.14. Find the solution of $18x \equiv 5 \pmod{6}$, if it exists.

Suppose x_0 is a solution of the given congruence. Then

$$18x_0 \equiv 5 \pmod{6}.$$

$$\therefore 18x_0 - 5 = 6k \quad \text{for some } k \in \mathbb{Z}$$

$$\Rightarrow 18x_0 - 6k = 5$$

$$\Rightarrow 3|5 \quad \therefore 3|(18x_0 - 6k).$$

which is not true.

Hence the given congruence does not have a solution.

Observe that in the above illustration $(a, m) = (12, 15) = 3 \neq 1$.

The following theorem gives the condition under which a given congruence always has a solution.

Theorem 4.31. The linear congruence $ax \equiv b \pmod{m}$ has a solution if and only if $d|b$, where $d = (a, m)$. If a solution exists, it is a unique solution modulo m_1 , where $m = m_1d$. In fact, there are exactly d solutions x_i , $0 \leq x_i < m$, no two of which are congruent modulo m . \square

Proof: Let $d = (a, m)$ (1)

Suppose the given linear congruence has a solution x_0 .

$$\text{Then } ax_0 \equiv b \pmod{m}$$

$$\text{So that } m|(ax_0 - b).$$

$$\Rightarrow ax_0 - b = mk, \quad \text{for some } k \in \mathbb{Z}$$

$$\Rightarrow ax_0 - mk = b \dots (2)$$

$$\text{Now (1)} \Rightarrow d|a \text{ and } d|m$$

$$\Rightarrow d|(ax_0 - mk)$$

$$\Rightarrow d|b \text{ using (2)}$$

Thus $d|b$.

Conversely, let $d|b$. $\therefore b = dk$, for some $k \in \mathbb{Z}$. Also (1) gives that there exists $\lambda, \mu \in \mathbb{Z}$ such that

$$\begin{aligned} a\lambda + m\mu &= d \\ \Rightarrow a\lambda k + m\mu k &= dk \\ \Rightarrow ax_0 + m\mu k &= b \text{ where } x_0 = \lambda k \\ \Rightarrow & m | (ax_0 - b) \\ \Rightarrow ax_0 &\equiv b \pmod{m} \\ \Rightarrow ax &\equiv b \pmod{m} \text{ has a solution} \end{aligned}$$

namely $x = x_0$.

Hence $ax \equiv b \pmod{m}$ has a solution if and only if $d|b$.

Suppose $ax \equiv b \pmod{m}$ has two solutions x_0 and x_1 . Then

$$\begin{aligned} ax_0 &\equiv b \pmod{m} \\ \text{and } ax_1 &\equiv b \pmod{m} \end{aligned}$$

So that $a(x_0 - x_1) \equiv 0 \pmod{m}$

$$\therefore a(x_0 - x_1) = km, \text{ for some } k \in \mathbb{Z} \dots (3)$$

Since $d = (a, m)$, \therefore By Theorem there exists $a_1, m_1 \in \mathbb{Z}$ such that $a = a_1d$, $m = m_1d$, $(a_1, m_1) = 1$. \therefore (3) becomes

$$\begin{aligned} a_1d(x_0 - x_1) &= km_1d \\ \Rightarrow a_1(x_0 - x_1) &= km_1 \quad \because d \neq 0 \\ \Rightarrow m_1 | a_1(x_0 - x_1) \\ \Rightarrow m_1 | (x_0 - x_1) \text{ as } (a_1, m_1) = 1. \\ \Rightarrow x_0 &\equiv x_1 \pmod{m_1} \end{aligned}$$

Hence there exists a unique solution. Call it x_0 .

The solution set is $[x_0] = \{x_0 + m_1t : t \in \mathbb{Z}\}$.

When any integer t is divided by d , the remainders can be any one of $0, 1, \dots, d - 1$.

Any element of the solution set is of the form.

$$x_0 + m_1(kd + r), \text{ where } r = 0, 1, \dots, d - 1$$

$$\text{i.e. } x_0 + m_1dk + m_1r, \text{ where } r = 0, 1, \dots, d - 1$$

$$\text{i.e. } (x_0 + m_1r) + m_1dk, \text{ where } r = 0, 1, \dots, d - 1$$

$$\text{i.e. } x_0 + m_1r, \text{ where } r = 0, 1, \dots, d - 1; \pmod{m}$$

These are d solutions, no two of which are congruent \pmod{m} .

Hence the given congruence has d distinct solutions, no two of which are congruent modulo m .

Example 4.15. Solve the linear congruence $24x \equiv 9 \pmod{81}$.

Step 1

$$24x \equiv 9 \pmod{81} \tag{4.41}$$

$$\Leftrightarrow 24x - 9 = 81k \quad \text{for some } k \in \mathbb{Z}$$

$$\Leftrightarrow 8x - 3 = 27k$$

$$\Leftrightarrow 8x \equiv 3 \pmod{27} \tag{4.42}$$

Thus x_0 is a solution of (4.41) if and only if x_0 is a solution of (4.42).

We now solve (4.42).

Step 2

$$\text{Here } a = 8, \quad b = 3, \quad m = 27$$

$$(a, m) = 1.$$

Find r and s such that

$$ar + ms = 1 \tag{4.43}$$

In fact $8 \times (-10) + 27 \times 3 = 1$, so that $r = -10$, $s = 3$.

Multiplying (4.43) by b , we get

$$\begin{aligned} arb + msb &= b \\ \Rightarrow arb - b &= m(-sb) \\ \Rightarrow ax_0 - b &= m(-sb), \text{ where } x_0 = rb. \\ \Rightarrow ax_0 &\equiv b \pmod{m} \end{aligned}$$

Thus $x_0 = -30$ is a solution of (4.42), But $x_0 \equiv 24 \pmod{27}$.

$\therefore 24$ is the unique solution mod m of (4.42).

The solutions are $24 + 27t$, $t \in \mathbb{Z}$

Step 3

We now obtain all the solutions of (4.41).

Since $(24, 81) = 3$

\therefore There are 3 non-congruent solutions modulo 81.

To obtain these solutions, we proceed as follows:

Any integer is of the form

$$3k, 3k + 1, 3k + 2$$

Thus, any solution of (4.41) is of the form

$$24 + 27(3k), 24 + 27(3k + 1), 24 + 27(3k + 2).$$

i.e. $24 + 81k, 51 + 81k, 78 + 81k$.

Thus the solutions are

$$\begin{aligned} x &\equiv 24 \pmod{81} \\ x &\equiv 51 \pmod{81} \\ x &\equiv 78 \pmod{81}. \end{aligned}$$

To solve a linear congruence $ax \equiv b \pmod{m}$, first we check whether a solution exists or not.

Existence of Solution

There are three cases arises:

Case 1.

$$(a, m) = 1.$$

In this case there is a unique solution modulo m .

Case 2.

$$(a, m) = d > 1.$$

If $d \nmid b$ then there is no solution.

Case 3.

$$(a, m) = d > 1 \text{ and } d|b.$$

In this case there are d non-congruent solutions modulo m .

We now give the steps to find the solution in case 1 and case 3.

Steps involved for Case 1.

Step 1 Since $(a, m) = 1$

Find integers r and s such that $ar + ms = 1$.

Multiplying by b , we get

$$\begin{aligned} arb + msb &= b \\ \Rightarrow ax_0 - b &= m(-sb), \text{ where } x_0 = rb. \\ \Rightarrow ax_0 &\equiv b \pmod{m}. \end{aligned}$$

Thus x_0 is a solution.

Step 2 The solution set is

$$\{x_0 + km; k \in \mathbb{Z}\} \quad (4.44)$$

If x_0 does not satisfy $0 \leq x_0 < m$, then reduce it so that it satisfy the above condition.

This can always be done by adding multiples of m to x_0 , and it will still remain a solution, because of (4.44).

Steps involved for Case 2.

Step1 Since $(a, m) = d$

$$\therefore a = da_1, \quad m = dm_1, \quad b = db_1 \text{ and } (a_1, m_1) = 1.$$

$$\text{Then,} \quad ax \equiv b \pmod{m} \quad (4.45)$$

$$\Leftrightarrow a_1x \equiv b_1 \pmod{m_1} \quad (4.46)$$

Solve (4.46) as in Case 1.

Step 2 Obtain a solution of $a_1x \equiv b_1 \pmod{m_1}$.

where $(a_1, m_1) = 1$.

The steps have been outlined earlier. If x_0 is a solution, then the solution set is

$$\{x_0 + tm_1 : t \in \mathbb{Z}\}.$$

Step 3 To obtained all the non-congruent modulo m solutions of (4.45).

Since $(a, m) = d$, therefore there are d non-congruent solutions of (4.45). which are

$$\begin{aligned} x &\equiv x_0 \pmod{m} \\ x &\equiv x_0 + m_1 \pmod{m} \\ x &\equiv x_0 + 2m_1 \pmod{m} \\ &\vdots \\ &\vdots \\ &\vdots \\ x &\equiv x_0 + (d-1)m_1 \pmod{m}. \end{aligned}$$

Problem 4.12. Let m, n are fixed integers greater than 1, and $a, b \in \mathbb{Z}$. If $a \equiv b \pmod{mn}$, then $a \equiv b \pmod{m}$ and $a \equiv b \pmod{n}$.

Is the converse true? if not, is it true under certain condition on m and n ? what are they?

Solution: $a \equiv b \pmod{mn}$

$$\Rightarrow a - b = qmn \text{ for some } q \in \mathbb{Z}$$

$$\Rightarrow m \mid (a - b) \text{ and } n \mid (a - b)$$

$$\Rightarrow a \equiv b \pmod{m} \text{ and } a \equiv b \pmod{n}.$$

Conversely, If $a \equiv b \pmod{m}$ and $a \equiv b \pmod{n}$, we need to prove that $a \equiv b \pmod{mn}$

Take $a = 18, b = 14, m = 2, n = 4$

then,

$$18 \equiv 14 \pmod{4}$$

$$18 \equiv 14 \pmod{2}$$

but, $18 \not\equiv 14 \pmod{2 \times 4}$

$$18 \not\equiv 14 \pmod{8}$$

We shall find condition on m and n so that we get the result,

$$\begin{aligned} a &\equiv b \pmod{m} \Rightarrow m \mid (a - b) \\ \Rightarrow a - b &= mq_1 \text{ for some } q_1 \in \mathbb{Z} \\ \text{Similarly } a &\equiv b \pmod{n} \\ \Rightarrow a - b &= nq_2 \text{ for some } q_2 \in \mathbb{Z} \\ \text{Thus } a - b &= mq_1 = nq_2 \quad (1) \\ \text{If } \gcd(m, n) &= 1, \text{ then} \\ mq_1 &= nq_2 \\ \Rightarrow m \mid nq_2 &\text{ and } n \mid mq_1 \\ \Rightarrow n \mid q_1, &\text{ since } \gcd(m, n) = 1 \\ (1) \Rightarrow q_1 &= nk_1, \text{ for some } k_1 \in \mathbb{Z} \\ \Rightarrow a - b &= mnk_1 \\ \Rightarrow mn \mid (a - b) \\ \Rightarrow a &\equiv b \pmod{mn} \end{aligned}$$

Thus the converse holds when m and n are co-prime.

Problem 4.13. If 'a' is any integer then $a^3 \equiv 0, 1, \text{ or } 8 \pmod{9}$.

Solution: Any integer 'a' is of the form $3k, 3k + 1, 3k + 2$.

Three cases arises:

$$\begin{aligned} \text{Case 1.} \quad a &= 3k \\ \text{Then} \quad a^3 &= 27k^3 \\ \therefore a^3 &\equiv 0 \pmod{9}. \end{aligned}$$

$$\begin{aligned} \text{Case 2.} \quad a &= 3k + 1 \\ \therefore a^3 &= (3k + 1)^3 \\ &= 9(3k^3 + 3k^2 + k) + 1 \end{aligned}$$

$$\text{Hence } a^3 \equiv 1 \pmod{9}.$$

$$\begin{aligned} \text{Case 3.} \quad a &= 3k + 2 \\ \therefore a^3 &= (3k + 2)^3 \\ &= 9(3k^3 + 6k^2 + 8k) + 8 \end{aligned}$$

$$\text{Hence } a^3 \equiv 8 \pmod{9}. \text{ Thus if 'a' is any integer, then } a^3 \equiv 0, 1, \text{ or } 8 \pmod{9}.$$

Problem 4.14. Find the remainder when $1! + 2! + \dots + 200!$ is divided by 12.

Solution: Clearly $12 \mid n!$ for all $n \geq 4$, therefore $12 \mid (4! + \dots + 200!)$

$$\Rightarrow 4! + \dots + 200! = 12k \text{ for some } k \in \mathbb{Z} \Rightarrow 1! + 2! + 3! + 4! + \dots + 200! =$$

$$12k + 1! + 2! + 3! = 12k + 9$$

$$\Rightarrow 1! + 2! + \dots + 200! \equiv 9 \pmod{12}$$

Thus, the remainder obtained on dividing $1! + 2! + \dots + 200!$ by 12 is 9.

4.8 Exercise

1. Calculate
 - (i) $7 + 8 \pmod{11}$
 - (ii) $2 + 3 + 4 + 5 \pmod{6}$
 - (iii) $8.9 \pmod{10}$
 - (iv) $4.5.6 \pmod{7}$.

2. Find the least positive integer modulo n to which the following expressions are congruent.
 - (i) $5.7.13.23.413, n = 11$
 - (ii) $6 + 18 + 29 + 346, n = 13$
 - (iii) $5.6 + 8.11 + 19.23, n = 9$
 - (iv) $123.13.2 + 481.6 - 239.11 + 17.11 - 14.239, n = 15.$

3. Evaluate
 - (i) $(2517 \times 4328) \pmod{14}$
 - (ii) $(2610 + 3929) \pmod{9}$
 - (iii) $(1718)^5 \pmod{13}$
 - (iv) $(5621 - 7398) \pmod{12}$
 - (v) $2^{20} \pmod{11}$
 - (vi) $10^{24} \pmod{9}$
 - (vii) $8126^{100} \pmod{7}$

4. Calculate $a + b, a.b, (a + b)^2, (a + b)^3 \pmod{n}$, for
 - (i) $a = 11528, b = 17332, n = 91.$
 - (ii) $a = -11528, b = -17332, n = 91.$

5. Examine which of the following are true.
 - (i) $-6 \equiv 18 \pmod{12}$
 - (ii) $111 \equiv 12 \pmod{11}$
 - (iii) $111 \equiv 11 \pmod{11}$
 - (iv) $-4 \equiv -4 \pmod{n}, n \in \mathbb{N}$
 - (v) $100 \equiv 10 \pmod{20}$
 - (vi) $1625 \equiv 15 \pmod{25}.$

6. Write the congruence classes of integers modulo 12. To which class does 5876 belong? Does -5876 also belong to the same class?

7. What is the general form of an integer in $[3]$ relative to the congruence $\pmod{11}$.

8. Write 3 negative integers in $[2]$ relative to the relation congruent $\pmod{9}$.

9. Write the multiplication table of equivalence classes modulo 5.

10. In *ABC* university each student is assigned an enrolment number. The last three digits of the enrolment number of a male student born in month m on date b is $71m + 2b + 1$ and that of a female student is $71m + 2b$. Find the date of birth and sex corresponding to the numbers.
 - (i) 480
 - (ii) 911
 - (iii) 716
 - (iv) 717
 - (v) 172

11. The last seven digits of the identification number of an employee gives the date of birth. The four digits preceding the last three digits is $4Y$, where Y is the year of birth. The last three digits of the identification number of a female employee born in month m on date b is $67m + 2b + 1$ and that of a male employee is $71m + 2b$. Find the date of birth and sex corresponding to the numbers.

(i) 7792572

(ii) 7936703

12. Solve the following congruences if the solution exists. If no solution exists, explain why.

(i) $3x \equiv 1 \pmod{7}$

(ii) $8x \equiv 4 \pmod{6}$

(iii) $8x \equiv 5 \pmod{6}$

(iv) $27x \equiv 8 \pmod{9}$

(v) $27x \equiv 15 \pmod{9}$

(vi) $4x \equiv 1 \pmod{6}$

(vii) $4x \equiv 2 \pmod{6}$

(viii) $8x \equiv 4 \pmod{12}$

(ix) $8x \equiv 3 \pmod{27}$

(x) $12x \equiv 9 \pmod{15}$.

13. Find the solution of

(i) $2x \equiv 3 \pmod{9}$

(ii) $4x \equiv 6 \pmod{9}$.

Do you see some relation between the solutions of (i) and (ii).

14. Find the solutions of

(i) $5x \equiv 1 \pmod{12}$

(ii) $10x \equiv 2 \pmod{12}$

(iii) $15x \equiv 3 \pmod{12}$.

Do you see any relation between their solution sets.

15. Construct linear congruences modulo 12 with

(i) One solution $\pmod{12}$

(ii) No solution

(iii) More than one solution $\pmod{12}$.

16. Find the remainder when $(2138)^9$ is divided by 31.

17. Find the remainder when

(i) $(2)^{50}$ is divided by 7.

(ii) $(41)^{65}$ is divided by 6.

4.9 Supplementary Problems

1. State whether the following statements are true or false. Justify.
 - (i) Cancellation law for multiplication holds in the set of natural numbers.
 - (ii) The equation $x + m = n$ has a solution in the set of natural numbers, for all $m, n \in \mathbb{N}$.
 - (iii) If $a, b, c \in \mathbb{Z}$, then

$$ac = bc \Rightarrow a = b.$$
 - (iv) The first principle of induction is a special case of the second principle of induction.
 - (v) The principle of induction is used to prove results about numbers only.
 - (vi) The first principle of induction is equivalent to the Well ordering principle.
 - (vii) The relation of divisibility is an equivalence relation on \mathbb{N} .
 - (viii) A number $p \in \mathbb{N}$ is prime if the only divisors of p are ± 1 and $\pm p$.
 - (ix) A number which is not prime is a composite number.
 - (x) Division algorithm holds in the set of natural numbers.
 - (xi) If $a \in \mathbb{Z}$, then $\gcd(a, 0) = a$.
 - (xii) The least element of the set $\{ax + by \mid x, y \in \mathbb{Z}, ax + by > 0\}$ is the \gcd of a and b .
 - (xiii) If $a, b, c \in \mathbb{Z}$, then $\gcd(ac, bc) = c \gcd(a, b)$.
 - (xiv) If $a, b, c \in \mathbb{Z}$, such that $ac \equiv bc \pmod{m}$ for some integer $m > 1$, then $a \equiv b \pmod{m}$.
 - (xv) If $a, b, c \in \mathbb{Z}$, such that $a|c$ and $b|c$ then $ab|c$.
 - (xvi) $25x \equiv 1 \pmod{49}$ has no solution.
 - (xvii) $2x \equiv 3 \pmod{6}$ has no solution.
 - (xviii) If $a \in [b] \pmod{m}$ is equivalent to saying $[a] = [b]$.
 - (xix) If $a, b, c \in \mathbb{Z}$, such that $ab \equiv 1 \pmod{m}$ then either $a \equiv 1 \pmod{m}$ and $b \equiv 1 \pmod{m}$, or $a \equiv -1 \pmod{m}$ and $b \equiv -1 \pmod{m}$.
 - (xx) If $a \in \mathbb{Z}$, such that $a^3 \equiv 1 \pmod{m}$ then $a \equiv 1 \pmod{m}$.
 - (xxi) $\gcd(a, -a) = a, a \in \mathbb{Z}$.
 - (xxii) If two integers a, b are coprime then their lcm is ab .
 - (xxiii) Two consecutive integers are always coprime.
2. Prove that an integer is divisible by 4 if and only if the number formed by the last two digits (digits in the ten's and unit's place) is divisible by 4.
3. Prove that an integer n is divisible by 9 if and only if the sum of the digits of n is divisible by 9.
4. Show that
 - (i) $(2)^n > n^2$ for all $n \geq 5$.
 - (ii) $n! \geq n^3$ for all $n \geq 6$.
5. If n is any integer, then prove that $5n + 3$ and $7n + 4$ are coprime.
6. If p is a prime and a, b are integers such that $ab \equiv 0 \pmod{p}$, prove that either $a \equiv 0 \pmod{p}$ or $b \equiv 0 \pmod{p}$. What can you say if p is not prime?
7. If $a \equiv b \pmod{n}$, prove that $\gcd(a, n) = \gcd(b, n)$.

8. If $a, b \in \mathbb{Z}$, and k is an odd integer such that $k \mid (a - b)$ and $k \mid (a + b)$ then prove that $k \mid \gcd(a, b)$.
9. If a, b are co-prime, prove that
- $\gcd(2a + b, a + 2b) = 1$ or 3 .
 - $\gcd(a + b, a^2 - ab + b^2) = 1$ or 3 .
 - $\gcd(a^m, b^n) = 1$, $m, n \in \mathbb{N}$.
10. If ' a ' is an odd integer prove that $a^2 \equiv 1 \pmod{8}$
11. Find the \gcd of a and b . Also find integers m and n such that $ma + nb = \gcd(a, b)$, for the following pairs.
- $a = 578, b = -442$
 - $a = -826, b = 1890$
 - $a = 741, b = 1079$.
12. Let a be an odd integer. Show that $a^{2^n} \equiv 1 \pmod{2^{n+2}}$ for all $n \in \mathbb{N}$ using induction.
13. Solve the linear congruence

$$40x \equiv 15 \pmod{135}$$

14. Prove that

$$84^{10} \equiv 1 \pmod{11}$$

15. Solve the following linear congruences. Obtain all the non congruent solutions.
- $258x + 18 \equiv 5 \pmod{7}$
 - $222x + 7 \equiv 19 \pmod{18}$
 - $12x \equiv 6 \pmod{16}$
 - $18x + 24 \equiv 15 \pmod{33}$.
16. Suppose $P(n)$ is a statement about the natural number n such that
- $P(1)$ is true.
 - For any $k \geq 1$, $P(k)$ is true $\Rightarrow P(2k)$ is true.
 - For any $k \geq 2$, $P(k)$ is true $\Rightarrow P(k - 1)$ is true.
- Prove that $P(n)$ is true for all $n \in \mathbb{N}$.

4.10 Answers to Exercises

Exercise - 4.7

- 11.
- $13, 13=7a - 4b$
 - $13, 13=-7a - 4b$
 - $1, 1=-40a + 79b$
 - $1, 1=40a - 79b$
 - $592, 592=2a - b$
 - $592, 592=2a + b$
 - $2, 2=97a + 68b$

12. (i) $27, m = -3, n = 4; m = -3 - kb, n = 4 + ka$, for $k = 1, 2$
(ii) $13, m = -16, n = 11; m = -16 - kb, n = 11 + ka$, for $k = 1, 2$
(iii) $73, m = -5, n = 17; m = -5 - kb, n = 17 + ka$, for $k = 1, 2$.
13. $m = 553, n = -280$.
No, other pair is $m = 867, n = -439$.
14. $m = -2, n = 2$.
15. *Hint:* $\gcd(219, 153) = 3$ and $3 \nmid 5$.
16. *Hint:* $7 \mid (10m + n) \Leftrightarrow 10m + n = 7q \Leftrightarrow m + 5n = 7(5q - 7m) \Leftrightarrow 7 \mid (m + 5n)$
18. (i) Yes
(ii) Yes
(iii) Yes
(iv) No

Exercise - 4.8

1. (i) 4 (ii) 2 (iii) 2 (iv) 1.
2. (i) 2 (ii) 9 (iii) 6 (iv) 11.
3. (i) 8 (ii) 5 (iii) 6 (iv) 11.
4. (i) 13, 56, 78, 13
(ii) 78, 56, 78, 78
5. (i) T (ii) T (iii) F (iv) T (v) F (vi) F.
6. $[k] = \{k + 12t \mid t \in \mathbb{Z}\}$ for $k = 0, 1, 2, \dots, 11$.
 $5876 \equiv 8 \pmod{12}$
 $\therefore 5876 \in [8]$. No.
 $-5876 \in [4]$.
7. $11k + 3, k \in \mathbb{Z}$.
8. $-7, -16, -25$.
- 9.
- | | | | | | |
|-----------|-----------|-----------|-----------|-----------|-----------|
| $\bar{0}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
| $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ |
| $\bar{1}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
| $\bar{2}$ | $\bar{0}$ | $\bar{2}$ | $\bar{4}$ | $\bar{1}$ | $\bar{3}$ |
| $\bar{3}$ | $\bar{0}$ | $\bar{3}$ | $\bar{1}$ | $\bar{4}$ | $\bar{2}$ |
| $\bar{4}$ | $\bar{0}$ | $\bar{4}$ | $\bar{3}$ | $\bar{2}$ | $\bar{1}$ |
10. (i) F, 27th June
(ii) M, 29th Dec
(iii) F, 3rd Oct
(iv) M, 3rd Oct
(v) Incorrect

11. (i) M, 2 August, 1948
(ii) F, 16 October, 1984
12. (i) $5(\text{mod } 7)$
(ii) $2(\text{mod } 6), 5(\text{mod } 6)$
(iii) No solution
(iv) No solution
(v) No solution
(vi) No solution
(vii) $2(\text{mod } 6), 5(\text{mod } 6)$
(viii) $2(\text{mod } 12), 5(\text{mod } 12), 8(\text{mod } 12), 11(\text{mod } 12)$
(ix) $24(\text{mod } 7)$
(viii) $2(\text{mod } 15), 7(\text{mod } 15), 12(\text{mod } 15)$.
13. (i) $6(\text{mod } 9)$
(ii) $6(\text{mod } 9)$
- Both have the same solution.
14. (i) $5(\text{mod } 12)$
(ii) $5(\text{mod } 12), 11(\text{mod } 12)$
(iii) $5(\text{mod } 12), 1(\text{mod } 12), 9(\text{mod } 12)$.
- (i), (ii) and (iii) have a common solution.
15. Other answers are possible
(i) $5x \equiv 7(\text{mod } 12)$
(ii) $2x \equiv 3(\text{mod } 12)$
(iii) $10x \equiv 2(\text{mod } 12)$.
16. 30
17. (i) 4 (iii) 5.

Supplementary Exercises

1. (i) T
(ii) F, $x + 5 = 1$ does not have a solution.
(iii) F, not true for $c = 0$.
(iv) T
(v) F, any statement involving natural numbers.
(vi) T
(vii) F, it is not symmetric.
(viii) F, $p, 1$
(ix) F, it can be unity also.

- (x) F, No quotient exists when 3 is divided by 5.
- (xi) F, $\gcd(a, 0) = |a|$.
- (xii) T
- (xiii) F, $\gcd(ac, bc) = |c|\gcd(a, b)$.
- (xiv) F, $7 \times 3 \equiv 5 \times 3 \pmod{6}$, but $7 \not\equiv 5 \pmod{6}$.
- (xv) F, $8|24$, $12|24$ but $8 \times 12 \nmid 24$.
- (xvi) F, $x \equiv 2 \pmod{49}$ is a solution.
- (xvii) T
- (xviii) T
- (xix) F, $2 \times 3 \equiv 1 \pmod{5}$, but $2 \not\equiv \pm 1 \pmod{5}$ and $3 \not\equiv \pm 1 \pmod{5}$.
- (xx) F, $5^3 \equiv 1 \pmod{31}$, but $5 \not\equiv 1 \pmod{31}$.
- (xxi) F, $\gcd(a, -a) = |a|$.
- (xxii) F, lcm is $|ab|$.
- (xxi) T

6. Not true. $2 \times 3 \equiv 0 \pmod{6}$, but $2 \not\equiv 0 \pmod{6}$ and $3 \not\equiv 0 \pmod{6}$.

9. *Hint:*

- (a) $d = \gcd(a + b, a^2 - ab + b^2)$
 $\Rightarrow d|(a + b), d|(a^2 - ab + b^2)$
 $\Rightarrow d|((a + b)^2 - (a^2 - ab + b^2))$
 $\Rightarrow d|(3ab)$
 Let $\gcd(d, a) = d'$
 $\therefore d'|d$ and $d| \Rightarrow d'|b$
 $d' = 1 \therefore \gcd(a, b) = 1$.
 Similarly $\gcd(d, b) = 1$.
 Hence $d|3 \Rightarrow d=1, 3$.
- (b) Use induction to prove $(a^m, b) = 1$.

11.

- (i) $34, m = -3, n = -4$
- (ii) $14, m = 16, n = 7$
- (iii) $13, m = -16, n = 11$

13. 24, 51, 78, 105, 132 ($\pmod{135}$)

14.

- (i) $6 \pmod{7}$
- (ii) 2, 5, 8, 11, 14, 17 ($\pmod{18}$)
- (iii) No solution exists.
- (ii) 5, 16, 27 ($\pmod{33}$)

This page is intentionally left blank.

UNIT - 2

Chapter 5

Group

Definition and Examples

In this chapter we shall study different algebraic structures with one binary operation and the relationship amongst them. The simplest algebraic structure is a groupoid. We begin with a few definitions.

5.1 Definition of Group

Definition 5.1. A non-empty set G equipped with a binary operation $*$ is called a groupoid, that is, $a * b \in G \forall a, b \in G$. This is also referred to as: G is closed with respect to $*$.

Definition 5.2. A non-empty set G equipped with a binary operation $*$ is called a semigroup if $*$ is associative, i.e.

$$a * (b * c) = (a * b) * c \quad \forall a, b, c \in G. \quad (5.1)$$

Definition 5.3. A non-empty set G equipped with a binary operation $*$ is called a monoid if

(i) $*$ is associative, i.e.

$$a * (b * c) = (a * b) * c \quad \forall a, b, c \in G. \quad (5.2)$$

(ii) $*$ has an identity element, i.e., there exists an element $e \in G$ such that

$$a * e = e * a = a \quad \forall a \in G.$$

Definition 5.4. A non-empty set G equipped with a binary operation $*$ is called a group if

(i) $*$ is associative, i.e.

$$a * (b * c) = (a * b) * c \quad \forall a, b, c \in G. \quad (5.3)$$

(ii) $*$ has an identity element, that is, there exists an element $e \in G$ such that

$$a * e = e * a = a \quad \forall a \in G.$$

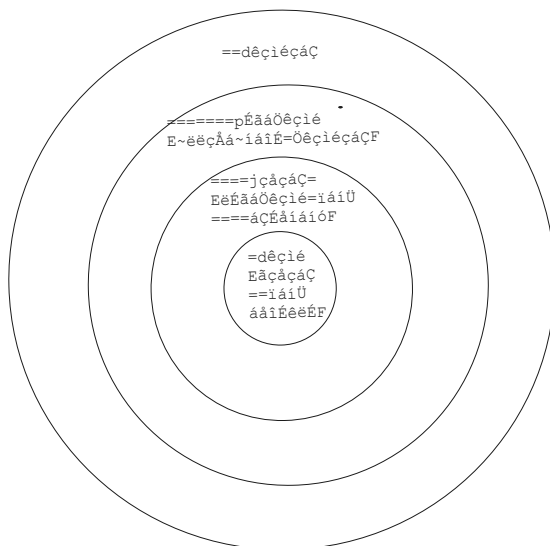
(iii) each element of G has an inverse with respect to $*$, that is, for every $a \in G$ there exists some $b \in G$ such that

$$a * b = b * a = e.$$

' b ' is called an inverse of ' a ' and is denoted by a^{-1} .

Each of these algebraic structures are denoted by $(G, *)$. When the binary operation $*$ is understood from the context, we simply say that G is an algebraic structure, and if there is no confusion we prefer to write $a * b$ as ab .

From the above definitions it is clear that an associative groupoid is a semi-group; a semi-group with an identity element is a monoid; and a monoid in which every element has inverse is a group. This relationship is shown in the following diagram.



Definition 5.5. In a group $(G, *)$, if the set G is finite, the number of elements in G is called the order of G and is denoted by $o(G)$ (or $|G|$). If G is not a finite set, then the group G is said to be infinite or a group of infinite order.

A groupoid is the simplest algebraic structure. Certain important results about groups also hold true in more general structures like semigroups and monoids, and therefore these structures have been discussed here. These structures have not been studied in details as our principal interest is to study groups.

Definition 5.6. A group $(G, *)$, or a semigroup $(G, *)$, is said to be commutative or Abelian group, if

$$a * b = b * a \quad \forall a, b \in G. \tag{5.4}$$

Having given a formal definition of a group, we shall now build up a good stock of examples. These examples will be used throughout to illustrate results

for their better comprehension. The reader is advised to study them carefully since the best way to feel the essence of a theorem is to see what happens in specific cases. To develop a complete understanding of these examples, you may supply the missing details.

Whenever we check whether a given set is a group with respect to a given binary operation $*$, we shall proceed as follows:

Step 1 (Closure) Verify that $*$ is a binary operation on G , that is, $a * b \in G$, for all $a, b \in G$.

Step 2 (Associativity) Verify that $*$ is associative.

Step 3 (Existence of identity) Verify the existence of identity element $e \in G$ with respect to $*$.

Step 4 (Existence of inverse) Verify that every element of G has an inverse in G , with respect to $*$.

In case it has to be seen whether the group is Abelian, we must do the following additional step.

Step 5 (Commutativity) Verify that $a * b = b * a$, for all $a, b \in G$.

We observe that the associative law holds with respect to the usual addition and multiplication in the set of complex numbers. Therefore associative law holds for every subset thereof. Hence, for any set of numbers with respect to the usual addition and multiplication, it is required to check only the existence of identity and inverse. Moreover the commutative law holds with respect to the usual addition and multiplication in the set of complex numbers, so that this law holds for every subset thereof. Hence for any sets of numbers, if $(S, +)$ or (S, \cdot) is a group, it will be an Abelian group.

Remark 5.1. *If only Step 1 holds then $(G, *)$ is a groupoid. If for a groupoid Step 2 also holds then $(G, *)$ is a semigroup. If for a semigroup, Step 3 holds then $(G, *)$ is a monoid. Finally a monoid for which Step 4 holds is a group. For a groupoid (semigroup, monoid, group respectively) if Step 5 holds then it is an Abelian groupoid (semigroup, monoid, group respectively).*

Remark 5.2. *In case of finite groups of small order, sometimes it is convenient to prepare a multiplication table to verify the above steps. This table is known as the Cayley table. It is prepared as follows:*

Let $G = \{x_1, x_2, \dots, x_n\}$ be a set and let $*$ be an operation defined on G , then the Cayley table of G with respect to $*$ is prepared as:

*	x_1	x_2	.	.	.	x_n
x_1	$x_1 * x_1$	$x_1 * x_2$.	.	.	$x_1 * x_n$
x_2	$x_2 * x_1$	$x_2 * x_2$.	.	.	$x_2 * x_n$
.
.
.
x_n	$x_n * x_1$	$x_n * x_2$.	.	.	$x_n * x_n$

Remark 5.3. *If $(G, *)$ is a group, then in the multiplication table in each row and each column, every element of G appears exactly once. Note that this is only a necessary condition, and not a sufficient condition for G to be a group.*

5.2 Exercise

Give examples to justify the statements, in Q1 to Q5.

1. The set of all odd integers is not a groupoid with respect to addition.
2. $(\mathbb{Z}, -)$ is a groupoid but not a semigroup.
3. (E, \cdot) is a semigroup but not a monoid, where E is the set of even integers.
4. $(\mathbb{N}, +)$ is a semigroup but not a monoid.
5. (\mathbb{N}, \cdot) is a monoid but not a group.
6. What algebraic structure does (Q^*, \div) possess?
7. On \mathbb{N} , define $m * n = m^n$. What algebraic structure does $(\mathbb{N}, *)$ possess?
8. Let $G = \{a, b, c\}$. The binary operation $*$ on G is defined by the following table:

$*$	a	b	c
a	a	b	c
b	b	b	b
c	c	c	c

Prove that $(G, *)$ a monoid. Is it a group?

9. Let $G = \{a, b, c, d\}$. The binary operation $*$ on G is defined by the following table:

$*$	a	b	c	d
a	a	b	c	d
b	b	c	d	a
c	c	d	a	a
d	d	b	a	c

Is $(G, *)$ a group? If not, why?

5.3 Groups of Numbers

Example 5.1. *The set of integers \mathbb{Z} , the set of rational numbers \mathbb{Q} , the set of real numbers \mathbb{R} and the set of complex numbers \mathbb{C} are all groups under ordinary addition. The identity element in each case is 0. The inverse of any element x is $-x$. Since addition of numbers is commutative, therefore they all are Abelian groups. Since \mathbb{Z} , \mathbb{Q} , \mathbb{R} and \mathbb{C} are infinite sets, these are infinite Abelian groups.*

Each of the sets \mathbb{Z} , \mathbb{Q} , \mathbb{R} and \mathbb{C} is closed with respect to the usual multiplication of numbers. The associative law holds, the identity element exists and is 1. Thus each of (\mathbb{Z}, \cdot) , (\mathbb{Q}, \cdot) , (\mathbb{R}, \cdot) and (\mathbb{C}, \cdot) is a monoid.

In (\mathbb{Z}, \cdot) none of the elements, except 1 and -1 are invertible. So (\mathbb{Z}, \cdot) is not a group. In (\mathbb{Q}, \cdot) every non-zero element $\frac{m}{n}$ has an inverse, namely $\frac{n}{m}$. But 0 does not have an inverse as there does not exist any $q \in \mathbb{Q}$ such that $0 \cdot q = 1$. Thus (\mathbb{Q}, \cdot) is not a group. For the similar reason, (\mathbb{R}, \cdot) and (\mathbb{C}, \cdot) are also not groups. Since 0 is the only element which does not have an inverse, is it possible that the set of non-zero numbers forms a group with respect to multiplication? This is answered in the following example.

Example 5.2. Let \mathbb{Q}^* denote the set of non-zero rational numbers. Then (\mathbb{Q}^*, \cdot) is a group. Since the product of two non-zero rational numbers is a non-zero rational number, \mathbb{Q}^* is closed with respect to multiplication. The identity element is 1 and the inverse of $\frac{m}{n} \in \mathbb{Q}^*$ is $\frac{n}{m}$. Since multiplication is commutative, therefore (\mathbb{Q}^*, \cdot) is an Abelian group. Moreover \mathbb{Q}^* is infinite. Thus (\mathbb{Q}^*, \cdot) is an infinite Abelian group. Similarly, (\mathbb{R}^*, \cdot) and (\mathbb{C}^*, \cdot) are also infinite Abelian groups.

Example 5.3. For any fixed integer m , let $m\mathbb{Z} = \{mz \mid z \in \mathbb{Z}\}$. Clearly $(m\mathbb{Z}, +)$ is closed. In fact, $(m\mathbb{Z}, +)$ is a group. For, the identity element is 0 and the inverse of $mz \in m\mathbb{Z}$ is $-mz$. Since addition is commutative, therefore $(m\mathbb{Z}, +)$ is an Abelian group. Further, $m\mathbb{Z}$ being an infinite set, it is an infinite Abelian group.

Note that the set \mathbb{E} of even integers is a group with respect to addition (taking $m = 2$ in the above example, $\mathbb{E} = 2\mathbb{Z}$). What can we say about the set \mathbb{O} of odd integers? Is $(\mathbb{O}, +)$ a group? Since the sum of two odd integers is even, therefore addition is not a binary operation on \mathbb{O} . Thus $(\mathbb{O}, +)$ is not even a groupoid. What about $\mathbb{O}^* = \mathbb{O}$.

Example 5.4. Let \mathbb{Q}^+ denote the set of all positive rational numbers. Then (\mathbb{Q}^+, \cdot) is a group. Clearly \mathbb{Q}^+ is closed. The identity is 1 and the inverse of $\frac{m}{n} \in \mathbb{Q}^+$ is $\frac{n}{m}$. It is an infinite Abelian group. Similarly (\mathbb{R}^+, \cdot) is an infinite Abelian group.

Example 5.5. Let $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$. Then $(\mathbb{Z}[\sqrt{2}], +)$ is a group. For, $\mathbb{Z}[\sqrt{2}]$ is closed with respect to addition. Addition is associative in this set. The identity element is 0 and the inverse of $a + b\sqrt{2}$ is $-a - b\sqrt{2}$. Similarly $\mathbb{Z}[\sqrt{3}]$, $\mathbb{Z}[\sqrt{5}]$, $\mathbb{Z}[\sqrt{6}]$, ... etc. are groups with respect to the usual addition. $(\mathbb{Z}[\sqrt{2}], \cdot)$ is a monoid, the identity element being 1. However, $(\mathbb{Z}[\sqrt{2}], \cdot)$ is not a group. Since $2 = 2 + 0\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$, it does not have an inverse in $\mathbb{Z}[\sqrt{2}]$.

Example 5.6. Let $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. Then $(\mathbb{Q}[\sqrt{2}] \setminus \{0\}, \cdot)$ is a group. The identity element is 1 and the inverse of $a + b\sqrt{2}$ is $(\frac{a}{a^2 - 2b^2}) + (\frac{-b}{a^2 - 2b^2})\sqrt{2}$. Since multiplication is commutative, $(\mathbb{Q}[\sqrt{2}] \setminus \{0\}, \cdot)$ is an Abelian group. Similarly $(\mathbb{Q}[\sqrt{3}] \setminus \{0\}, \cdot)$, $(\mathbb{R}[\sqrt{2}] \setminus \{0\}, \cdot)$, $(\mathbb{R}[\sqrt{3}] \setminus \{0\}, \cdot)$ are all Abelian groups with respect to the usual multiplication.

Example 5.7. Let $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$. Then $(\mathbb{Z}[i], +)$ is a group. The identity element is 0, as $0 = 0 + 0i \in \mathbb{Z}[i]$ and the inverse of $a + bi \in \mathbb{Z}[i]$ is $(-a) + (-b)i \in \mathbb{Z}[i]$. It is an infinite Abelian group. Similarly $(\mathbb{Q}[i], +)$ is also an infinite Abelian group.

Example 5.8. Let $\mathbb{Q}[i]^* = \mathbb{Q}[i] \setminus \{0\}$. Then $\mathbb{Q}[i]^*$ is an Abelian group, the identity element being $1 = 1 + 0i \in \mathbb{Q}[i]^*$. The inverse of $a + bi \in \mathbb{Q}[i]^*$ is $\frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}i \in \mathbb{Q}[i]^*$.

Example 5.9. The n th roots of unity form a finite Abelian group of order n , with respect to multiplication. They are given by

$$z_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n},$$

$$k = 0, 1, \dots, n-1.$$

$$\text{Let } \alpha = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}.$$

Then $\alpha \neq 1$, $\alpha^m \neq 1$ for $0 < m < n$, and $\alpha^n = 1$.

Further, $z_k = \alpha^k$, for $k = 0, 1, 2, \dots, n-1$. Let $G = \{1, \alpha, \dots, \alpha^{n-1}\}$. Then $o(G) = n$.

Step 1 (Closure) If $\alpha^r, \alpha^s \in G$ then $0 \leq r, s \leq n-1$, and

$$\alpha^r \alpha^s = \alpha^{r+s} = \begin{cases} \alpha^{r+s} & \text{if } r+s < n \\ \alpha^{r+s-n} & \text{if } r+s \geq n. \end{cases}$$

Since $r+s-n \leq n-2$, hence $\alpha^r \alpha^s \in G$, the multiplication is a binary operation on G .

Step 2 (Associativity) Associativity in G follows from the associativity (with respect to multiplication) of complex numbers.

Step 3 (Existence of identity) The identity element of G is 1.

Step 4 (Existence of inverse) Inverse of $1 \in G$ is 1. If $\alpha^r \in G$, $0 < r < n$ then $\alpha^{n-r} \in G$ and $\alpha^r \alpha^{n-r} = \alpha^{n-r} \alpha^r = \alpha^n = 1$. Thus the inverse of α^r is α^{n-r} .

Hence every element of G has an inverse in G . Thus G is a group. Moreover, it is Abelian as multiplication of complex numbers is commutative. Since G has n distinct elements, it is a finite Abelian group of order n .

Note that the above example helps us in constructing a finite Abelian group of any given order. This is the first example of a finite Abelian group. So far we have not seen an example of a non Abelian group. This does not mean that all the groups are Abelian. Later on, we shall have plenty of examples of such groups.

Example 5.10. The set $\mathbb{R}^n = \{(a_1, a_2, \dots, a_n) \mid a_1, \dots, a_n \in \mathbb{R}\}$ is a group under the componentwise addition, that is, $(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$. Since addition is a binary operation on \mathbb{R} , therefore it is a binary operation on \mathbb{R}^n . Associativity in \mathbb{R} implies associativity in \mathbb{R}^n . The identity element is $(0, 0, \dots, 0)$ and the inverse of $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ is $(-a_1, -a_2, \dots, -a_n) \in \mathbb{R}^n$.

5.4 Exercise

- If $G = \{-1, 1\}$, show that G is a group with respect to multiplication.
- If $G = \{1, -1, i, -i\}$, where $i^2 = -1$, show that G is a group with respect to multiplication.
- Prove that the set $\mathbb{R}^3 = \{(a_1, a_2, a_3) : a_1, a_2, a_3 \in \mathbb{R}\}$ is a group under componentwise addition, i.e. $(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$.
- Let $S = \mathbb{R} \setminus \{-1\}$. Define $*$ on S by $a * b = a + b + ab$. Show that $(S, *)$ is a group.
 - Find the inverses of 3 and 4.
 - Find a solution of the equation $4 * x * 3 = 5$ in S . Is it unique?
- If \mathbb{Q}^+ denotes the set of positive rational numbers, show that $(\mathbb{Q}^+, *)$, where $a * b = \frac{ab}{2}$ for all $a, b \in \mathbb{Q}^+$ is a group. What is the identity element? What is the inverse of an element $q \in \mathbb{Q}^+$?

6. Show that the six 6th roots of unity form an Abelian group of order 6.
7. Give an example of a group of order (i) 53, (ii) 4021.
8. Show that (G, \cdot) is a group, where $G = \{2^n : n \in \mathbb{Z}\}$.
9. Prove that the set of all rational numbers of the form $3^m 6^n$, where m and n are integers, is a group under multiplication
10. If G is the group of all the 20 roots of unity, what are the pairs of inverses? Can you give a general formula for them
11. Show that $(G, *)$, where $G = \{0, 1, 2\}$ and $a * b = |a - b|$ is not a group. Which of the properties fail to hold?

5.5 Groups of Residues

Before discussing the groups of residues, we shall define a new type of addition and multiplication on \mathbb{Z} .

Addition Modulo n

We now define a new type of addition called “addition modulo n ” and written as $a \oplus_n b$ where a and b are integers and $n > 1$ is a positive integer. Define

$$a \oplus_n b = r, \quad 0 \leq r < n$$

where, r is the least non-negative remainder obtained on dividing $a + b$ by n . Clearly $a \oplus_n b = (a + b) \pmod{n}$. For example, $18 \oplus_6 10 = 4$ since $18 + 10 = 28 = 6 * 4 + 4$. Similarly, $-28 \oplus_3 3 = 2$, as $-28 + 3 = -25 = -9 * 3 + 2$.

Multiplication Modulo n

Let $n > 1$ be a positive integer. Define *multiplication modulo n* , to be written as $a \odot_n b$, as follows:

If a, b are integers, then

$$\begin{aligned} a \odot_n b &= r, \\ 0 &\leq r < n. \end{aligned}$$

where r is the least non-negative remainder obtained on dividing ab by n . Clearly $a \odot_n b = (ab) \pmod{n}$. For example, $9 \odot_7 6 = 5$, because $9 \times 6 = 54 = 7 \times 7 + 5$, and $-3 \odot_6 9 = 3$, since $(-3) \times 9 = -27 = -5 \times 6 + 3$.

Example 5.11. Let $G = \{0, 1, 2, 3, 4, 5\}$. \oplus_6 is a binary operation on G . Given below is the multiplication table for G :

\oplus_6	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

Associativity holds because

$$(a \oplus_6 b) \oplus_6 c = a \oplus_6 (b \oplus_6 c) = (a + b + c)(\text{mod}6)$$

Here the identity element is 0. The pairs of inverses are: 1,5; 2,4; 0,0; 3,3. Note that 0 and 3 are their own inverses. Thus (G, \oplus_6) is a group. It is a finite group of order 6. It is also Abelian.

The group (G, \oplus_6) is denoted by (\mathbb{Z}_6, \oplus_6) . (\mathbb{Z}_6, \oplus_6) is a finite Abelian group of order 6.

Example 5.12. Let $n > 1$ be any integer. Let $G = \{0, 1, 2, \dots, n-1\}$. We prove that (G, \oplus_n) is a group.

Step 1 (Closure) For $a, b \in G$, since the remainder obtained on dividing $a+b$ by n is a non-negative integer less than n , so that $a \oplus_n b \in G$.

Step 2 (Associativity) Associativity of \oplus_n follows from the corresponding property for addition in integers.

Step 3 (Existence of identity) The identity element is '0' because

$$a \oplus_n 0 = 0 \oplus_n a = a \quad \forall a \in G.$$

Step 4 (Existence of inverse) For each $m \in G$, $n-m \in G$ is such that $m \oplus_n (n-m) = (n-m) \oplus_n m = 0$. Hence $n-m$ is the inverse of m .

Thus (G, \oplus_n) is a group. This group is called additive group of integers modulo n and denoted by (\mathbb{Z}_n, \oplus_n) . It is a finite group of order n . Since \oplus_n is commutative, therefore (\mathbb{Z}_n, \oplus_n) is an Abelian group.

The above example helps us to construct a group of any given order.

Is (\mathbb{Z}_6, \odot_6) a group? It can be verified that \mathbb{Z}_6 is closed with respect to multiplication modulo 6. Construct the multiplication table. The identity element is 1. Therefore (\mathbb{Z}_6, \odot_6) is a monoid. The element $0 \in \mathbb{Z}_6$ does not have a multiplicative inverse. Hence (\mathbb{Z}_6, \odot_6) is not group. What can we say about $(\mathbb{Z}_6^*, \odot_6)$. Since $2, 3 \in \mathbb{Z}_6^*$, but $2 \odot_6 3 = 0 \notin \mathbb{Z}_6^*$, thus \mathbb{Z}_6^* is not even closed with respect to \odot_6 . Thus $(\mathbb{Z}_6^*, \odot_6)$ is not even a groupoid. In fact, if n is a composite number, $n = m_1 m_2$ for some $0 < m_1, m_2 < n$. and $\mathbb{Z}_n^* = \{1, 2, \dots, n-1\}$ then, $m_1, m_2 \in \mathbb{Z}_n^*$ but $m_1 \odot_n m_2 = 0 \notin \mathbb{Z}_n^*$. Therefore \odot_n is not a binary operation on \mathbb{Z}_n^* . What can we say when n is prime? The following examples suggest an answer to this question.

Example 5.13. Let $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$.

The multiplication table for \mathbb{Z}_7^* with respect to \odot_7 is:

\odot_7	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

Note that 1 is the identity element. Each element of \mathbb{Z}_7^* is invertible, the pair of inverses are 1,1; 2,4; 3,5 and 6,6. Thus $(\mathbb{Z}_7^*, \odot_7)$ is a group. It is a finite group of order 6.

Example 5.14. Let p be a fixed prime and let $\mathbb{Z}_p^* = \{1, 2, \dots, p-1\}$. We prove that $(\mathbb{Z}_p^*, \odot_p)$ is a group.

Step 1 \odot_p is a binary operation on \mathbb{Z}_p^* , for if $a, b \in \mathbb{Z}_p^*$ then by division algorithm, there exist integers q and r , such that

$$ab = qp + r, \quad 0 \leq r < p$$

If $r = 0$ then $ab = qp \Rightarrow p$ divides ab . Since p is prime, therefore p divides a or p divides b . This is not possible since both a, b are positive integers less than p . Hence $r \neq 0$. Thus $0 < r < p$ i.e. $a \odot_p b = r \in \mathbb{Z}_p^*$.

Step 2 Associativity of \odot_p follows from the corresponding property for multiplication in natural numbers.

Step 3 The identity element is 1 because

$$1 \odot_p a = a \odot_p 1 = a \quad \forall a \in \mathbb{Z}_p^*.$$

Step 4 We now show that each element of \mathbb{Z}_p^* is invertible. Let $a \in \mathbb{Z}_p^*$. Then $1 \leq a < p$. Hence a and p are coprime, so that by Euclid's Algorithm, there exist integers m and n such that $am + pn = 1$. By division algorithm applied to p and m , there exist integers q and r such that $m = pq + r$, $0 \leq r < p$. In case $r = 0$, $m = pq$ so that $apq + pn = 1$, that is $p(aq + n) = 1$. This is impossible, as $p > 1$ and $aq + n$ is an integer. Hence $r \neq 0$. i.e. $0 < r < p$, giving $r \in \mathbb{Z}_p^*$. Now

$$1 = am + pn = a(pq + r) + pn = p(aq + n) + ar.$$

Thus $a \odot_p r = 1$ so that r is the inverse of a . Thus $(\mathbb{Z}_p^*, \odot_p)$ is a group. This group is called the multiplicative group of non-zero integers modulo p and is denoted by $(\mathbb{Z}_p^*, \odot_p)$.

In fact, for any positive integer $n > 1$, the set $\mathbb{Z}_n^* = \{1, 2, \dots, n-1\}$ is a group under multiplication modulo n if and only if n is prime. When n is prime $(\mathbb{Z}_n^*, \odot_n)$ is a group has been proved above.

Conversely, suppose $(\mathbb{Z}_n^*, \odot_n)$ is a group and n is not prime, then $n = rs$ for some integers r and s such that $1 \leq r, s < n$. Then $r, s \in \mathbb{Z}_n^*$ and $r \odot_n s = 0 \notin \mathbb{Z}_n^*$. This contradicts the fact that \mathbb{Z}_n^* is a group, hence n must be prime.

Example 5.15. Let $\mathcal{U}(15) = \{n \in \mathbb{Z} | (n, 15) = 1, 0 < n < 15\}$
Then $(\mathcal{U}(15), \odot_{15})$ is an Abelian group.

$$\mathcal{U}(15) = \{1, 2, 4, 7, 8, 11, 13, 14\}.$$

The multiplication table is:

\odot_{15}	1	2	4	7	8	11	13	14
1	1	2	4	7	8	11	13	14
2	2	4	8	14	1	7	11	13
4	4	8	1	13	2	14	7	11
7	7	14	13	4	11	2	1	8
8	8	1	2	11	4	13	14	7
11	11	7	14	2	13	1	8	4
13	13	11	7	1	14	8	4	2
14	14	13	11	8	7	4	2	1

From the table it is clear that \odot_{15} is a binary operation on $\mathcal{U}(15)$. The binary operation \odot_{15} is associative. 1 is the identity element. The elements 1, 4, 11

and 14 are their own inverses, whereas the inverses of 2, 7, 8, 13 are 8, 13, 2, 7 respectively. Also $a \odot_{15} b = b \odot_{15} a$ for all $a, b \in \mathcal{U}(15)$. Thus $(\mathcal{U}(15), \odot_{15})$ is an Abelian group of order 8.

Example 5.16. Let $n > 1$ be a fixed integer and let $\mathcal{U}(n) = \{m \in \mathbb{N} | m < n, (m, n) = 1\}$. Then

- (i) $(\mathcal{U}(n), \odot_n)$ is a group.
(ii) For $n > 2$, there are at least two elements in $\mathcal{U}(n)$ satisfying $x^2 = 1$. Clearly $1 \in \mathcal{U}(n)$.

- (i) *Step 1* Let $a, b \in \mathcal{U}(n)$. We shall prove that $a \odot_n b \in \mathcal{U}(n)$. Let $a \odot_n b = c$. Then there exists $q \in \mathbb{Z}$ such that $ab = nq + c$.

If $c = 0$, then $n|ab$. Since $(n, a) = 1$, therefore $n|b$, which is a contradiction, because $(n, b) = 1$. Hence $c \neq 0$. We now prove that $(c, n) = 1$. If $(c, n) \neq 1$ then there exists a prime p such that $p|(c, n)$, so that $p|c$ and $p|n$. Hence $p|nq + c$ i.e. $p|ab$. Since p is prime, therefore $p|a$ or $p|b$. Thus $p|(a, n)$ which contradicts the fact that $a \in \mathcal{U}(n)$. Hence $(c, n) = 1$ so that $c = a \odot_n b \in \mathcal{U}(n)$. Thus \odot_n is a binary operation on $\mathcal{U}(n)$.

Step 2 Let $a, b, c \in \mathcal{U}(n)$. It can be easily proved that $(a \odot_n b) \odot_n c =$ remainder obtained on dividing $(ab)c$ by n . Similarly $a \odot_n (b \odot_n c) =$ remainder obtained on dividing $a(bc)$ by n . Since multiplication in \mathbb{Z} is associative, therefore $(a \odot_n b) \odot_n c = a \odot_n (b \odot_n c)$. Hence \odot_n is associative.

Step 3 $1 \in \mathcal{U}(n)$ and $1 \odot_n a = a \odot_n 1 = a \forall a \in \mathcal{U}(n)$. Hence 1 is the identity element in $\mathcal{U}(n)$.

Step 4 Let $a \in \mathcal{U}(n)$. Then $(a, n) = 1$. By Euclid's algorithm, there exist integers x and y such that $ax + ny = 1$. Also by division algorithm applied to n and x , there exist integers q and r such that $x = nq + r$, $0 \leq r < n$. Hence $aqn + ar + ny = 1$ i.e. $n(aq + y) + ar = 1$, so that $a \odot_n r = 1$. Similarly $r \odot_n a = 1$. We claim that $(r, n) = 1$. Suppose $(r, n) \neq 1$. Then there exists a prime p such that $p|r$ and $p|n$ so that $p|(nq + r)$ i.e. $p|x$. Hence $p|(ax + ny)$ i.e. $p|1$ which is not possible as $p > 1$. Thus $(r, n) = 1$, hence $r \in \mathcal{U}(n)$. Also $a \odot_n r = r \odot_n a = 1$. Thus r is the inverse of a .

We have proved that $(\mathcal{U}(n), \odot_n)$ is a group.

- (ii) Clearly $1^2 = 1$. Now $(n-1) \in \mathcal{U}$ and $(n-1)^2 = n^2 - 2n + 1 = n(n-2) + 1$. Thus $(n-1) \odot_n (n-1) = 1$. Thus $x=1$ and $x = n-1$ satisfy $x^2 = 1$. Moreover, for $n > 2$, $n-1 \neq 1$.

5.6 Exercise

1. Prove that $(\mathbb{Z}^*_5, \odot_5)$ is an Abelian group.
2. Let $2\mathbb{Z}^*_5 = \{2, 4, 6, 8\}$. Prove that $(2\mathbb{Z}^*_5, \odot_{10})$ is a group.
3. Let $S = \{2, 4, 8\}$. Prove that (S, \odot_{14}) is a group.
4. Let $G = \{1, 2, 3, 4, 5\}$. Is (G, \odot_6) a group?
5. Show that $\mathcal{U}(10)$ is a group of order 4.
6. What is the order of the group
(i) $\mathcal{U}(20)$ (ii) $\mathcal{U}(30)$ (iii) $\mathcal{U}(40)$.

7. In the group $\mathcal{U}(30)$ find the inverse of the elements 7, 11, 19 and 23.
8. When I was typing a list of nine integers which form a group under multiplication modulo 91, I missed out one element and typed only the eight elements 1, 9, 16, 29, 53, 74, 79 and 81. Can you tell which integer was left out?

5.7 Groups of Matrices

All the groups considered so far are Abelian. Does that lead us to believe that every group is Abelian? Certainly not. The next example answers this question.

Example 5.17. Let $GL(2, \mathbb{Q})$ be the set of all 2×2 non-singular matrices over \mathbb{Q} , the set of rational numbers. This is a group with respect to the usual multiplication of matrices.

Step 1 (Closure) Let $A, B \in GL(2, \mathbb{Q})$, then $|A| \neq 0$ and $|B| \neq 0$. Now AB is a 2×2 matrix over \mathbb{Q} , and $|AB| = |A||B| \neq 0$. Hence $AB \in GL(2, \mathbb{Q})$. Thus multiplication of matrices is a binary operation on $GL(2, \mathbb{Q})$.

Step 2 (Associativity) Associativity in $GL(2, \mathbb{Q})$ follows from the associativity of multiplication of matrices.

Step 3 (Existence of identity) $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in GL(2, \mathbb{Q})$ is the identity element.

Step 4 (Existence of inverse) If $A \in GL(2, \mathbb{Q})$, then $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, for some $a, b, c, d \in \mathbb{Q}$, such that $ad - bc \neq 0$. Let $k = ad - bc$. If $B = \begin{pmatrix} \frac{d}{k} & \frac{-b}{k} \\ \frac{-c}{k} & \frac{a}{k} \end{pmatrix}$, then $|B| = \frac{ad-bc}{k^2} = \frac{1}{k} \neq 0$. Thus $B \in GL(2, \mathbb{Q})$ such that $AB = BA = I$. Hence B is inverse of A and so every element of $GL(2, \mathbb{Q})$ is invertible. We have proved that $GL(2, \mathbb{Q})$ is a group. Clearly it is non-Abelian, for if $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, then $A, B \in GL(2, \mathbb{Q})$ and $AB \neq BA$.

Notation: $GL(n, \mathbb{Q})$ denotes the set of all $n \times n$ non-singular matrices over \mathbb{Q} , and $SL(n, \mathbb{Q})$ denotes the set of all $n \times n$ matrices over \mathbb{Q} with determinant 1.

Clearly, $SL(n, \mathbb{Q}) \subseteq GL(n, \mathbb{Q})$. It can be proved that $SL(n, \mathbb{Q})$ is also a group with respect to multiplication of matrices. Moreover both are non-Abelian. These groups provide a rich source of non-Abelian groups.

Example 5.18. Let $\mathbb{M}_2(\mathbb{Z}_5)$ denote the set of all 2×2 matrices over \mathbb{Z}_5 , integers modulo 5. Then with respect to the usual addition of matrices, where the elements are reduced modulo 5, the set $\mathbb{M}_2(\mathbb{Z}_5)$ is a group. For, clearly addition in $\mathbb{M}_2(\mathbb{Z}_5)$ is a binary operation. Also associative law holds, as it holds for addition of matrices. The null matrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is the identity element. If $A \in \mathbb{M}_2(\mathbb{Z}_5)$,

say $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$; $a, b, c, d \in \mathbb{Z}_5$ then $B = \begin{pmatrix} 5-a & 5-b \\ 5-c & 5-d \end{pmatrix} \in \mathbb{M}_2(\mathbb{Z}_5)$ is such that B is the additive inverse of A .

Since addition of natural numbers is commutative, therefore addition in $\mathbb{M}_2(\mathbb{Z}_5)$ is commutative. Hence $\mathbb{M}_2(\mathbb{Z}_5)$ is an Abelian group. Moreover, it is a finite

group. Each entry in a matrix has 5 choices, hence the number of elements in the group is $5^4 = 625$. Thus $\mathbb{M}_2(\mathbb{Z}_5)$ is a finite Abelian group of order 625.

Notation: $\mathbb{M}_n(F)$ denotes the set of all $n \times n$ matrices over F . If F is a finite set of order m , then $o(\mathbb{M}_n(F)) = m^{n^2}$.

Example 5.19. Let $G = \left\{ \begin{pmatrix} a & 0 \\ 2a & 0 \end{pmatrix} \mid a \in \mathbb{Q}^* \right\}$. Then G is an Abelian group with respect to the usual multiplication of matrices.

Step 1 (Closure) Let $A, B \in G$. Then $A = \begin{pmatrix} a & 0 \\ 2a & 0 \end{pmatrix}, B = \begin{pmatrix} b & 0 \\ 2b & 0 \end{pmatrix}$ for some $a, b \in \mathbb{Q}^*$. Further, $AB = \begin{pmatrix} ab & 0 \\ 2ab & 0 \end{pmatrix} \in G$. Hence multiplication is a binary operation in G .

Step 2 (Associativity) Since multiplication of matrices is associative, therefore associative law holds in G .

Step 3 (Existence of identity) The matrix $E = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \in G$ is identity of G , for, if $A = \begin{pmatrix} a & 0 \\ 2a & 0 \end{pmatrix} \in G$, then $EA = AE = A$.

Step 4 (Existence of inverse) Let $A \in G$. Then $A = \begin{pmatrix} a & 0 \\ 2a & 0 \end{pmatrix}$ for some $a \in \mathbb{Q}^*$. If $B = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{a} & 0 \end{pmatrix}$, then $B \in G$, such that $AB = BA = E$. Hence B is an inverse of A , so that each element of G is invertible. Thus G is a group. Clearly G is Abelian because $AB = BA$ for all $A, B \in G$.

We make two observations from this example. Firstly, that groups of matrices can be Abelian as well as non-Abelian. Secondly, if the determinant of a matrix is zero, then also it can be invertible (with respect to some identity element). Note that this is so because in this case, the identity element is not the usual unit matrix I_2 . In fact $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \notin G$.

Example 5.20. Let $G = \left\{ \begin{pmatrix} a & a \\ a & a \end{pmatrix} \mid a \in \mathbb{Q}^* \right\}$. Then G is an Abelian group with respect to multiplication of matrices.

Step 1 (Closure) If $A, B \in G$, then $A = \begin{pmatrix} a & a \\ a & a \end{pmatrix}, B = \begin{pmatrix} b & b \\ b & b \end{pmatrix}$ for some $a, b \in \mathbb{Q}^*$. Then $AB = \begin{pmatrix} 2ab & 2ab \\ 2ab & 2ab \end{pmatrix} \in G$. Thus multiplication is a binary operation on G .

Step 2 (Associativity) Multiplication in G is associative as multiplication of matrices is associative.

Step 3 (Existence of identity) The matrix $E = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \in G$ is identity element of G .

Step 4 (Existence of inverse) Let $A \in G$. Then $A = \begin{pmatrix} a & a \\ a & a \end{pmatrix}$ for some $a \in \mathbb{Q}^*$. Now $B = \begin{pmatrix} \frac{1}{a} & \frac{1}{a} \\ \frac{1}{a} & \frac{1}{a} \end{pmatrix} \in G$ is such that $AB = BA = E$. Hence B is the inverse of A .

The above steps prove that G is a group. Moreover $AB = BA$ holds for all $A, B \in G$. Hence G is an Abelian group.

Note that every element of G is a singular matrix but still they are invertible in G .

Example 5.21. Let $G = GL(3, \mathbb{Z}_5)$ be the set of all 3×3 non-singular matrices over \mathbb{Z}_5 , integers modulo 5. Then G is a finite non-Abelian group with respect to the usual multiplication of matrices where entries are added and multiplied modulo 5.

Step 1 (Closure) If $A, B \in G$, then $|A| \neq 0$ and $|B| \neq 0$. In fact $|A|, |B| \in \mathbb{Z}_5^*$. Since $|A|, |B| \in \mathbb{Z}_5^*$, therefore $|A|, |B|$ are not multiples of 5 so that $|A||B|$ is not a multiple of 5, that is, $|AB|$ is not a multiple of 5 (as $|AB| = |A||B|$.) Hence AB is non-singular, so that $AB \in G$.

Step 2 (Associativity) Since multiplication of matrices is associative, therefore associative law holds in G .

Step 3 (Existence of identity) The 3×3 unit matrix I_3 is non-singular and is over \mathbb{Z}_5 , so that $I_3 \in G$. Also

$$I_3A = AI_3 = A \text{ for every } A \in G.$$

Hence I_3 is the identity element of G .

Step 4 (Existence of inverse) For any $A \in G$, $|A| \neq 0$. Since $(\mathbb{Z}_5^*, \odot_5)$ is a group so that $|A|$ has a multiplicative inverse in \mathbb{Z}_5^* , say b .

Now,

$$\begin{aligned} & A \text{adj}(A) = |A|I \\ \Rightarrow & bA \text{adj}(A) = b|A|I = I \\ \Rightarrow & Ab \text{adj}(A) = I \\ \Rightarrow & AB = I, \end{aligned}$$

where $B = b \text{adj}(A)$. Similarly $\text{adj}(A)A = |A|I$.
 $\Rightarrow BA = I$, Therefore $A^{-1} = B \in G$. Hence G is a group.

Step 5 G is non-Abelian, as $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$

and $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ are elements of G such that $AB \neq BA$.

Example 5.22. Let $Q = \{I, A, B, C, D, E, F, G\}$, where $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,
 $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,
 $E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $F = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $G = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ under the usual multiplication of matrices. Then Q is a non-Abelian group.

The multiplication table of Q is:

\cdot	I	A	B	C	D	E	F	G
I	I	A	B	C	D	E	F	G
A	A	B	C	I	G	D	E	F
B	B	C	I	A	F	G	D	E
C	C	I	A	B	E	F	G	D
D	D	E	F	G	I	A	B	C
E	E	F	G	D	C	I	A	B
F	F	G	H	E	B	C	I	A
G	G	D	E	F	A	B	C	I

Observe that I is the identity element. The elements I, B, D, E, F and G are their own inverses whereas A, C are inverses of each other.

In a later chapter, when we discuss linear transformations we shall note that these 8 matrices are the matrices of linear transformations of \mathbb{R}^2 . They are rotations about origin through 0° , 90° , 180° and 270° , reflection in the x-axis, y-axis, lines $y = x$ and $y = -x$.

5.8 Exercise

1. Prove that $(\mathbb{M}_2(\mathbb{R}), +)$ is an Abelian group. Is $\mathbb{M}_2(\mathbb{R})$ a group with respect to multiplication? If not, what algebraic structure does $(\mathbb{M}_2(\mathbb{R}), \cdot)$ have? Is it commutative?
2. Let G be the set of all diagonal matrices over \mathbb{R}^* . Prove that G is a group with respect to multiplication of matrices.
3. Let $G = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} : a \in F \right\}$ where
 - (i) $F = \mathbb{Q}^*$ (ii) $F = \mathbb{R}^*$ (iii) $F = \mathbb{C}^*$ (iv) $F = \mathbb{Z}_5^*$
 Prove in each case that G is a group with respect to multiplication of matrices.
4. Prove that $G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$ is a group with respect to matrix multiplication. Is it Abelian?
5. Test whether the following are groups or not. Which of them are Abelian?
 - (i) $(GL(n, \mathbb{Q}), \cdot)$
 - (ii) $(\mathbb{M}_2(\mathbb{Z}), +)$
 - (iii) $(\mathbb{M}_n(\mathbb{R}), +)$
6. Prove that $(\mathbb{M}_3(\mathbb{Z}_6), +)$ is a finite Abelian group. Compute its order.
7. Prove that $(\mathbb{M}_n(\mathbb{Z}_m), +)$ is a finite Abelian group. Compute its order.

5.9 Groups of Functions

Example 5.23. Let S be a non-empty set and let G be the set of all bijective functions from S onto S . Then G is a non-Abelian group with respect to the operation \circ the composition of functions.

Step 1 (Closure) If $f, g \in G$ then f, g are bijective functions on S , so that $f \circ g$ is also a bijective function on S . Hence $f \circ g \in G$. Thus \circ is a binary operation.

Step 2 (Associativity) Since composition of functions is associative, therefore \circ is associative.

Step 3 (Existence of identity) The identity function e on S i.e. $e(x) = x \forall x \in S$ being a bijective function, $e \in G$. Moreover $f \circ e = e \circ f = f \forall f \in G$. Thus e is the identity element of G .

Step 4 (Existence of inverse) Let $f \in G$. Then f is a bijective function, so that f^{-1} is also a bijective function on S . Thus $f^{-1} \in G$, and $f \circ f^{-1} = f^{-1} \circ f = e$. Therefore, every element in G has an inverse.

Thus (G, \circ) is a group. However, it is not Abelian, for, consider $S = \mathbb{R}$, $f(x) = x^3, g(x) = 1 + x$. Then $(f \circ g)(x) = (1 + x)^3, (g \circ f)(x) = 1 + x^3$ so that $f \circ g \neq g \circ f$. Thus (G, \circ) is not Abelian. These groups are called transformation groups and are denoted by $A(S)$.

If S is a finite set having n elements then $A(S)$ is a finite group of order $n!$.

Example 5.24. Let $T_{a,b} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, be defined by

$$T_{a,b}(x, y) = (x + a, y + b) \text{ and let } G = \{T_{a,b} : a, b \in \mathbb{R}\}.$$

Then G is an Abelian group with respect to the composition of mappings as the binary operation.

Step 1 (Closure) Let $T_{a,b}, T_{c,d} \in G$ then for any $(x, y) \in \mathbb{R}^2$

$$\begin{aligned} (T_{a,b}T_{c,d})(x, y) = T_{a,b}(T_{c,d}(x, y)) &= T_{a,b}(x + c, y + d) = (x + c + a, y + d + b) \\ &= (x + a + c, y + b + d) = T_{(a+c), (b+d)} \end{aligned}$$

$\therefore T_{a,b}T_{c,d} = T_{a+c, b+d} \in G$. Hence composition of mappings is a binary operation on G .

Step 2 (Associativity) Associativity holds because composition of mappings is associative.

Step 3 (Existence of identity) The mapping $T_{0,0}$ is the identity of G .

Step 4 (Existence of inverse) For $T_{a,b} \in G, T_{-a,-b} \in G$ is such that $T_{a,b}T_{-a,-b} = T_{0,0}$ using Step 1.

Also $T_{-a,-b}T_{a,b} = T_{0,0}$. Thus $T_{-a,-b}$ is the inverse of $T_{a,b}$.

Step 5 (Commutativity) Since addition of real numbers is commutative

$$T_{a,b}T_{c,d} = T_{a+c, b+d} = T_{c+a, d+b} = T_{c,d}T_{a,b}.$$

Steps 1–5 prove that G is an Abelian group.

Notation: The group G discussed in the above example is usually denoted by $T(\mathbb{R}^2)$. The elements of G are called translations.

Example 5.25. Let G be the set consisting of the six functions f_1, f_2, \dots, f_6 defined on $\mathbb{R} \setminus \{0, 1\}$ by $f_1(x) = x, f_2(x) = 1 - x, f_3(x) = \frac{1}{x}, f_4(x) = \frac{1}{1-x}, f_5(x) = \frac{x-1}{x}, f_6(x) = \frac{x}{x-1}$ and let \circ be the composition of functions. Then (G, \circ) is a non-Abelian group.

Solution: Given $G = \{f_1, f_2, f_3, f_4, f_5, f_6\}$. We construct the composition table

for G :

\circ	f_1	f_2	f_3	f_4	f_5	f_6
f_1	f_1	f_2	f_3	f_4	f_5	f_6
f_2	f_2	f_1	f_5	f_6	f_3	f_4
f_3	f_3	f_4	f_1	f_2	f_6	f_5
f_4	f_4	f_3	f_6	f_5	f_1	f_2
f_5	f_5	f_6	f_2	f_1	f_4	f_3
f_6	f_6	f_5	f_4	f_3	f_2	f_1

Observe the following steps:

Step 1 (Closure) From the multiplication table it is clear that \circ is a binary operation.

Step 2 (Associativity) Since the composition of functions is associative, therefore \circ is associative.

Step 3 (Existence of identity) The mapping f_1 is the neutral element, because

$$f_1 \circ f = f \circ f_1 = f \quad \forall f \in G.$$

Hence f_1 is the identity of G .

Step 4 (Existence of inverse) f_4, f_5 are inverses of each other, whereas others are their own inverses. Thus (G, \circ) is a group.

Also we see from the table that $f_3 \circ f_4 \neq f_4 \circ f_3$. Hence (G, \circ) is a non-Abelian group.

We shall show later that every group of order up to 5 is Abelian. Thus the smallest non-Abelian group is of order 6. The above group is an example of such a group.

5.10 Exercise

1. Show that the set $G = \{f_1, f_2, f_3, f_4\}$ where $f_1(x) = x, f_2(x) = -x, f_3(x) = \frac{1}{x}, f_4(x) = \frac{-1}{x}$ for all $x \in \mathbb{R} \setminus \{0\}$ is a group with respect to the composition of mappings.
2. Let $S = \mathbb{R} \setminus \{0, 1\}$ and let f_i , for $i = 1, 2, \dots, 6$ be functions on S defined by $f_1(x) = x, f_2(x) = 1 - x, f_3(x) = \frac{1}{x}, f_4(x) = \frac{1}{1-x}, f_5(x) = \frac{x-1}{x}, f_6(x) = \frac{x}{x-1}$. If \circ is the operation 'composite of functions', determine which of the following are groups? In case they are groups, are they Abelian?
 - (i) (G_1, \circ) where $G_1 = \{f_1, f_2\}$
 - (ii) (G_2, \circ) where $G_2 = \{f_1, f_3\}$
 - (iii) (G_3, \circ) where $G_3 = \{f_1, f_4\}$
 - (iv) (G_4, \circ) where $G_4 = \{f_1, f_5\}$
 - (v) (G_5, \circ) where $G_5 = \{f_1, f_6\}$
 - (vi) (G_6, \circ) where $G_6 = \{f_2, f_6\}$
 - (vii) (G_7, \circ) where $G_7 = \{f_1, f_4, f_5\}$
 - (viii) (G_8, \circ) where $G_8 = \{f_1, f_2, f_3, f_6\}$

5.11 Group of Subsets of a Set

Example 5.26. Let S be any set and $\mathcal{P}(S)$ be the power set of S . Define Δ on $\mathcal{P}(S)$ by

For $A, B \in \mathcal{P}(S)$, $A\Delta B = (A \setminus B) \cup (B \setminus A)$. Then $(\mathcal{P}(S), \Delta)$ is an Abelian group.

Step 1 (Closure) If $A, B \in \mathcal{P}(S)$ then $A \setminus B$ and $B \setminus A$ are both subsets of S so that $(A \setminus B) \cup (B \setminus A)$ is also a subset of S . Hence $(A \setminus B) \cup (B \setminus A) \in \mathcal{P}(S)$ so that $A\Delta B \in \mathcal{P}(S)$.

Note that $A \setminus B = A \cap B'$, so that $A\Delta B = (A \cap B') \cup (B \cap A')$

Step 2 (Associativity)

Step 3 (Existence of identity) $\phi \in \mathcal{P}(S)$. For, $A \in \mathcal{P}(S)$,

$$A\Delta\phi = (A \cap \phi') \cup (\phi \cap A') = A \cup \phi = A$$

Similarly, $\phi\Delta A = A$ so that

$$A\Delta\phi = \phi\Delta A = A \text{ for all } A \in \mathcal{P}(S).$$

Hence ϕ is the identity element of $\mathcal{P}(S)$.

Step 4 (Existence of inverse) If $A \in \mathcal{P}(S)$, then

$$A\Delta A = (A \setminus A) \cup (A \setminus A) = \phi \cup \phi = \phi$$

Thus $A\Delta A = \phi$. Thus each element of $\mathcal{P}(S)$ is its own inverse.

Step 5 (Commutativity) If $A, B \in \mathcal{P}(S)$ then

$$A\Delta B = (A \setminus B) \cup (B \setminus A) = (B \setminus A) \cup (A \setminus B) = B\Delta A$$

Hence $A\Delta B = B\Delta A$. So that Δ is commutative.

Thus $(\mathcal{P}(S), \Delta)$ is an Abelian group.

If S is an infinite set, then $\mathcal{P}(S)$ is also infinite. But if S is a finite set with n elements then $\mathcal{P}(S)$ has 2^n elements. Hence $(\mathcal{P}(S), \Delta)$ is a finite group and its order is 2^n , where n is the order of S .

5.12 Exercise

1. Let S be any set. Is $(\mathcal{P}(S), \cap)$ a group? If not, why? Which algebraic structure does $(\mathcal{P}(S), \cap)$ have? Is it commutative?
2. Let S be any set and A be a subset of S . Define $G = \{B \subseteq S : B \cap A = \phi\}$. Prove that (G, Δ) is an Abelian group. When will G be a finite group? What is G when A is the null set?

5.13 Groups of Symmetries

These groups have been studied in the previous unit. We shall mention them here for the sake of completeness.

Example 5.27. The 8 symmetries of a square form a group with respect to composition of motions.

It is called the dihedral group of order 8 and is denoted by D_8 . It is a non-Abelian group.

Example 5.28. *The 6 symmetries of an equilateral triangle form a group with respect to composition of motions.*

It is called the dihedral group of order 6 and is denoted by D_6 . It is a non-Abelian group.

Example 5.29. *The 4 symmetries of a non-square rectangle form a group with respect to composition of motions.*

It is known as the Klein's four group and is denoted by V_4 . It is an Abelian group.

Example 5.30. *The $2n$ symmetries of a regular n -gon form a group with respect to composition of motions.*

It is called dihedral group of order $2n$ and is denoted by D_{2n} .

S. No.	Set	Binary Operation	Form of the Element	Identity Element	Inverse	A/NA	Finite/Infinite
1	\mathbb{Z}	+	$z \in \mathbb{Z}$	0	$-z$	A	Infinite
2	\mathbb{Q}	+	$\frac{p}{q}; p, q \in \mathbb{Z}, q \neq 0$	0	$-\frac{p}{q}$	A	Infinite
3	\mathbb{R}	+	$x \in \mathbb{R}$	0	$-x$	A	Infinite
4	\mathbb{C}	+	$a + ib; a, b \in \mathbb{R}$	0	$(-a) + i(-b)$	A	Infinite
5	\mathbb{Q}^*	\times	$\frac{p}{q}; p, q \in \mathbb{Q}, p \neq 0, q \neq 0$	1	$\frac{q}{p}$	A	Infinite
6	\mathbb{R}^*	\times	$x \in \mathbb{Z}, x \neq 0$	1	$\frac{1}{x}$	A	Infinite
7	\mathbb{C}^*	\times	$a + ib; a \neq 0, b \neq 0, a, b \in \mathbb{R}$	1	$\frac{a}{a^2+b^2} + i\frac{-b}{a^2+b^2}$	A	Infinite
8	$m\mathbb{Z}$	+	$mz, z \in \mathbb{Z}$	0	$(-m)z$	A	Infinite
9	$\mathbb{Z}\sqrt{2}$	+	$m + n\sqrt{2}; m, n \in \mathbb{Z}$	0	$(-m) + (-n)\sqrt{2}$	A	Infinite
10	$\mathbb{Q}(\sqrt{2}) \setminus \{0\}$	\times	$p + q\sqrt{2}; p, q \in \mathbb{Q}, p$ and q not both zero	1	$\frac{p}{p^2-2q^2} + i\frac{-q}{p^2-2q^2}\sqrt{2}$	A	Infinite
11	$\mathbb{Z}[i]$	+	$a + ib; a, b \in \mathbb{Z}$	0	$(-a) + i(-b)$	A	Infinite
12	$\mathbb{Q}[i] \setminus \{0\}$	\times	$a + ib; a, b \in \mathbb{Q}, a$ and b not both zero	1	$\frac{a}{a^2+b^2} + i\frac{-b}{a^2+b^2}$	A	Infinite
13	n th root of unity	\times	$\alpha^k, k = 0, 1, \dots, n-1$ where $\alpha = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$	1	α^{n-k}	A	Finite of order n
14	\mathbb{R}^n	+	$(a_1, a_2, \dots, a_n), a_i \in \mathbb{R}$	$(0, 0, \dots, 0)$	$(-a_1, -a_2, \dots, -a_n)$	A	Infinite
15	\mathbb{Z}_m, n is an integer > 1	\oplus_n	$m \in \mathbb{Z}, 0 \leq m < n$	0	$n - m$	A	Finite of order n
16	\mathbb{Z}_p^*, p is prime	\odot_p	$a \in \mathbb{Z}, 1 \leq a < p$	1	$r \in \mathbb{Z}_p^* : m \equiv r \pmod{p},$ where $am + pn = 1$	A	Finite of order $p-1$

Notation: A-Abelian, NA-Non-Abelian, \times - Multiplication.

S. No.	Set	Binary Operation	Form of the Element	Identity Element	Inverse	A / NA	Finite/ Infinite
17	$\mathcal{U}(n), n \in \mathbb{Z}, n \geq 2$	\odot_n	$1 \leq k < n$ s.t. $\gcd(k, n) = 1$	1	solution of $kx \equiv 1 \pmod{n}$	A	Finite of order $\phi(n)$
18	$GL(2, \mathbb{Q})$	\times of matrices	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ s.t. $a, b, c, d \in \mathbb{Q}$ and $ad - bc \neq 0$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} d & -b \\ -c & a \\ ad-bc & ad-bc \end{pmatrix}$	NA	Infinite
19	$\left\{ \begin{pmatrix} a & 0 \\ 2a & 0 \end{pmatrix} : a \in \mathbb{Q}^* \right\}$	\times of matrices	$\begin{pmatrix} a & 0 \\ 2a & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$	$\begin{pmatrix} a^{-1} & 0 \\ 2a^{-1} & 0 \end{pmatrix}$	A	Infinite
20	$M_2(\mathbb{Z}_5)$	$+$ of matrices mod 5	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, a, b, c, d \in \mathbb{Z}_5$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 5-a & 5-b \\ 5-c & 5-d \end{pmatrix}$	A	Finite of order 625
21	$\left\{ \begin{pmatrix} a & a \\ a & a \end{pmatrix} : a \in \mathbb{Q}^* \right\}$	\times of matrices	$\begin{pmatrix} a & a \\ a & a \end{pmatrix}, a \in \mathbb{Q}^*$	$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} \frac{4}{a} & \frac{4}{a} \\ \frac{4}{a} & \frac{4}{a} \end{pmatrix}$	A	Infinite
22	$\mathcal{P}(S)$, where S is any set	symmetric difference, Δ	A , where A is a subset of S .	ϕ , the null set	A	A	$*$ ¹
23	$T(\mathbb{R}^2)$	composite of mappings	For $a, b \in \mathbb{R}$, $T_{a,b}(x, y) = (x + a, y + b)$	$T_{0,0}$	$T_{-a, -b}$	A	Infinite
24	$A(S)$, the set of all bijective mappings on a non-empty set S	composite of mappings	f such that $f : S \rightarrow S$ is a bijective mapping	e , identity mapping	Inverse mapping of f	NA	$*$ ²
25	D_{2n} , the dihedral group	composite of motions	R_α, L_i	R_0	$R_{2\pi-\alpha}, L_i$	NA	Finite of order $2n$

^{*1}: Infinite if S is infinite IF S is a finite having n elements, then $\mathcal{P}(S)$ is finite having 2^n elements.

^{*2}: Infinite if S is infinite. If S is finite having n elements, then $A(S)$ is finite having $n!$ elements.

5.14 Supplementary Exercise

1. State whether the following statements are true or false and justify your answer.
 - (i) The set of rational numbers is a group with respect to multiplication.
 - (ii) The set of integers is a semigroup with respect to subtraction.
 - (iii) In a group some elements may have more than one inverses.
 - (iv) There exists a non-Abelian group of order 6.
 - (v) Every group of order two is Abelian.
 - (vi) The identity element of $(\mathbb{Z}, *)$ where $*$ is defined by $a * b = a - b + 1 \forall a, b \in \mathbb{Z}$ is 1.
 - (vii) (\mathbb{Q}^*, ϕ) is a monoid where ϕ is defined by $a\phi b = |ab| \forall a, b \in \mathbb{Q}^*$.
 - (viii) (\mathbb{N}, \cdot) is a monoid but not a group.
 - (ix) A group may have more than one identity elements.
 - (x) In a group $(G, *)$ for $a, b \in G$ the equation $a * x = b$ has a solution in G .
 - (xi) In a group $(G, *)$ for $a, b \in G$ the equation $(a * x) + b = 0$ has a solution in G .
 - (xii) The null set can be considered to be a group.
 - (xiii) Every semigroup is a monoid.
 - (xiv) Every monoid is a semigroup.
 - (xv) If S is the null set, then $(\mathcal{P}(S), \cup)$ is a group.
 - (xvi) $(\mathbb{Z}_{10}^*, \odot_{10})$ is a group.
2. Give an example of a group in which every element is its own inverse.
3. Can you give an example of a non-Abelian group of order 4?
4. Give five examples of each of the following:
 - (i) Finite Abelian group.
 - (ii) Finite non-Abelian group.
 - (iii) Infinite Abelian group.
 - (iv) Infinite non-Abelian group.
5. Give an example of a group of order
 - (i) 81
 - (ii) 2^9
 - (iii) 5^{16}
 - (iv) p^{n^2}
 - (v) p^{mn} , where p is a prime and m, n are natural numbers.
6. Let $(G, *)$ be a finite group with even number of elements. Show that there exists at least one $a \in G$, different from the identity element e such that $a * a = e$.
7. Give a multiplication table for the binary operation on the set $S = \{e, a, b\}$ of three elements satisfying the properties of the existence of identity and existence of inverse but not the associative law.
8. Let $SL(2, \mathbb{Q})$ be the set of all 2×2 matrices over \mathbb{Q} with determinant 1. Prove that $SL(2, \mathbb{Q})$ is a group with respect to multiplication of matrices. Is it Abelian?
9. Let G be the set of all diagonal matrices over \mathbb{R}^* . Prove that G is a group with respect to multiplication of matrices. Is the group Abelian?
10. Let $G = \{\pm 1, \pm i, \pm j, \pm k\}$ where $i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j$. Prove that G is a group. Is it Abelian? What is the order of the group?
(This is called the Quaternion group and is denoted by Q_8 .)
11. Prove that every group of order 3 is Abelian.
12. Let $G = \{p, q, r, s, t\}$. If G is a group with respect to the binary operation $*$,

then complete the following table, given that p is the identity element.

*	p	q	r	s	t
p	-	-	-	-	-
q	-	r	-	-	p
r	-	s	t	p	-
s	-	t	-	q	r
t	-	-	-	-	-

13. Complete the following table:

Set	Binary Operation	Groupoid	Semi-group	Monoid	Group	Commutative law holds
$\mathbb{N} \cup \{0\}$	Addition	✓	✓	✓	×	✓
\mathbb{Z}	Addition					
\mathbb{Z}	Subtraction					
\mathbb{Q}	Multiplication					
\mathbb{R}^*	Multiplication					
\mathbb{R}^*	Division					
\mathbb{Z}_6	Multiplication modulo 6					
Irrational numbers	Multiplication					
Odd integers	Addition					
$\mathcal{U}(8)$	Multiplication modulo 8					
Odd integers	Multiplication					

14. Let G be a group and let $g \in G$. Define

$$f_g : G \rightarrow G \text{ by } f_g(x) = gxg^{-1} \forall x \in G.$$

- (i) Show that f_g is a bijective function.
 - (ii) Define $\text{Inn}(G) = \{f_g : g \in G\}$. Is $\text{Inn}(G)$ a group with respect to composite of mappings?
 - (iii) Is it Abelian?
15. Let $S = \{5, 15, 25, 35\}$. Show that (S, \odot_{40}) is a group. What is the identity element of this group? Can you see any relationship between S and $\mathcal{U}(8)$?
16. Determine whether each of the following sets form a group under the indicated operation on the elements of the set. In case they do not form a group, state which property fails to hold.

- (i) $S = \{\frac{p}{2^n} : p \in \mathbb{Z}, n \in \mathbb{N}\}$ under addition.
 - (ii) $S = \{x \in \mathbb{C} : x^n = 1, n \in \mathbb{N}\}$ under multiplication.
 - (iii) $S =$ Set of all $n \times n$ matrices over \mathbb{Z} under multiplication.
 - (iv) $S =$ Set of all $n \times n$ matrices over \mathbb{Z} with determinant ± 1 under multiplication.
 - (v) $S = \{a, b \in \mathbb{R}^+ : a * b = a^b\}$ under multiplication.
 - (vi) $S = \{x \in \mathbb{R} : 0 \leq x < 1\}$. On S define $*$ as $x * y = x + y - [x + y]$, where $[]$ denotes the greatest integer function.
 - (vii) Let n be an arbitrary but fixed positive integer. Let $S =$ set of all real polynomials of degree $\leq n$ in x including the zero polynomial; under addition.
 - (viii) Let n be an arbitrary but fixed positive integer. Let $S =$ set of all real polynomials of degree n in x including the zero polynomial; under addition.
 - (ix) Let $S =$ set of all non-zero real polynomials under multiplication.
 - (x) $S =$ Set of all rotations of the plane \mathbb{R}^2 about the origin, with respect to composition of mappings. That if $s * t =$ rotation t followed by the rotation s .
 - (xi) $S =$ set of all translations of the plane \mathbb{R}^2 parallel to a fixed line, with respect to composition of mappings. That is, $s * t =$ translation t followed by translation s .
17. Let $G = \left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} : a \in \mathbb{Z} \right\}$. Prove that G is a group with respect to the usual multiplication of matrices? Is it Abelian?
18. Let S be any set. Prove that $(\mathcal{P}(S), \cup)$ is an Abelian monoid. Is it a group? Justify.
19. Let $S = \mathbb{R} \setminus \{0, 1\}$ and let $g_i, i = 1, 2, 3$ be functions defined on S by $g_1(x) = x, g_2(x) = \frac{1}{1-x}, g_3(x) = \frac{x-1}{x}, x \in S$. If $G = \{g_1, g_2, g_3\}$. Prove that G is a group with respect to composition of mappings. What is the order of G ?
20. Give three examples of groups of order 6.
21. The following are the “definitions” of a group given by students. Are they fully correct? If not, correct them.
- (a) A group is a set G such that
 - (i) the operation is associative.
 - (ii) there is an identity element $\{e\}$ in G .
 - (iii) for any $a \in G$, there is an a' (inverse for each element).
 - (b) A group is a set with a binary operation such that
 - (i) the operation is associative.
 - (ii) An inverse exists.
 - (iii) An identity element exists.
 - (c) A set $(G, *)$ is called a group such that
 - (i) $*$ is associative.
 - (ii) there exists an element e such that $a * e = e * a = e \forall a$
 - (iii) for every element a there exists an element a' such that $a * a' = a' * a = e$.
 - (d) A group G is a set of elements together with a binary operation $*$ such that the following conditions are satisfied.
 - (i) $*$ is associative under addition
 - (ii) There exists $e \in G$ such that $e * x = x * e = x$
 - (iii) There exists an element a' (inverse) such that $a * a' = a' * a = e$ for every $a \in G$.

- (e) A set G is called a group over the binary operation $*$ if
- (i) $*$ has an identity element.
 - (ii) for $a \in G$ there exists $a' \in G$ such that $a * a' = a' * a = e$
 $\forall a \in G$
 - (iii) $*$ is associative over G .

5.15 Answers to Exercises

Exercise - 5.2

1. Since 3, 5 are odd but $3 + 5 = 8$ is not odd, so the closure property fails.
2. $(2 - 3) - 4 \neq 2 - (3 - 4)$. Associativity does not hold.
3. Identity element does not exist.
4. Identity element does not exist.
5. $2 \in N$ does not have an inverse.
6. Groupoid.
7. Groupoid.
8. $b \in G$ does not have an inverse.
9. $(b * c) * d \neq b * (c * d)$. Moreover c has two inverses c and d .

Exercise - 5.4

4. 0 is the identity element. Inverse of $a \in S$ is $-\frac{a}{1+a}$.
 - (i) Inverse of 3 is $-\frac{3}{4}$, Inverse of 4 is $-\frac{4}{5}$
 - (ii) $x = -\frac{7}{10}$, Yes.
5. 2; $\frac{4}{q}$.
7. n^{th} roots of unity for
 - (i) $n = 53$, (ii) $n = 4021$.
10. For $w = \cos \frac{2\pi}{20} + i \sin \frac{2\pi}{20}$; w^r ; w^{20-r} , $0 \leq r \leq 10$ are the pairs of inverses.
For $w = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$, w^r ; w^{n-r} are inverses of each other.
11. $(2 * 1) * 1 \neq 2 * (1 * 1)$, $*$ is not associative. Identity element is 0. Each element is its own inverse.

Exercise - 5.6

1. Construct multiplication table.
2. Construct multiplication table.
3. Construct multiplication table.
4. No, $2 \odot_6 3 = 0 \notin G$.
5. Construct multiplication table.
6. (i)8 (ii)8 (iii)15.
7. 13, 11, 19, 17.
8. 22, Construct multiplication table.

Exercise - 5.8

1. No, Monoid, No.
2. Yes
4. Yes
5. (i) Non-Abelian group
(ii) Abelian group
(iii) Abelian group
6. 6^9 .
7. m^{n^2}

Exercise - 5.10

2. (i), (ii), (v) and (vii) are Abelian groups. Others are not even groups.

Exercise - 5.12

1. Inverse of only identity element S exists. Monoid. Yes.
2. S finite. $\mathcal{P}(S)$.

Supplementary Exercises

1.
 - (i) False, 0, has no inverse.
 - (ii) False, not associative.
 - (iii) False, inverse is unique.
 - (iv) True, D_6 .
 - (v) True,
 - (vi) False, since $1 * 2 = 0 \neq 2$.
 - (vii) False, for \cdot if e is identity, $(-2)\phi e = 2|e| \neq -2$.
 - (viii) True.
 - (ix) False, identity is unique.
 - (x) True.
 - (xi) False, since $(a * x) + b$ is not defined in G .
 - (xii) False, A group is always a non-empty set.
 - (xiii) False. $(\mathbb{N}, +)$.
 - (xiv) True.
 - (xv) True, when $S = \phi$, $\mathcal{P}(S) = \{\phi\}$.
 - (xvi) False, $2, 5 \in \mathbb{Z}_{10}^*$, $2 \circ_{10} 5 \notin \mathbb{Z}_{10}^*$, so not closed.
2. $\mathcal{P}(S)$
3. Does not exist.
4. Look-up the examples in the text.
5. Other answers are also possible.
 - (i) $(\mathbb{Z}_{81}, \oplus_{81}); 81 = 3^4 = 3^{2^2} (\mathbb{M}_2(\mathbb{Z}_3), +)$
 - (ii) $(\mathbb{M}_3(\mathbb{Z}_2), +)$
 - (iii) $(\mathbb{M}_4(\mathbb{Z}_5), +)$
 - (iv) $(\mathbb{M}_n(\mathbb{Z}_p), +)$
 - (v) $(\mathbb{M}_{m \times n}(\mathbb{Z}_p), +)$.
6. Since inverses exist in pairs and e is its own inverse.

7. Table is given by

	<i>e</i>	<i>a</i>	<i>b</i>
<i>e</i>	<i>e</i>	<i>a</i>	<i>b</i>
<i>a</i>	<i>a</i>	<i>a</i>	<i>e</i>
<i>b</i>	<i>b</i>	<i>e</i>	<i>b</i>

8. No.

9. Yes.

10. No. $o(G) = 8$.

12. The complete table is given below:

*	<i>p</i>	<i>q</i>	<i>r</i>	<i>s</i>	<i>t</i>
<i>p</i>	<i>p</i>	<i>q</i>	<i>r</i>	<i>s</i>	<i>t</i>
<i>q</i>	<i>q</i>	<i>r</i>	<i>s</i>	<i>t</i>	<i>p</i>
<i>r</i>	<i>r</i>	<i>s</i>	<i>t</i>	<i>p</i>	<i>q</i>
<i>s</i>	<i>s</i>	<i>t</i>	<i>p</i>	<i>q</i>	<i>r</i>
<i>t</i>	<i>t</i>	<i>p</i>	<i>q</i>	<i>r</i>	<i>s</i>

13. The complete table is given below:

Set	Binary Operation	Groupoid	Semi-group	Monoid	Group	Commutative law holds
$\mathbb{N} \cup \{0\}$	addition	✓	✓	✓	×	✓
\mathbb{Z}	Addition	✓	✓	✓	✓	✓
\mathbb{Z}	Subtraction	✓	×	×	×	×
\mathbb{Q}	Multipli-cation	✓	✓	✓	×	✓
\mathbb{R}^*	Multipli-cation	✓	✓	✓	✓	✓
\mathbb{R}^*	Division	✓	×	×	×	×
\mathbb{Z}_6	Multipli-cation modulo 6	✓	✓	✓	×	✓
Irrational numbers	Multipli-cation	×	×	×	×	×
Odd integers	Addition	×	×	×	×	×
$\mathcal{U}(8)$	Multipli-cation modulo 8	✓	✓	✓	✓	✓
Odd integers	Multipli-cation	✓	✓	✓	×	✓

14. (ii) Yes. (iii) Abelian if G is Abelian.

15. 25 is the identity element.

$\mathcal{U}(8) = \{1, 3, 5, 7\}$. Every element of S is an element of $\mathcal{U}(8)$ multiplied by 5. Every element of the multiplication table is multiplied by 5^2 modulo 40.

16.

- (i) Yes.
- (ii) Yes.
- (iii) No. Existence of Inverse.
- (iv) Yes.
- (v) No. Associativity.
- (vi) Yes. 0 is the identity element, $0^{-1} = 0$ $x^{-1} = 1 - x$, $x \neq 0$.
- (vii) Yes.
- (viii) No. Closure.
- (ix) No. Existence of Inverse.
- (x) Yes.
- (xi) Yes.

17. Unity is I_2 . $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}$. No.

18. ϕ is the identity element. Only identity element is invertible.

19. Identity element is $= g_1$. The elements g_2 , g_3 are inverses of each other. Order of the group is 3.

20. (\mathbb{Z}_6, \oplus_6) , 6th roots of unity, D_6 other examples are possible. First is Abelian, while second is non-Abelian.

21. Look at the definition of a group carefully.

Chapter 6

Group

Properties and Characterization

We have studied a variety of examples of groups in the previous chapter. We will now study some properties shared by all groups. It will be proved that in a groupoid if an identity element exists, it is unique. In examples of groups it was observed that every element had only one inverse. This was not by chance. In fact, we shall prove that, in a group, every element has a unique inverse.

6.1 Properties of Groups

Before discussing the properties of groups, some notations, which will be used throughout, will be in order.

Notation: For a group $(G, *)$ it is tedious to keep on writing the operation $*$ throughout our calculations. Thus, except where necessary, juxtaposition will be used for the binary operation and $a * b$ will be written as ab . In this case we will say that (G, \cdot) , or simply G , is a group. When dealing with special groups, the given group operation will be used.

In view of the generalized associative law, the product of three or more elements of a group will not be bracketed. For the sake of completeness, we prove the uniqueness of identity element.

Theorem 6.1. (*Uniqueness of identity*) *In a monoid G identity element is unique.*

Proof: Let e_1, e_2 be two identity elements in G . Then

$$e_1 a = a e_1 = a \quad \forall a \in G \dots(1)$$

$$e_2 a = a e_2 = a \quad \forall a \in G \dots(2)$$

In (1), taking $a = e_2$, we get $e_1 e_2 = e_2 e_1 = e_2$, and in (2), taking $a = e_1$, we get $e_2 e_1 = e_1 e_2 = e_1$.

Hence $e_1 = e_2$, so that identity element is unique.

In view of the above result we may speak of the identity element in a group. We denote it by e . □

Theorem 6.2. (*Uniqueness of inverse*) In a group G every element has a unique inverse.

Proof: Let G be a group. Suppose an element $a \in G$ has two inverses b and c . Let e be the identity element in G . Then

$$ab = ba = e \quad \dots(1)$$

$$ac = ca = e \quad \dots(2)$$

Now,

$$\begin{aligned} ab &= e \\ \Rightarrow c(ab) &= ce \quad \text{pre multiplying by } c \\ \Rightarrow (ca)b &= c \quad \text{using associative law and property of identity} \\ \Rightarrow eb &= c \quad \text{using (1)} \\ \Rightarrow b &= c \quad \text{using property of identity.} \end{aligned}$$

Hence $a \in G$ has a unique inverse. Since ‘ a ’ is an arbitrary element of G , therefore every element of G has a unique inverse. \square

From Theorems 6.1 and 6.2 it follows that:

- (i) in a monoid, the identity element is unique.
- (ii) in a group, the identity element is unique and every element has a unique inverse.

As a consequence of the above theorem, we can now speak of ‘the inverse’ of an element of a group. We denote the inverse of an element $g \in G$ by g^{-1} . We now define the integral powers of an element of a group.

Definition 6.1. In a group G , for any $g \in G$ and any non-negative integer m , we define

1. $g^0 = e$
2. $g^m = gg \cdots g$ (m -times)
3. $g^{-m} = (g^{-1})^m$

In view of the above definition, we have the following theorem.

Theorem 6.3. In a group G , for any $g \in G$ and for any integers m and n ,

1. $g^m \cdot g^n = g^{m+n} = g^n g^m$
2. $(g^m)^n = g^{mn} = (g^n)^m$
3. $(g^n)^{-1} = (g^{-1})^n$
4. $e^n = e$

Proof: Left to the reader. \square

The above theorem tells us that the familiar laws of exponents for real numbers also hold true for all elements in a group. The laws fail to hold for expressions involving two elements of the group, because, in general, $(ab)^n \neq a^n b^n$. In fact, $(ab)^2 \neq a^2 b^2$. This is shown by the following example.

Example 6.1. Let $G = GL(2, \mathbb{R})$, the group of 2×2 non-singular matrices over the set of real numbers. Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Then $A, B \in G$. It can be verified that $(AB)^2 = \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix}$, $A^2B^2 = \begin{pmatrix} 34 & 21 \\ 21 & 13 \end{pmatrix}$. Hence, in general, $(AB)^2 \neq A^2B^2$.

When dealing with a group whose binary operation is addition denoted by $+$, the inverse of an element g is $-g$. When g is added n -times, it is written as “ ng ”. This should not be confused with $n \cdot g$, as the group operation is addition, not multiplication. Moreover n may not be an element of the group at all. Note that we do not permit non-integral exponents. The following Table 6.1 shows the notations used for multiplicative and additive groups, respectively.

Multiplicative Group	Additive Group
$a \cdot b$ or ab for multiplication	$a + b$ for addition
e or 1 for identity/unity	0 for identity/zero
a^{-1} for inverse of ‘ a ’	$-a$ for inverse of ‘ a ’
a^n for n th power of a	na for a added n -times
ab^{-1} for quotient	$a - b$ for difference

Table 6.1

Theorem 6.4. (Cancellations laws) In a group G left and right cancellation laws hold, that is for $a, b, c \in G$

- $ba = ca \Rightarrow b = c$ (right cancellation law)
- $ab = ac \Rightarrow b = c$ (left cancellation law).

Proof: Let $a, b, c \in G$ be such that

$$ba = ca \dots (1)$$

Since a is invertible, therefore it has an inverse, say a' . Post multiplying (1) by a' we get

$$\begin{aligned} (ba)a' &= (ca)a' \\ \text{Hence } b(aa') &= c(aa') \text{ using associativity.} \\ \text{This gives } be &= ce \text{ using property of inverse.} \\ \text{Thus } b &= c \text{ using property of identity.} \end{aligned}$$

Hence $ba = ca \Rightarrow b = c$, so the right cancellation law holds.

Similarly we can prove that the left cancellation law also holds. \square

As a consequence of the cancellation property, we find that in the multiplication table for a finite group, each element of the group occurs exactly once in each row and in each column. Thus the multiplication table is a Latin square.¹

Theorem 6.5. If G is a group and a, b are any elements of G then

- $(a^{-1})^{-1} = a$

¹Let $S = \{a_1, a_2, \dots, a_n\}$. Then an $n \times n$ array is said to be a Latin square over S if each of its rows and columns is a permutation of a_1, a_2, \dots, a_n .

2. $(ab)^{-1} = b^{-1}a^{-1}$
3. $(a_1 \dots a_n)^{-1} = a_n^{-1} \dots a_1^{-1}$ for every integer n .

Proof:

1. Let $a^{-1} = c$. By definition of inverse of an element $ac = ca = e$, the identity element of G . Then $ca = ac = e$ means that $a = c^{-1} = (a^{-1})^{-1}$. Hence $(a^{-1})^{-1} = a$.
2. Denote $b^{-1}a^{-1}$ by c . Proving the result amounts to showing that c is the inverse of ab , that is, proving that $(ab)c = c(ab) = e$. Now

$$\begin{aligned}
 (ab)c &= (ab)(b^{-1}a^{-1}) \\
 &= ((ab)b^{-1})a^{-1} \text{ using associativity} \\
 &= (a(bb^{-1}))a^{-1} \text{ using associativity} \\
 &= (ae)a^{-1} \text{ using property of inverse} \\
 &= aa^{-1} \text{ using property of identity} \\
 &= e \text{ using property of inverse}
 \end{aligned}$$

Similarly it follows that $c(ab) = e$.

So $(ab)c = c(ab) = e$ and hence $(ab)^{-1} = c = b^{-1}a^{-1}$.

3. Prove by induction on n . □

Remark 6.1. Note that $(ab)^{-1} = b^{-1}a^{-1}$, but in general $(ab)^{-2} \neq b^{-2}a^{-2}$.

For example, in the group $GL(2, \mathbb{R})$, if $A = \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix}$,

then $A^2 = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}$, $B^2 = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, so $A^{-2} = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$, $B^{-2} =$

$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $B^{-2}A^{-2} = \begin{pmatrix} 8 & 11 \\ 5 & 7 \end{pmatrix}$, $AB = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$,

$(AB)^2 = \begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix}$, $(AB)^{-2} = \begin{pmatrix} 7 & 12 \\ 4 & 7 \end{pmatrix}$.

Thus we see that $(AB)^{-2} \neq B^{-2}A^{-2}$.

Theorem 6.6. In a group G , the equations $ax = b$ and $ya = b$ have unique solutions in G for all $a, b \in G$.

Proof: Since G is a group, for each $a \in G$, $a^{-1} \in G$. Consider the equation $ax = b$. Existence of the solution $x \in G$:

Since $a, b \in G$, therefore $a^{-1}b \in G$.

$$\begin{aligned}
 \text{Let } c &= a^{-1}b. \text{ Then} \\
 ac &= a(a^{-1}b) \\
 &= (aa^{-1})b \text{ using associativity} \\
 &= eb \text{ using property of inverse} \\
 &= b \text{ using property of identity}
 \end{aligned}$$

Thus $ac = b$ so that $c \in G$ is a solution of $ax = b$ in G .

Uniqueness of the solution: Suppose c_1, c_2 are two solutions of $ax = b$ in G .

Then $ac_1 = b$, and $ac_2 = b$. Thus $ac_1 = ac_2$. Using left cancellation law in G , we get $c_1 = c_2$. Hence the solution is unique.

Similarly it can be proved that the equation $ya = b$ also has a unique solution $y = ba^{-1}$ in G . \square

Example 6.2. Let G be the group of quaternions. That is $G = \{\pm 1, \pm i, \pm j, \pm k\}$, where $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, $kl = -lk = j$. Verify that:

The equation $x^2 = -1$ has 6 solutions in G , namely $x = \pm i, \pm j, \pm k$.

Similarly $x^4 = 1$ has 8 solutions. In fact, every element of G is a solution.

6.2 Solved Problems

We shall now give some sufficient conditions for a group to be Abelian.

Problem 6.1. If G is a group, and $a, b \in G$ are such that $b = xax^{-1}$ for some $x \in G$, then $b^n = xa^n x^{-1}$, for every integer $n \in \mathbb{Z}$.

Solution: Three cases arise:

Case 1. If n is a positive integer. We prove the result by induction on n . The result obviously holds for $n = 1$.

Let the result hold for $n = k$, i.e.

$$b^k = xa^k x^{-1}$$

Now

$$\begin{aligned} b^{k+1} &= b^k b \\ &= (xa^k x^{-1})(xax^{-1}) \\ &= xa^k (x^{-1}x)ax^{-1} \\ &= xa^k ax^{-1} \\ &= xa^{k+1} x^{-1} \end{aligned}$$

Hence the result holds for $n = k + 1$. The induction is complete and thus, the result holds for every positive integer n .

Case 2. If n is a negative integer, then $n = -m$, where $m > 0$.

$$\begin{aligned} \therefore \quad b^n &= b^{-m} \\ &= (b^m)^{-1} \\ &= (xa^m x^{-1})^{-1} \\ &= xa^{-m} x^{-1} && \text{using the law of inverse for a product} \\ & && \text{of elements.} \\ &= xa^n x^{-1} \\ \therefore \quad b^n &= xa^n x^{-1} \end{aligned}$$

Case 3. If $n = 0$, then $b^0 = e$ and $xa^0 x^{-1} = xex^{-1} = xx^{-1} = e$. Hence $b^0 = xa^0 x^{-1}$.

Combining all the three cases we get $b^n = xa^n x^{-1}$ for all $n \in \mathbb{Z}$.

Problem 6.2. Let G be a group such that if $a, b, c \in G$ and $ab = ca \Rightarrow b = c$ then G is Abelian (this property is called the cross cancellation law).

Solution: Let $a, b \in G$. By the associative property $a(ba) = (ab)a$. The given condition gives that $ba = ab$, so that G is Abelian.

Problem 6.3. If G is a group satisfying $a^2 = e \forall a \in G$, then G is Abelian.

Solution: *Step 1* Let $a \in G$. Then

$$\begin{aligned} a^2 &= e. \\ \Rightarrow a^{-1}(aa) &= a^{-1}e \quad \text{premultiplying by } a^{-1} \\ \Rightarrow (a^{-1}a)a &= a^{-1} \quad \text{using associativity and property of identity} \\ \Rightarrow ea &= a^{-1} \\ \Rightarrow a &= a^{-1}. \end{aligned}$$

Hence every element is its own inverse.

Step 2 Let $a, b \in G$. Then $ab \in G$ implies that $a = a^{-1}$, $b = b^{-1}$ and $ab = (ab)^{-1}$.

Now $ab = (ab)^{-1} = b^{-1}a^{-1}$, by property of inverses.

Hence $ab = ba \forall a, b \in G$. $\therefore G$ is Abelian.

Problem 6.4. If G is a group satisfying $(ab)^2 = a^2b^2 \forall a, b \in G$, then G is Abelian.

Solution: Let $a, b \in G$. Then

$$\begin{aligned} (ab)^2 &= a^2b^2 \\ \Rightarrow (ab)(ab) &= aabb \\ \Rightarrow a(ba)b &= a(ab)b \quad \text{using associativity} \\ \Rightarrow ba &= ab \quad \text{using cancellation laws} \end{aligned}$$

Hence G is Abelian.

Problem 6.5. Let G be a group and a, b be two elements of G satisfying $(ab)^i = a^i b^i$ for three consecutive integers i . Then $ba = ab$. If G has this property for all $a, b \in G$, then G is Abelian.

Solution: Let $n, n+1$ and $n+2$ be three consecutive integers for which the given condition holds. That is,

$$(ab)^n = a^n b^n \tag{6.1}$$

$$(ab)^{n+1} = a^{n+1} b^{n+1} \tag{6.2}$$

$$(ab)^{n+2} = a^{n+2} b^{n+2} \tag{6.3}$$

Then

$$\begin{aligned} (ab)(ab)^{n+1} &= a^{n+2} b^{n+2} \quad \text{using (6.1)} \\ (ab)(a^{n+1} b^{n+1}) &= a^{n+2} b^{n+2} \quad \text{using (6.2)} \\ \Rightarrow ba^{n+1} &= a^{n+1} b \quad \text{using cancellation laws} \end{aligned}$$

Thus

$$ba^{n+1} = a^{n+1} b. \tag{6.4}$$

$$\begin{array}{llll}
\text{Similarly,} & (ab)(ab)^n & = & a^{n+1}b^{n+1} \quad \text{using (6.2)} \\
\Rightarrow & (ab)(a^n b^n) & = & a^{n+1}b^{n+1} \quad \text{using (17.18)} \\
\Rightarrow & ba^n & = & a^n b \quad \text{using cancellation laws} \\
\Rightarrow & ba^{n+1} & = & a^n ba \quad \text{post multiplying by } a \\
\Rightarrow & a^{n+1}b & = & a^n ba \quad \text{using (17.17)} \\
\Rightarrow & ab & = & ba \quad \text{using cancellation laws.}
\end{array}$$

Thus $ab = ba$. If the condition holds for all $a, b \in G$ then we get $ab = ba \quad \forall a, b \in G$, so that G is Abelian.

Remark 6.2. *If the above result holds just for two consecutive integers, then G may not be Abelian. Give an example to prove this.*

Problem 6.6. *A group G is Abelian if and only if $(ab)^n = a^n b^n$ for all $a, b \in G$ and for every positive integer n .*

Solution: Let G be Abelian. Let $a, b \in G$. We prove the result by induction on n . The result obviously holds true for $n = 1$. Let the result holds for $n = k$, i.e.

$$(ab)^k = a^k b^k \quad (6.5)$$

Now

$$\begin{aligned}
(ab)^{k+1} &= (ab)(ab) \cdots (ab), \quad (k+1) - \text{times} \\
&= a(ba)(ba) \cdots (ba)b \\
&= a(ba)^k b \\
&= a(ab)^k b \quad \text{since } G \text{ is Abelian} \\
&= a(a^k b^k)b \quad \text{using (6.5)} \\
&= aa^k b^k b \\
&= a^{k+1} b^{k+1}
\end{aligned}$$

Hence the result holds for $n = k+1$. The induction is complete. Thus the result holds for every positive integer n . Conversely, if $(ab)^n = a^n b^n$ for all $a, b \in G$ and for all $n \in \mathbb{N}$, then the consequence of $n = 2$ is that G is Abelian, as proved in problem 6.4.

Remark 6.3. *The above result does not hold for non-Abelian groups. For example, let $G = GL(2, \mathbb{R})$, the set of all 2×2 non-singular matrices over the set of real numbers. If $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, then $A, B \in G$ are such that $(AB)^2 \neq A^2 B^2$.*

Problem 6.7. *Let G be a group and m, n be two relatively prime integers such that $a^m b^m = b^m a^m$ and $a^n b^n = b^n a^n$ for all $a, b \in G$, then G is Abelian.*

Solution: Since m, n are two relatively prime integers, therefore (by division algorithm) there exist integers x, y such that

$$mx + ny = 1. \quad (6.6)$$

Step 1 We prove that

$$(a^m b^n)^{mk} = (b^n a^m)^{mk} \quad \text{and} \quad (a^m b^n)^{nk} = (b^n a^m)^{nk} \quad \text{for all integers } k.$$

Let $r = mk$. Two cases arise: $k > 0$ or $k < 0$, since for $k = 0$, both sides are trivially equal to the identity e .

Case 1. If $k > 0$, then $r > 0$, so that

$$\begin{aligned}
 (a^m b^n)^r &= (a^m b^n)(a^m b^n) \cdots (a^m b^n) \quad (r - \text{times}) \\
 &= a^m (b^n a^m)(b^n a^m) \cdots (b^n a^m) b^n \\
 &= a^m (b^n a^m)^{r-1} b^n \\
 &= a^m (b^n a^m)^r (b^n a^m)^{-1} b^n \\
 &= a^m (b^n a^m)^r a^{-m} b^{-n} b^n \\
 &= a^m (b^n a^m)^r a^{-m} \\
 &= a^m (b^n a^m)^{mk} a^{-m} \\
 &= a^m [(b^n a^m)^k]^m a^{-m} \\
 &= a^m (c^m a^{-m}) \quad \text{where } c = (b^n a^m)^k \\
 &= (a^m c^m) a^{-m} \\
 &= (c^m a^m) a^{-m} \quad \text{by the given condition} \\
 &= c^m \\
 &= [(b^n a^m)^k]^m \\
 &= (b^n a^m)^{mk} \\
 &= (b^n a^m)^r
 \end{aligned}$$

$$\text{Hence, } (a^m b^n)^r = (b^n a^m)^r. \quad (6.7)$$

Case 2. If $k < 0$, then $r < 0$.

Let $r = -r'$ where $r' > 0$. Observe that

$$\begin{aligned}
 (a^m b^n)^r &= (a^m b^n)^{-r'} \\
 &= [(a^m b^n)^{r'}]^{-1} \\
 &= [(b^n a^m)^{r'}]^{-1} \quad \text{using case 1} \\
 &= (b^n a^m)^{-r'} \\
 &= (b^n a^m)^r
 \end{aligned}$$

Thus we have proved that for every integer k ,

$$(a^m b^n)^{mk} = (b^n a^m)^{mk} \quad (6.8)$$

Similarly,

$$(a^m b^n)^{nk} = (b^n a^m)^{nk} \quad (6.9)$$

Step 2 We shall now prove that for all $a, b \in G$, a^m and b^n commute, i.e.

$$a^m b^n = b^n a^m \quad \text{for all } a, b \in G.$$

Observe that

$$\begin{aligned}
 (a^m b^n) &= (a^m b^n)^1 \\
 &= (a^m b^n)^{mx+ny} \quad \text{using (6.6)} \\
 &= (a^m b^n)^{mx} (a^m b^n)^{ny} \quad \text{using laws of exponents} \\
 &= (b^n a^m)^{mx} (b^n a^m)^{ny} \quad \text{using (11.10) and (11.10)} \\
 &= (b^n a^m)^{mx+ny} \\
 &= (b^n a^m)^1 \\
 &= b^n a^m
 \end{aligned}$$

Therefore,

$$a^m b^n = b^n a^m \quad (6.10)$$

Step 3 Finally let $a, b \in G$. Then

$$\begin{aligned} ab &= a^1 b^1 \\ &= a^{mx+ny} b^{mx+ny} \\ &= a^{mx} a^{ny} b^{mx} b^{ny} \\ &= a^{mx} p^n q^m b^{ny} \quad \text{where } p = a^y, q = b^x \\ &= a^{mx} q^m p^n b^{ny} \quad \text{using (6.10)} \\ &= a^{mx} b^{mx} a^{ny} b^{ny} \\ &= (a^x)^m (b^x)^m (a^y)^n (b^y)^n \\ &= (b^x)^m (a^x)^m (a^y)^n (b^y)^n \quad \text{by the given property in } G \\ &= (b^x)^m (a^x)^m (b^y)^n (a^y)^n \quad \text{by the given property in } G \\ &= (b^x)^m (b^y)^n (a^x)^m (a^y)^n \quad \text{using (6.10)} \\ &= b^{mx} b^{ny} a^{mx} a^{ny} \\ &= b^{mx+ny} a^{mx+ny} \\ &= b^1 a^1 \quad \text{using (6.6)} \\ &= ba \end{aligned}$$

Hence $ab = ba$ for all $a, b \in G$. Thus G is Abelian.

6.3 Exercise

- Let G be a group and $a, b \in G$ such that $ab = ba$. Then prove that
 - $a^{-1}b^{-1} = b^{-1}a^{-1}$.
 - $a^{-1}b = ba^{-1}$.
 - $ab^{-1} = b^{-1}a$.
- Prove that a group G is Abelian if and only if $(ab)^{-1} = a^{-1}b^{-1}$ for all $a, b \in G$.
- If G is a group and $a_1, a_2, \dots, a_n \in G$, prove that is

$$(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} a_{n-1}^{-1} \cdots a_2^{-1} a_1^{-1}.$$

- Let G be a group, and $a, b, c \in G$. Solve the following equations for x , in G :
 - $a^{-1}xa = c$
 - $axb = c$.
- Let G be a finite group having even number of elements. Show that there is at least one element in G , other than identity which is its own inverse.
- In a group G , prove that

$$(x^{-1}ax)^n = x^{-1}a^n x \text{ for all } a, x \in G, n \in \mathbb{Z}.$$
- Prove that in a group, the identity element is the only idempotent element. (Recall that an element $g \in G$ is said to be an idempotent if $g^2 = g$.)

8. Let G be a group and a, b be two elements of G such that $(ab)^n = a^n b^n$ for two consecutive integers n . Do a and b necessarily commute? Justify your answer.

6.4 Characterization of Groups

The axioms used to define a group can actually be weakened considerably. This will give a stronger version of the definition of a group. The following theorems characterize groups with the weaker axioms. First we give some definitions.

Definition 6.2. Let G be a groupoid.

- (i) An element $e_r \in G$ is said to be a right identity if $ae_r = a \forall a \in G$.
(ii) An element $e_l \in G$ is said to be a left identity if $e_l a = a \forall a \in G$.

Definition 6.3. Let G be a groupoid and let $e_r \in G$ be a right identity of G . Then

- (i) an element $a \in G$ has a right inverse with respect to right identity e_r . if there exists $b \in G$ such that $ab = e_r$.
(ii) an element $a \in G$ has a left inverse with respect to right identity e_r . if there exists $c \in G$ such that $ca = e_r$.

Theorem 6.7. Let G be a semigroup. Then G is a group if and only if the following conditions hold:

- (i) There exists $e_r \in G$ such that $ae_r = a \forall a \in G$, i.e. a right identity element e_r exists.
(ii) For each $a \in G$, there exists $a'_r \in G$ such that $aa'_r = e_r$, i.e every element of G has a right inverse with respect to right identities.

Proof: If G is a group then the conditions hold as identity element e is also a right identity and inverse of an element is also a right inverse. Conversely, let the conditions hold. We shall prove that G is a group.

Step 1 Let $a \in G$. By (ii) there exists $a' \in G$ such that $aa' = e_r \dots (1)$

Since $a' \in G$, therefore there exists $a'' \in G$ such that $a'a'' = e_r \dots (2)$

Now

$$\begin{aligned}
 a'a &= a'(ae_r) \\
 &= a'(a(a'a'')) \quad \text{using (2)} \\
 &= a'((aa')a'') \quad \text{using associativity} \\
 &= a'(e_r a'') \quad \text{using (1)} \\
 &= (a'e_r)a'' \\
 &= a'a'' \quad \text{using condition (i)} \\
 &= e_r \quad \text{using (2)}.
 \end{aligned}$$

$\therefore a'a = e_r$, so that we get $aa' = a'a = e_r \dots (3)$.

$$\begin{aligned}
 \text{Step 2} \quad e_r a &= (aa')a \quad \text{using (1)} \\
 &= a(a'a) \\
 &= ae_r \quad \text{using (3)} \\
 &= a \quad \text{using condition (i)}.
 \end{aligned}$$

Thus $ae_r = e_r a = a$ for all $a \in G$. Hence e_r is the identity element of G . Let us denote it by e .

Step 3 By (3) we get

$$aa_r = a_r a = e.$$

Hence a_r is the inverse of a . Thus every element of G has an inverse.

Step 4 Steps 2 and 3 above show that the semigroup G has an identity element and every element is invertible. Hence G is a group. \square

This is a very important result. It tells us that if a binary composition on a set G is associative and there exists a right identity and every element has a right inverse then G is a group. Thus even by assuming weaker conditions, we are able to prove that the algebraic structure is a group. In this sense we can say that this is a stronger version of the definition of a group. The word 'right' in the above theorem can be replaced by 'left', as the following theorem shows.

Remark 6.4. *We used Theorem 6.7 here. In that theorem it is essential for G to be a semigroup. If G is not a semigroup then it may not be a group even if right identity and right inverses exist in it. This can be seen in the following example.*

Example 6.3. *Let \mathbb{Q}^+ be the set of positive rational numbers. Then (\mathbb{Q}^+, \div) is a groupoid, but not a semigroup, since $2 \div (3 \div 4) \neq (2 \div 3) \div 4$.*

Further $a \div 1 = a \forall a \in \mathbb{Q}^+$. Hence 1 is a right identity. Moreover $a \div a = 1$. Hence a is a right inverse of a . Thus \mathbb{Q}^+ has a right identity and every element of \mathbb{Q}^+ has a right inverse, but (\mathbb{Q}^+, \div) is not a group. Note that this is because associative law does not hold in \mathbb{Q}^+ , w.r.t the operation \div .

Theorem 6.8. *Let G be a semigroup. Then G is a group if and only if the following conditions hold:*

- (i) *There exists $e_l \in G$ such that $e_l a = a \forall a \in G$, i.e left identity element e_l exists.*
- (ii) *For each $a \in G$, there exists $a' \in G$ such that $a' a = e_l$, i.e every element of G has a left inverse with respect to some left identity.*

Remark 6.5. *If a semigroup G has an one sided identity and the other sided inverse, then it need not be a group. In other words, if a semigroup G has a left identity and every element has a right inverse, then G need not be a group. This can be seen from the following example.*

Example 6.4. *Let $G = \{a, b, c, d\}$. Define \circ on G as: $x \circ y = y$ for all $x, y \in G$. The multiplication table for G is*

\circ	a	b	c	d
a	a	b	c	d
b	a	b	c	d
c	a	b	c	d
d	a	b	c	d

If $x, y, z \in G$ then

$$(x \circ y) \circ z = y \circ z = z$$

$$x \circ (y \circ z) = x \circ z = z$$

Thus $(x \circ y) \circ z = x \circ (y \circ z)$, so that the operation is associative. Hence (G, \circ) is a semigroup. We see that $a \circ x = x$ for all $x \in G$. Thus a is left identity. For any $x \in G$, from the table

$$x \circ a = a.$$

Thus every element $x \in G$ has a right inverse, namely, 'a'. If G were a group, then by Theorem 6.4, cancellation laws must hold in G . However from the table $b \circ a = a$ and $c \circ a = a$. Then $b \circ a = c \circ a$ but $b \neq c$. Thus cancellation laws do not hold in G and therefore G can not be a group.

Theorem 6.9. *Let G be a semigroup. Then G is a group if and only if the equations $ax = b$ and $ya = b$ are solvable in G for all $a, b \in G$.*

Proof: Let G be a semigroup. First suppose that G is a group, then the equations $ax = b$ and $ya = b$ are solvable in G for all $a, b \in G$, follows from Theorem 6.6.

Conversely, suppose that for all $a, b \in G$ the equations $ax = b$ and $ya = b$ are solvable in G .

Step 1 (G has a right identity element). Let $a \in G$. Consider the equation $ax = a$. This equation is solvable in G . Thus there exists $e \in G$ such that $ae = a$. We shall prove that e is a right identity for G . Let $g \in G$. Then the equation $ya = g$ has a solution in G , so that there exists $h \in G$ such that $ha = g$. Now $ge = (ha)e = h(ae) = ha = g$. Thus $ge = g$ for all $g \in G$, so that e is a right identity in G .

Step 2 (Every element has a right inverse in G .) Let $a \in G$. Then the equation $ax = e$ has a solution in G , say a' . Thus, we get $aa' = e$, so that every element in G has a right inverse in G .

Step 3 Steps 1 and 2 above show that G is a semigroup with a right identity and that every element in G has a right inverse. Thus by Theorem 6.7, G is a group. \square

Theorem 6.10. *A finite semigroup is a group if and only if cancellation laws hold.*

Proof: Let G be a finite semigroup. If G is a group, then cancellation laws hold in G by theorem 6.4. Conversely, suppose cancellation laws hold in G . Since G is finite, let $G = \{a_1, a_2, \dots, a_n\}$.

Step 1 Let $a \in G$. Then $aa_i \in G$ for all $i = 1, 2, \dots, n$. Let $P = \{aa_1, aa_2, \dots, aa_n\}$. Then $P \subseteq G$. We assert that all the elements of P are distinct, for if $aa_i = aa_j$ for some $i \neq j$; then by left cancellation law $a_i = a_j$, which is a contradiction. Hence P has exactly n elements. Now $P \subseteq G$. Since G is finite and both P and G have the same number of elements, therefore $P = G$. Let $b \in G$ and since $G = P$, therefore $b \in P$ so there exists $a_i \in G$ such that $aa_i = b$. Hence for all $a, b \in G$, $ax = b$ has a solution in G .

Step 2 As in Step 1, considering $Q = \{a_1a, \dots, a_na\}$ we can prove that for all $a, b \in G$, the equation $ya = b$ has a solution in G .

The two steps using Theorem 6.9, prove that G is a group. \square

Remark 6.6. In Theorem 6.10, none of the following conditions on G can be dropped:

- (i) finiteness of G .
- (ii) associativity in G .

(i) (\mathbb{Z}, \cdot) is an infinite semigroup in which cancellation laws hold but is not a group.

(ii) (\mathbb{Q}^+, \div) is a groupoid in which cancellation laws hold but (\mathbb{Q}^+, \div) is not a group.

In view of the above theorems, the task of verifying a given system for being a group becomes bit easier. For example, for the non-commutative systems we need not worry about finding both-sided identities and inverses (in the commutative systems, one-sided identities and inverses automatically become both-sided identities and inverses).

6.5 Solved Problems

Problem 6.8. Let G be a semigroup such that there exists $e \in G$ satisfying $eg = g \forall g \in G$. Also let for each pair of distinct elements $a, b \in G$, there exist a solution of the equation $ya = b$ in G . Prove that G is a group.

Solution: We have $eg = g \forall g \in G$. Hence e is a left identity in G . Let $a \in G$. Then $ya = e$ has a solution (according to the given condition) in G , say b , so that $ba = e$. Hence a has a left inverse. Thus G is a semigroup having a left identity and each element having a left inverse. Hence G is a group.

Problem 6.9. Let $G = \{e, x, y, z\}$. A binary composition \circ on G is defined by the following table as:

\circ	e	x	y	z
e	e	x	y	z
x	y	z	e	x
y	x	y	z	e
z	z	e	x	y

Prove that cancellation laws hold in G . Is (G, \circ) a group?

Solution: In the given composition table, in any row or column no two elements are repeated. This implies that cancellation laws hold in G . However, G is not a group, as there is no identity element.

Remark 6.7. In the above example G , is a finite groupoid in which cancellation laws hold, but it is not a group. This shows that for the conclusion of theorem 6.10 to hold true, the assumption of associativity of G cannot be dropped.

Problem 6.10. Let G be a finite group with identity e . Prove that

- (a) the number of elements x of G such that $x^2 \neq e$ is even.
- (b) the number of non-identity elements that satisfy the equation $x^3 = e$ is even.
- (c) the number of non-identity elements that satisfy the equation $x^5 = e$ is a multiple of 4.

Solution:

(a) Let $g \in G$. If $g^2 = e$ then $g = g^{-1}$. Thus, if $g^2 \neq e$, then the inverse h of g , is such that $h \neq g$, and $g^2 \neq e$, $h^2 \neq e$. Thus such elements can be paired $\{g, g^{-1}\} \forall g \neq g^{-1}$. Hence the number of such elements is even.

(b) Consider $g \in G$, $g \neq e$ such that $g^3 = e$. We assert that g is such that:

- (i) $g^2 \neq e$ (ii) $g \neq g^2$ (iii) $(g^2)^3 = e$. For,
 (i) if $g^2 = e$ then $g = e$, which is a contradiction. Hence $g^2 \neq e$.
 (ii) if $g = g^2 \Rightarrow g^3 = e$ implies $g = e$, a contradiction. Hence $g \neq g^2$.
 (iii) Clearly $g^6 = (g^2)^3 = (g^3)^2 = e$. Thus g, g^2 are both non identity distinct elements satisfying $x^3 = e$. Such elements always occur in pairs, as $\{g, g^2\}$ so they are even in number and result is proved.

(c) Consider $g \in G, g \neq e$ such that $g^5 = e$. We assert that

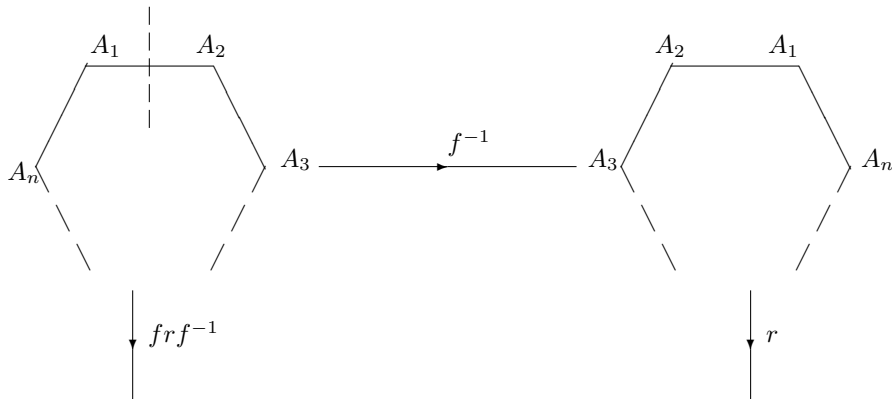
- (i) each of g, g^2, g^3, g^4 is different from identity.
 $g \neq e$ by assumption. $g^2 = e$
 $\Rightarrow g^5 = e$
 $\Rightarrow g = (g^2)^3(g^5)^{-1} = e.g^3 = e$
 $\Rightarrow g^5 = e$
 $\Rightarrow g = (g^3)^2(g^5)^{-1} = e.g^4 = e, g^5 = e$
 $\Rightarrow g = (g^4)^{-1} = e$.
 (ii) the elements g, g^2, g^3, g^4 are all distinct.
 Any two powers of g are equal implies $g = e$ or $g^2 = e$ or $g^3 = e$.
 (iii) each of the elements g, g^2, g^3, g^4 satisfy $x^5 = e$.

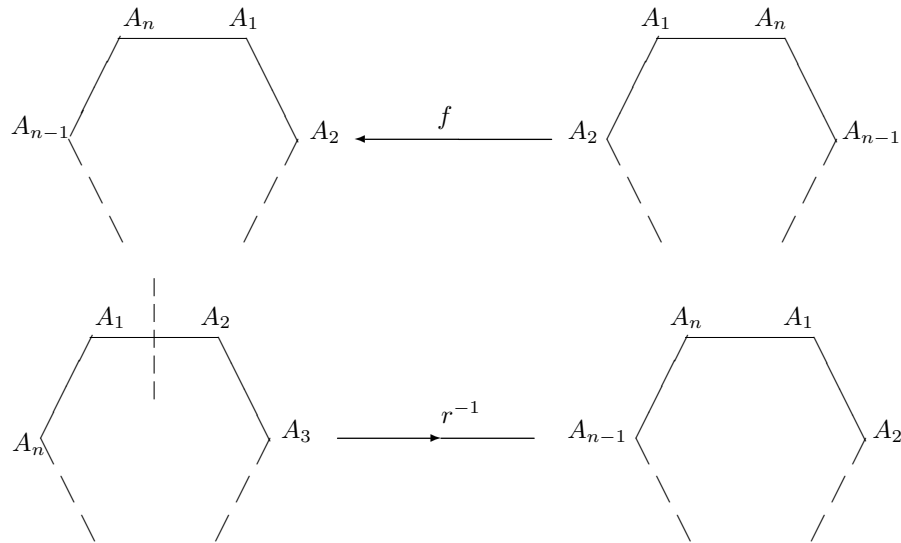
The proof is similar to the proof of part (b).

Problem 6.11. In D_{2n} , let r be a rotation by an angle $\frac{2\pi}{n}$ radians, in the anticlockwise direction. Use a diagram to verify $frf^{-1} = r^{-1}$, where f is any reflection. Use this relation to write the following elements in the form r^i or $r^i f$, where $0 \leq i < n$.

- (a) in D_8 , $fr^{-2}fr^5$
 (b) in D_{10} , $r^{-3}fr^4fr^{-2}$
 (c) in D_{12} , $fr^5fr^{-2}f$.

Solution: Let us consider the reflection f about the line which is the perpendicular bisector of one of the sides and rotation r in the anticlockwise direction by an angle of $\frac{2\pi}{n}$ radians. Observe the following diagrams:





We find that $frf^{-1} = r^{-1}$.

Since f is any reflection, then f^2 is the identity motion, that is, no motion at all. We shall say $f^2 = e$, the identity of the group D_{2n} . Hence $f^{-1} = f$. Further, $(frf^{-1})^k = fr^k f^{-1} \forall k \in \mathbb{Z}$. In particular in D_{2n} , $r^n = e \Rightarrow (frf^{-1})^n = e$. We shall use these two facts for proving the properties.

(a) In D_8 , $r^4 = e$. If $g = fr^{-2}fr^5$ then

$$g = (fr^{-2}f^{-1})r^5 = (frf^{-1})^{-2}r^5 = (r^{-1})^{-2}r^5 = r^2r^5 = r^7 = r^3.$$

(b) In D_{10} , $r^5 = e$. If $h = r^{-3}fr^4fr^{-2}$ then

$$h = r^{-3}fr^4f^{-1}r^{-2} = r^{-3}(frf^{-1})^4r^{-2} = r^{-3}r^{-4}r^{-2} = r^{-9} = r^{-9(mod5)} = r.$$

(c) In D_{12} , $r^6 = e$. If $x = fr^5fr^{-2}f$ then

$$x = fr^5f^{-1}r^{-2}f = (frf^{-1})^5r^{-2}f = r^{-5}r^{-2}f = r^{-7}f = r^{-1}f = r^5f.$$

Problem 6.12. Let G be a group and $g \in G$. Define $f_g : G \rightarrow G$ by $f_g(x) = gxg^{-1}$ for all $x \in G$. For $g, h \in G$ prove that $f_{gh} = f_g f_h$.

Solution: Let $x \in G$. Then

$$\begin{aligned} f_{gh}(x) &= (gh)x(gh)^{-1} \\ &= (gh)x(h^{-1}g^{-1}) \\ &= ghxh^{-1}g^{-1} \\ &= g(hxh^{-1})g^{-1} \\ &= f_g(hxh^{-1}) \\ &= f_g(f_h(x)) \\ &= (f_g f_h)(x) \end{aligned}$$

Thus, $f_{gh} = f_g f_h$.

Problem 6.13. Let G be a semigroup such that for every $a \in G$, there exists a unique $b \in G$, such that $aba = a$. Prove that G is a group.

Solution:

Step 1 There exists an idempotent in G .

Let $a \in G$. Then there exists a unique $b \in G$ such that

$$\begin{aligned} aba &= a \\ \Rightarrow abab &= ab && \text{on post - multiplying by } b \\ \Rightarrow (ab)(ab) &= ab \\ \Rightarrow (ab)^2 &= ab \end{aligned}$$

Hence ab is an idempotent element. Let $ab = e$. Thus $e^2 = e$.

Step 2 The idempotent is unique.

Claim: e of *Step 1* is the only idempotent element of G . Let f be another idempotent in G . Then $f^2 = f$. Hence there exists a unique $g \in G$ such that

$$\begin{aligned} (ef)g(ef) &= ef && \dots(1) \\ \Rightarrow ef(gef)g &= efg && \text{on post - multiplying by } g \\ &&& \text{and using associativity} \\ \Rightarrow ef(gef)g(ef) &= (ef)g(ef) && \text{on post - multiplying by } ef \\ &&& \text{and using associativity} \\ \Rightarrow ef(gef)g(ef) &= (ef)g(ef) \\ \Rightarrow ef(gef)g(ef) &= ef && \text{using (1).} \end{aligned}$$

By uniqueness of g , we get

$$gef = g \quad \dots(2)$$

Also

$$\begin{aligned} ef(ge)ef &= efge^2f \\ &= efgef \quad \text{since } e^2 = e \\ &= (ef)g(ef) \\ &= ef \quad \text{using (1)} \end{aligned}$$

$\therefore ef(ge)ef = ef$. Again by uniqueness of g , we get $ge = g \dots(3)$

Similarly, by proving $ef(fg)ef = (ef)g(ef) = ef$, we get $fg = g \dots(4)$

$$\begin{aligned} (3) \text{ and } (2) &\Rightarrow gfg = gef = g \dots(5) \\ \Rightarrow gg &= g && \text{using (4)} \\ \Rightarrow g^2 &= g && \dots(6) \\ \Rightarrow &&& g \text{ is an idempotent.} \end{aligned}$$

Now $g^3 = g^2g = gg = g^2 = g$ using (6).

$\therefore g^3 = g \implies ggg = g \dots(7)$

Post multiplying (6) by fg , we get

$$\begin{aligned} g^2fg &= gfg \\ \Rightarrow g(g)fg &= gfg \\ \Rightarrow g(ge)fg &= gfg && \text{(using (3))} \\ \Rightarrow g^2efg &= gfg \\ \Rightarrow g(ef)g &= gfg && \text{(using (6))} \\ \Rightarrow g(ef)g &= g && \text{(using (5))} \end{aligned}$$

By uniqueness of g in (7) we get

$$ef = g$$

Hence ef is an idempotent. Thus

$$\begin{aligned} & (ef)(ef) = ef \\ \Rightarrow & (eff)(ef) = ef \quad \text{since } f \text{ is an idempotent} \\ \Rightarrow & (ef)f(ef) = ef \quad \dots (8) \end{aligned}$$

Similarly, using that $g = ef$ is an idempotent we get

$$(ef)e(ef) = ef \quad \dots (9)$$

Using (8) and (9) and the uniqueness of f , we get

$$f = e.$$

Hence G has exactly one idempotent element.

Step 3 If $e = ab$ as in *Step 1*, then $e = ab = ba$.

Let $a \in G$. Then there exists a unique $b \in G$ such that $aba = a$. Pre-multiplying by b we get

$$\begin{aligned} & baba = ba \\ \Rightarrow & (ba)^2 = ba \end{aligned}$$

\Rightarrow ba is an idempotent.

This gives us three idempotents namely e , ab and ba . Since G has a unique idempotent e , we get $ba = e = ab$.

Step 4 We shall now prove that G has a left identity.

Let $a \in G$. Then there exists a unique $b \in G$ such that $aba = a$. Since $ab = e$, we get $ea = a$ for all $a \in G$. Thus e is a left identity of G .

Step 5 Every element of G has a left inverse.

$$\begin{aligned} & aba = a \\ \Rightarrow & b(aba) = ba \quad \text{pre multiplying by } b \\ \Rightarrow & (bab)a = e \quad \text{using Step 3 that } ba = e \end{aligned}$$

\Rightarrow bab is a left inverse of a .

Hence every element of G has a left inverse in G . We have proved that G is semigroup having a left identity and every element of G has a left inverse. Hence G is a group.

6.6 Exercise

1. Let G be a semigroup such that there exists $e \in G$ satisfying $ge = g \quad \forall g \in G$. If for each pair of distinct elements a, b of G , there exists a solution of $ax = b$ in G then prove that G is a group.
2. Let G be a semigroup such for all $a, b \in G$, there exists a solution of the equation $ax = b$ in G . Is G a group? Justify.

3. Let G be a semigroup such for all $a, b \in G$, there exists a solution of the equation $ya = b$ in G . Is G a group? Justify.
4. Let $G = \left\{ \begin{pmatrix} a & b \\ a & b \end{pmatrix} : a, b \in \mathbb{R}, a + b \neq 0 \right\}$, then show that
- G is a semigroup under matrix multiplication.
 - G has a left identity.
 - Each element of G has a right inverse.
 - G is a group.
5. Let $G = \{e, a, b, c, d\}$. A binary composition $*$ on G is defined by the following table:

$*$	e	a	b	c	d
e	e	a	b	c	d
a	a	b	d	e	c
b	b	c	e	d	a
c	c	d	a	b	e
d	d	e	c	a	b

Prove that the cancellation laws hold in G . Is $(G, *)$ a group?

6. If a semigroup has a right identity, is it necessarily unique? Justify.
7. Let where $\omega \neq 1$ is such that $\omega^3 = 1$. Show that
- G is a semigroup with respect to multiplication of matrices.
 - Cancellation laws hold in G .
 - G is a group.
- Is G Abelian? Justify.
8. Let G be a group with identity e . Let p be a prime number. Prove that the number of the non-identity elements satisfying $x^p = e$ is a multiple of $p - 1$.

6.7 Supplementary Exercises

1. State whether following statements are true or false. Justify your answers.
- The empty set is a group.
 - A group has at least one identity element.
 - A group can have more than one identity element.
 - Every group has at least one idempotent.
 - Every group has at most one idempotent.
 - Every group has exactly one idempotent.
 - In a group G , if $a, b \in G$ are such that $a^2 = b^2$ then $a = b$ or $a = -b$.
 - If every element of a group G is its own inverse, then G is Abelian.
 - If G is a group, then $(ab)^n = a^n b^n$ for all $a, b \in G$.
 - In a group G , every linear equation $ax + b = c$; $a, b, c \in G$ has a solution.
 - In a group G , every linear equation $ax = b$; $a, b \in G$ has a solution.
 - If every element 'a' of a group G satisfies $x^2 = e$, then G is Abelian.

- (xiii) In a group G , every equation $x^2 = x$, $x \in G$ has exactly two solutions 1 and 0.
 - (xiv) There exists a group in which cancellation laws do not hold.
 - (xv) In a group G , if $(ab)^2 = a^2b^2 \quad \forall a, b \in G$ then G is Abelian.
 - (xvi) In a group G , if $(ab)^n = a^n b^n \quad \forall a, b \in G$ and some positive integer $n > 2$, then G is Abelian.
 - (xvii) A semigroup G with a left identity in which every element of G has a left inverse is a group.
 - (xviii) A semigroup G with a left identity in which every element of G has a right inverse is a group.
 - (xix) A semigroup G with an identity element (i.e monoid) in which elements of G have a right inverse or a left inverse, is a group.
2. On the set \mathbb{R}^* of nonzero real numbers, define \circ by $a \circ b = |a|b$. Show that
- (i) \circ is a binary operation on \mathbb{R}^* .
 - (ii) \circ has a left identity.
 - (iii) every element of \mathbb{R}^* has a right inverse.
 - (iv) is (\mathbb{R}^*, \circ) a group?
- Explain the significance of this question.
3. In a semigroup one sided identity is unique? Justify.
4. Let G be a group such that $(ab)^n = a^n b^n$ for two consecutive integers n , and for all $a, b \in G$. Is then G Abelian?
5. Give an example of an infinite semigroup in which cancellation laws hold, but which is not a group.
6. Give an example of a finite groupoid which is not a group but in which cancellation laws hold.

6.8 Answers to Exercises

Exercise - 6.3

- 4. (i) aca^{-1} .
- (ii) $a^{-1}cb^{-1}$.
- 5. *Hint:* There is an even number of elements which are not their own inverses.
- 6. *Hint:* By induction prove that $(x^{-1}ax)^n = x^{-1}a^n x$ for all $n \in \mathbb{Z}^+$. Then take inverses of both sides.
- 7. *Hint:* Use cancellation law.
- 8. *Hint:* No. Consider the group of Quaternions, and take $a = i$, $b = j$, $n = 4, 5$.

Exercise - 6.6

- 1. e is the right identity. By solving $ax = e$, every element has a right inverse in G .
- 2. No. Consider $(G, *)$ with G having at least two elements $*$ is defined by $a * b = b \quad \forall a, b \in G$.

3. No. Consider $(G, *)$ with G having at least two elements a, b such that $a * b = a \forall a, b \in G$.
4. (ii) $\begin{pmatrix} a & 1-a \\ a & 1-a \end{pmatrix}$ for $a \in \mathbb{R}$ is a left identity.
 (iii) $\begin{pmatrix} \frac{1}{a+b} & 0 \\ \frac{1}{a+b} & 0 \end{pmatrix}$ is a right inverse of $\begin{pmatrix} a & b \\ a & b \end{pmatrix}$.
 (iv) No, left identity is not a right identity.
5. $(G, *)$ is not a group as $*$ is not associative as $a(bc) \neq (ab)c$.
6. No. $G = \left\{ \begin{pmatrix} x & x \\ y & y \end{pmatrix} : x, y \in \mathbb{R} \right\}$ is a semigroup under multiplication, in which $\begin{pmatrix} x & x \\ 1-x & 1-x \end{pmatrix}$ is a right identity for every $x \in \mathbb{R}$.
7. *Hint:* G is not an Abelian group. Form the multiplication table.
8. *Hint:* $x, x^2, \dots, x^{p-1}, x^p = e$ are all distinct.

Supplementary Exercises

2.
 - (i) False
 - (ii) True
 - (iii) False
 - (iv) True
 - (v) False
 - (vi) True
 - (vii) False, since in $U(10)$ we have $3^2 = 7^2$. Also, in Klein's 4 group, $G = \{e, a, b, ab\}$, in which $a^2 = b^2 = (ab)^2 = e$.
 - (viii) True
 - (ix) False
 - (x) False
 - (xi) True
 - (xii) True
 - (xiii) False
 - (xiv) False
 - (xv) True
 - (xvi) False, take $G = D_6, n = 6$.
 - (xvii) True
 - (xviii) False
 - (xix) False

2. (iv) No.

The significance of this problem is that it shows that formally weaker axioms for a group must either be all left axioms or all right axioms and not half and half.

3. No, find a counter-example.
4. No, group of Quaternions, with $n = 4, 5$.
5. (N, \cdot) is one such example.
6. See Exercise ?? , Q5.

Chapter 7

Subgroups

7.1 Criteria for Subgroups

While studying examples of groups we had groups contained within larger groups. For example, \mathbb{Z} the group of integers under addition is contained within the larger group \mathbb{Q} of rationals under addition, which in turn is contained within the group \mathbb{R} of reals under addition. The best way to study any algebraic structure is to study its subsets, which themselves have the same structure. Therefore, we study subsets of a group which are groups in their own right. They are called subgroups. Thus we have the following definition.

Definition 7.1. *Let G be a group. A subset H of G is called a subgroup of G if H is a group under the operation of G restricted to H .*

Notation: If H is a subgroup of a group G , then we shall write $H \leq G$. Further, if $H \neq G$ then we shall write $H < G$.

Since the operation of G has been restricted to H , therefore we shall denote the operation for the group G and for the subgroup H by the same symbol.

It is possible that H has the structure of a group with respect to some operation other than the operation on G restricted to H . For example, (\mathbb{Q}^+, \cdot) and $(\mathbb{R}, +)$ are groups, $\mathbb{Q}^+ \subseteq \mathbb{R}$ but $(\mathbb{Q}^+, \cdot) \not\subseteq (\mathbb{R}, +)$. This is because the operation \cdot is not a restriction of $+$ to \mathbb{Q}^+ .

A natural question which comes to our mind is: Can a group and its subgroup have different identity elements? The following theorem answers this question.

Theorem 7.1. *If H is a subgroup of a group G then*

- (i) *the identity element of H is the same as that of G .*
- (ii) *for any $a \in H$, inverses of a in H is the same as the inverse of a in G .*

Proof: Let e be the identity element of G .

$$\therefore \quad ae = ea = a \quad \forall \quad a \in G \quad (7.1)$$

- (i) Let e' be the identity of the subgroup H . Then

$$ae' = e'a = a \quad \forall \quad a \in H \quad (7.2)$$

Using 7.1, we get in particular

$$ae = ea = a \quad \forall a \in H \quad (7.3)$$

$$7.2 \text{ and } 7.3 \Rightarrow ae = ae' \quad \forall a \in H.$$

Since $H \subseteq G$, therefore this is an equation in G . Using cancellation laws in G , we get

$$e = e'.$$

(ii) Let $a \in H$.

Let a^{-1} be the inverse of a in G and b the inverse of a in H .

Then

$$aa^{-1} = a^{-1}a = e \quad (7.4)$$

and

$$ab = ba = e \quad (7.5)$$

$$\begin{aligned} 7.4 \text{ and } 7.5 &\Rightarrow aa^{-1} = ab \\ &\Rightarrow a^{-1} = b \quad \text{using left cancellation law in } G. \quad \square \end{aligned}$$

The above theorem tells us that the identity element of a group and its subgroup are the same. Moreover, for any element a of H , its inverse in H is the same as its inverse in G .

If e is the identity element of a group G , then trivially $\{e\}$ and G are subgroups of G . They are called trivial subgroups of G . A subgroup of G other than $\{e\}$ and G is called a non-trivial (or proper) subgroup of G .

Example 7.1.

1. $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Q}, +)$.
2. $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{R}, +)$.
3. (\mathbb{Q}^+, \cdot) is a subgroup of (\mathbb{R}^+, \cdot) .
4. (\mathbb{Q}^+, \cdot) is not a subgroup of $(\mathbb{R}, +)$ though $\mathbb{Q}^+ \subseteq \mathbb{R}$. This is because $2, 3 \in \mathbb{Q}^+$, $2 \cdot 3 = 6$ but $2, 3 \in \mathbb{R}$ and $2 + 3 = 5$, Thus \cdot is not the binary operation on \mathbb{Q}^+ obtained from the binary operation $+$ on \mathbb{R} restricted to \mathbb{Q}^+ .
5. V_4 , the Klein 4-group is a subgroup of the dihedral group D_8 .
6. (\mathbb{N}, \cdot) is not a subgroup of (\mathbb{Q}^*, \cdot) , as (\mathbb{N}, \cdot) is not a group in its own right.
7. (\mathbb{Z}_5, \oplus_5) is a group but it is not a subgroup of $(\mathbb{Z}, +)$, whereas $(5\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Z}, +)$.

We now give some tests to determine whether a subset is a subgroup or not.

Theorem 7.2. (Three steps test) A subset H of a group G is a subgroup of G if and only if

- (i) the identity element of G belongs to H .
- (ii) $ab \in H \quad \forall a, b \in H$.
- (iii) for every $a \in H, a^{-1} \in H$.

Proof: If H is a subgroup of G , then H is a group with respect to the restricted binary operation on H . Hence conditions (i) to (iii) hold.

Conversely, let H be a subset of G such that the conditions hold. Then $e \in H \Rightarrow H$ is nonempty. Thus for H to be a subgroup the only axiom to be checked is the associativity axiom. Let $a, b, c \in H$. Since $H \subseteq G$, therefore $a, b, c \in G$. By associative law in G , $a(bc) = (ab)c$. Hence associative law holds in H , so that H is a group in its own right. Thus H is a subgroup of G . \square

Sometimes we need to check that a given subset of a group is not a subgroup. How do we go about this? In view of the above theorem, a subset H of a group G is not a subgroup if any one of the following is true:

- (i) The identity e of G does not belong to H .
- (ii) For some pair of elements a, b of H , $ab \notin H$.
- (iii) For some element $a \in H$, $a^{-1} \notin H$.

The use of these conditions is illustrated by the following examples.

- (i) $(\mathbb{N}, +)$ is not a subgroup of $(\mathbb{Z}, +)$ because the identity element 0 of \mathbb{Z} does not belong to \mathbb{N} .
- (ii) (\mathbb{Z}_6, \oplus_6) is a group, where $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$. Let $H = \{0, 1, 2\}$. Then $H \subseteq \mathbb{Z}_6$. $1, 2 \in H$ but $1 \oplus_6 2 = 3 \notin H$. Hence H is not a subgroup of G .
- (iii) (\mathbb{Z}^+, \cdot) is not a subgroup of (\mathbb{Q}^+, \cdot) as $2 \in \mathbb{Z}^+$ but $2^{-1} = \frac{1}{2} \notin \mathbb{Z}^+$.

The three steps test can be simplified to testing of only 2 conditions instead of 3. This is given in the following theorem.

Theorem 7.3. (Two steps test) Let G be a group and H a non empty subset of G . Then H is a subgroup of G if and only if

- (i) $ab \in H$, for all $a, b \in H$.
- (ii) For each $a \in H$, $a^{-1} \in H$.

Proof: If H is a subgroup of G , then conditions (i) and (ii) must hold by definition of subgroup.

Conversely suppose H is a non-empty subset of G such that the conditions hold. Since H is non empty, therefore there exists some $a \in H$. Then for $a \in H$, by (ii) $a^{-1} \in H$. Now $a, a^{-1} \in H$ so that by (i), $aa^{-1} \in H$, that is, $e \in H$. Thus the identity element is in H . Hence, by Theorem 7.2, H is subgroup of G . \square

This two step test can be further reduced to a one step test, Great, isn't it?

Theorem 7.4. (One step test) Let G be a group. A non-empty subset H of G is a subgroup of G if and only if $a, b \in H \Rightarrow ab^{-1} \in H$.

Proof: If H is a subgroup of G , then the condition holds by the definition of a subgroup.

Conversely, let the condition hold. Since H is non-empty, therefore, there exists an element a in H . By the given condition $aa^{-1} \in H$, that is, $e \in H$.

$e, a \in H$, so that $ea^{-1} \in H$, that is $a^{-1} \in H$. Let $a, b \in H$. Then $b^{-1} \in H$. By the given condition $a(b^{-1})^{-1} \in H$, that is, $ab \in H$. Thus, by Theorem 7.3, H is a subgroup of G . \square

The following theorem gives a condition for a finite subset of a group to be a subgroup.

Theorem 7.5. (*Finite subgroup test*) *Let G be a group and H a finite non empty subset of G . Then H is a subgroup of G if and only if H is closed under the operation of G .*

Proof: If H is a subgroup of G then clearly it is closed.

Conversely, let H be closed. To prove that H is a subgroup of G , it is sufficient to prove that $a^{-1} \in H$, whenever $a \in H$. If $a = e$, then $a^{-1} = e^{-1} = e = a$, so that $a^{-1} \in H$. If $a \neq e$, consider the sequence a, a^2, a^3, \dots . Closure property of H implies that all powers of a are in H , since H is finite not all of these elements can be distinct. Suppose $a^i = a^j$ for some i, j such that $i > j$. Then $a^{i-j} = e$; and since $a \neq e$, therefore $i - j > 1$. Thus $a^{i-j} = a^{i-j-1} \cdot a = e$, so that $a^{-1} = a^{i-j-1}$. Since $i - j - 1 \geq 1$, therefore, $a^{i-j-1} \in H$ so that $a^{-1} \in H$. Thus the proof is complete by Theorem 7.3. \square

Note: For additive groups in above tests we replace e by 0 , ab by $a + b$, a^{-1} by $-a$ and ab^{-1} by $a - b$.

Theorem 7.6. *The intersection of two subgroups of a group is a subgroup.*

Proof: Let G be a group and H_1, H_2 be two subgroups of G . Let $H = H_1 \cap H_2$. Since H_1, H_2 are subgroups, therefore $e \in H_1$ and $e \in H_2$, so that $e \in H_1 \cap H_2 = H$. Hence $e \in H$. This proves that H is non-empty. Let $a, b \in H$. Then $a, b \in H_1$ and $a, b \in H_2$. Since H_1 and H_2 are subgroups, therefore $ab^{-1} \in H_1$ and $ab^{-1} \in H_2$, so that $ab^{-1} \in H_1 \cap H_2 = H$ i.e. $ab^{-1} \in H$. Hence by the one step test, H is a subgroup of G . \square

Note that the above result does not hold for the union of two subgroups. This is shown by the following example.

Example 7.2. *The union of two subgroups need not be a subgroup. Consider the group $(\mathbb{Z}, +)$. Then $2\mathbb{Z}, 3\mathbb{Z}$ are subgroups of \mathbb{Z} . If $H = 2\mathbb{Z} \cup 3\mathbb{Z}$, then H is not a subgroup of \mathbb{Z} , for $2 \in 2\mathbb{Z} \subseteq H$ and $3 \in 3\mathbb{Z} \subseteq H$. Thus $2, 3 \in H$ but $2 + 3 = 5 \notin H$ as $5 \notin 2\mathbb{Z}$ and $5 \notin 3\mathbb{Z}$.*

We would like to know under what condition is the union of two subgroups a subgroup. This is answered in the next theorem.

Theorem 7.7. *The union of two subgroups is a subgroup if and only if one of them is contained in the other. That is, if H_1, H_2 are two subgroups of a group G , then $H_1 \cup H_2$ is a subgroup of G if and only if $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.*

Proof: Let H_1, H_2 be subgroups of a group G . Suppose $H_1 \subseteq H_2$, then $H_1 \cup H_2 = H_2$ which is a subgroup of G . Similarly if $H_2 \subseteq H_1$, then $H_1 \cup H_2 = H_1$ is a subgroup of G .

Conversely, let $H = H_1 \cup H_2$ be a subgroup of G , we shall prove that either $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$. Assume the contrary, that is, $H_1 \not\subseteq H_2$ and $H_2 \not\subseteq H_1$. Thus, there exists $h_1 \in H_1$, such that $h_1 \notin H_2$ and $h_2 \in H_2$, such that $h_2 \notin H_1$. Now $h_1 \in H_1 \subseteq H$ and $h_2 \in H_2 \subseteq H$, so that $h_1 h_2 \in H$ (as H is a subgroup). Since $H = H_1 \cup H_2$, therefore, $h_1 h_2 \in H_1 \cup H_2$, so that $h_1 h_2 \in H_1$, or $h_1 h_2 \in H_2$. Suppose $h_1 h_2 \in H_1$. Then $h_1^{-1}(h_1 h_2) \in H_1$ (as H_1 is a subgroup), that is $h_2 \in H_1$ which is a contradiction to our choice of h_2 .

Similarly $h_1h_2 \in H_2$ will also give a contradiction. Hence our assumption is wrong, so that $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$. \square

The following theorem characterizes all subgroups of $(\mathbb{Z}, +)$.

Theorem 7.8. *For every integer $n \geq 0$, $n\mathbb{Z}$ is a subgroup of \mathbb{Z} . Moreover every subgroup of \mathbb{Z} is of the form $m\mathbb{Z}$ for some integer $m \geq 0$.*

Proof: $n\mathbb{Z} = \{nz | z \in \mathbb{Z}\}$. If $n = 0$, then $n\mathbb{Z} = \{0\}$, which is a subgroup of \mathbb{Z} .

If $n > 0$, then $0 = n0 \in n\mathbb{Z}$, so that $n\mathbb{Z}$ is non-empty. Let $x, y \in n\mathbb{Z}$ then $x = nz_1$ and $y = nz_2$ for some $z_1, z_2 \in \mathbb{Z}$. So $x - y = nz_1 - nz_2 = n(z_1 - z_2) \in n\mathbb{Z}$ ($\because z_1 - z_2 \in \mathbb{Z}$). Thus by the one step test $n\mathbb{Z}$ is a subgroup of \mathbb{Z} .

Let H be any subgroup of \mathbb{Z} . If $H = \{0\}$ then $H = m\mathbb{Z}$ where $m = 0$. Let $H \neq \{0\}$. Then H contains some non-zero integer t , since H is a subgroup of \mathbb{Z} $\therefore -t \in H$. Out of $t, -t$ one of them must be positive. Thus H contains a positive integer. Let m be the smallest positive integer in H . Clearly $m\mathbb{Z} \subseteq H$. Let $h \in H$. By division algorithm there exists $q, r \in \mathbb{Z}$ such that $h = mq + r$, $0 \leq r < m$. Then $r = h - mq$, $r \in H$. Now $h, m \in H$ so that $h - mq \in H$, that is, $r \in H$. If $r \neq 0$ then $0 < r < m$ which contradicts our choice of m . Thus $r = 0$, so that $h = mq \in m\mathbb{Z}$, i.e. $H \subseteq m\mathbb{Z}$. Thus we get $H = m\mathbb{Z}$. \square

An important relation amongst the subgroups of \mathbb{Z} is the following:
If $m\mathbb{Z}$ and $n\mathbb{Z}$ are two subgroups of \mathbb{Z} then $m\mathbb{Z} \leq n\mathbb{Z}$ if and only if n divides m . This can be proved as follows:

$$\begin{aligned} m\mathbb{Z} &\leq n\mathbb{Z} \\ m\mathbb{Z} &\subseteq n\mathbb{Z} \\ \Leftrightarrow m \in n\mathbb{Z} &\Leftrightarrow m = nz \text{ for some } z \in \mathbb{Z} \\ \Leftrightarrow n &\text{ divides } m. \end{aligned}$$

Example 7.3. *Find all the subgroups of \mathbb{Z}*

(i) *containing $20\mathbb{Z}$.*

(ii) *contained in $20\mathbb{Z}$.*

$$\begin{aligned} \text{(i)} \quad 20\mathbb{Z} &\subseteq n\mathbb{Z} \\ \Leftrightarrow n &| 20 \\ \Leftrightarrow n &= 1, 2, 4, 5, 10, 20. \end{aligned}$$

Thus subgroups of \mathbb{Z} containing $20\mathbb{Z}$ are $\mathbb{Z}, 2\mathbb{Z}, 4\mathbb{Z}, 5\mathbb{Z}, 10\mathbb{Z}, 20\mathbb{Z}$.

(ii) If $n\mathbb{Z}$ is a subgroup contained in $20\mathbb{Z}$, then

$$\begin{aligned} n\mathbb{Z} &\leq 20\mathbb{Z} \\ \Leftrightarrow 20 &\text{ divides } n \\ \Leftrightarrow n &\text{ is a multiple of } 20 \\ \Leftrightarrow n &= 20k, k \in \mathbb{Z}, k \geq 0. \\ \Leftrightarrow n &= 0, 20, 40, 60, \dots \end{aligned}$$

Thus $(0), 20\mathbb{Z}, 40\mathbb{Z}, 60\mathbb{Z}, \dots$ are all subgroups of $20\mathbb{Z}$. Thus $20\mathbb{Z}$ has infinitely many subgroups.

7.2 Solved Problems

Problem 7.1. *If H is a subgroup of a group G and $x \in G$, then $xHx^{-1} = \{xhx^{-1} : h \in H\}$ is a subgroup of G .*

Solution: Clearly $e = xex^{-1} \in xHx^{-1}$ so that xHx^{-1} is non-empty. Let $a, b \in xHx^{-1}$. Then $a = xh_1x^{-1}, b = xh_2x^{-1}$ for some $h_1, h_2 \in H$. Now $ab = (xh_1x^{-1})(xh_2x^{-1}) = xh_1h_2x^{-1} \in xHx^{-1}$ ($\because h_1h_2 \in H$ as H is a subgroup). Thus $ab \in xHx^{-1}$. Also $a^{-1} = (xh_1x^{-1})^{-1} = (x^{-1})^{-1}(h_1)^{-1}x^{-1} = xh_1^{-1}x^{-1} \in xHx^{-1}$ ($\because h_1^{-1} \in H$). Hence $a \in H \Rightarrow a^{-1} \in H$. Thus we have proved that H is a subgroup of G .

Problem 7.2. If $H = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a \neq 0, a, b \in \mathbb{R} \right\}$ then prove that H is a subgroup of $GL(2, \mathbb{R})$.

Solution: Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in H$. Thus H is a non-empty subset of $GL(2, \mathbb{R})$.

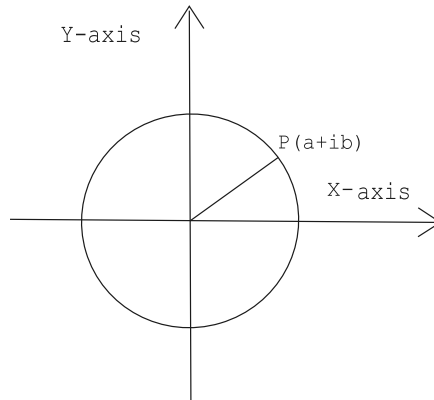
Let $A, B \in H$. Then $A = \begin{pmatrix} a_1 & b_1 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} a_2 & b_2 \\ 0 & 1 \end{pmatrix}$ for some $a_1, a_2, b_1, b_2 \in \mathbb{R}, a_1 \neq 0, a_2 \neq 0$. Now $AB = \begin{pmatrix} a_1a_2 & a_1b_2 + b_1 \\ 0 & 1 \end{pmatrix} \in H$ ($\because a_1a_2, a_1b_2 + b_1 \in \mathbb{R}, a_1a_2 \neq 0$). Thus H is closed.

If $A = \begin{pmatrix} a_1 & b_1 \\ 0 & 1 \end{pmatrix} \in H$,

then $A^{-1} = \begin{pmatrix} a_1^{-1} & -b_1a_1^{-1} \\ 0 & 1 \end{pmatrix}$ and $A^{-1} \in H$. Thus H is a subgroup of $GL(2, \mathbb{R})$.

Problem 7.3. Let $H = \{a + ib \in \mathbb{C} \mid a^2 + b^2 = 1\}$. Describe the elements of H geometrically. Is H a subgroup of \mathbb{C}^* under multiplication? Justify.

Solution: Let $z = a + ib \in H$. Then $|z| = \text{distance of } P(z) \text{ from origin} = \sqrt{a^2 + b^2} = 1$. Thus H represents all points on the circle of radius 1, centered at the origin.



\mathbb{C}^* , the set of non-zero complex numbers, is a multiplicative group with 1 as the identity element and $\frac{1}{z}$ as the multiplicative inverse of z . Thus H is the set of all complex numbers of modulus 1. We shall now prove that H is a subgroup of G . Clearly $1 + i0 \in H$, so that H is non-empty. Let $z_1, z_2 \in H$.

Then $|z_1| = 1, |z_2| = 1$. Now $|z_1 z_2| = |z_1||z_2| = 1$. Hence $z_1 z_2 \in H$. Also $|\frac{1}{z_1}| = \frac{1}{|z_1|} = 1$ so that $\frac{1}{z_1} \in H$. $\therefore z_1^{-1} = \frac{1}{z_1}$ so that $z_1^{-1} \in H$. Thus H is a subgroup of \mathbb{C}^* .

Problem 7.4. Let G be an Abelian group and let $H = \{x^2 | x \in G\}$. Then H is a subgroup of G .

Solution: *Step 1* Since $e^2 = e$, therefore $e \in H$. Hence H is a non-empty set.

Step 2 Let $a, b \in H$. Then $a = x^2$ and $b = y^2$ for some $x, y \in G$. Now $ab = x^2 y^2 = x(xy)y = x(yx)y$ (as G is an Abelian) $= (xy)(xy) = (xy)^2$, and $xy \in G$. Thus $ab \in H$.

Step 3 Let $a \in H$. Then $a = x^2$ for some $x \in G$. Now $a^{-1} = (x^2)^{-1} = (x^{-1})^2$. Since $x \in G$, therefore $x^{-1} \in G$, so that $(x^{-1})^2 \in H$, that is, $a^{-1} \in H$.

By the three step test, H is a subgroup of G .

Problem 7.5. Show that a group of order 6 cannot have a subgroup of order 4.

Solution: Let G be a group of order 6. Let, if possible, H be a subgroup of G of order 4. Let $H = \{e = h_1, h_2, h_3, h_4\}$, e being the identity element. Let $g \in G$ such that $g \notin H$ such an element g exists because $o(G) = 6$ and $o(H) = 4$. Consider the set $gH = \{gh_1, gh_2, gh_3, gh_4\}$. gH has at most 4 elements. We assert that all the elements of gH are distinct and different from elements of H . For if, $gh_i = gh_j$, $i \neq j$ ($1 \leq i, j \leq 4$) then $h_i = h_j$ by cancellation law in G which is a contradiction. Hence all the elements of gH are distinct so that gH has exactly 4 elements. Also $H \cap gH = \phi$, for if $h \in gH \cap H$, then $h \in H, h \in gH$. $\therefore h = h_i$ for some h_i and $h = gh_j$ for some $h_j \in H$. $\therefore gh_j = h_i$ so that $g = h_i h_j^{-1} \in H$ which contradicts the fact that $g \notin H$. Element of H and G account for 8 elements in a group of order 6 and this is not possible. Hence it is not possible for G to have a subgroup of order 4.

7.3 Exercise

- Find the flaw in the following argument: "Condition (i) of Theorem 7.2 is redundant since it can be derived from (ii) and (iii). For let $a \in H$. Then by (iii) $a^{-1} \in H$. By (ii) $aa^{-1} \in H$ i.e $e \in H$ which gives (i)."
- Determine which of the following subsets are subgroups of the group \mathbb{C} of complex numbers under addition.
 - \mathbb{R}
 - \mathbb{Q}^+
 - $7\mathbb{Z}$
 - $S = \{\pi^n : n \in \mathbb{Z}\}$
 - $\pi\mathbb{Q} = \{\pi x : x \in \mathbb{Q}\}$
 - The set $i\mathbb{R}$ of pure imaginary numbers together with zero.
- If $H = \{1, -1, i, -i\}$ prove that H is a subgroup of the group of Quaternions.

4. Consider $M_2(\mathbb{Z})$, the group of all 2×2 matrices over \mathbb{Z} , under addition.
- $$H_1 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z}) : a + b + c + d = 0 \right\}.$$
- $$H_2 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z}) : a + b + c + d = 1 \right\}.$$
- Are H_1, H_2 subgroups of $M_2(\mathbb{Z})$? Justify your answer.
5. Prove or disprove the following statements:
- every subgroup of a non-Abelian group is non-Abelian.
 - every subgroup of an infinite group is infinite.
6. If H is a subgroup of a group G and K is a subgroup of H , then prove that K is a subgroup of G . (Note that this shows that the relation “is a subgroup of” is transitive.)
7. Let H be a subgroup of \mathbb{R} under addition. Let $K = \{3^a : a \in H\}$. Prove that K is a subgroup of \mathbb{R}^* under multiplication.
8. Let \mathbb{R}^* be the group of non-zero real numbers under multiplication. If $H = \{x \in \mathbb{R}^* : x^2 \text{ is rational}\}$, prove that H is a subgroup of \mathbb{R}^* .
9. Let a, b, m be integer, $m > 1$ define $a \equiv b \pmod{m}$ if m divides $a - b$. Let $H = \{x \in \mathcal{U}(20) : x \equiv 1 \pmod{3}\}$. Is H a subgroup of $\mathcal{U}(20)$?
10. Let $G = GL(2, \mathbb{R})$. Test whether the subsets define below are subgroups of G .
- $H_1 = \{A \in G : |A| \text{ is an integral power of } 2\}$.
 - $H_2 = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} : a, b \text{ are nonzero integers} \right\}$
 - $H_3 = \{A \in G : |A| \text{ is rational}\}$
 - $H_4 = \{A \in G : |A| \text{ is an integer}\}$.
11. Let $H = \{a + ib : a, b \in \mathbb{R}, ab \geq 0\}$. Is H a subgroup of \mathbb{C} under addition? Justify.
12. Let G be the group of functions from \mathbb{R} to \mathbb{R}^* under multiplication. Set $H = \{f \in G : f(1) = 1\}$. Prove that H is a subgroup of G .
13. Let G be an Abelian group and let n be a fixed positive integer. Let $G^n = \{g^n : g \in G\}$. Prove that G^n is a subgroup of G .

7.4 Centralizers, Normalizers and Centre

Given a group how do we go about finding its subgroups? we now give some important families of subgroups of a group.

Centralizer of an Element

Although an element from a non-Abelian group need not necessarily commute with every element of the group, there are always some elements with which it will commute. For example, every element a commutes with all its powers. This observation prompts the following definition.

Definition 7.2. (Centralizer of an element):

Let G be a group and $a \in G$, be a fixed element of G . The centralizer of a in G , denoted by $C_G(a)$, is the set of all elements in G which commute with a . Symbolically, $C_G(a) = \{g \in G | ag = ga\}$. If the group G is understood then we simply write $C(a)$.

Example 7.4. Consider the dihedral group D (say).

$$\begin{aligned} C(R_0) &= D_8 \\ C(R_1) &= \{R_0, R_1, R_2, R_3\} \\ C(R_2) &= D_8 \\ C(R_3) &= \{R_0, R_1, R_2, R_3\} \\ C(H) &= \{R_0, H, R_2, V\} \\ C(V) &= \{R_0, H, R_2, V\} \\ C(D) &= \{R_0, D, R_2, D'\} \\ C(D') &= \{R_0, D, R_2, D'\}. \end{aligned}$$

Note that each of the centralizer is actually a subgroup of D_8 . Also two different elements may have the same centralizer.

Example 7.5. Consider the group $G = GL(2, \mathbb{R})$ under multiplication, and

$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = A \in GL(2, \mathbb{R})$. Then $C(A)$, the centralizer of A is

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G | c = b, d = a - b \right\}.$$

$$\begin{aligned} \text{If } B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ then } C(B) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G | d = a, c = b \right\} \\ &= \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} | a^2 - b^2 \neq 0; a, b \in \mathbb{R} \right\} \end{aligned}$$

Theorem 7.9. Let G be a group and $a \in G$. Then the centralizer $C(a)$ of a is a subgroup of G .

Proof: $C(a) = \{g \in G | ga = ag\}$. Since $e \in G$ is such that $ae = ea$, therefore $e \in C(a)$. Let $x, y \in C(a)$. Then $xa = ax$ and $ya = ay$. Now $ya = ay \Rightarrow y^{-1}yay^{-1} = y^{-1}ayy^{-1}$ (pre and post multiplication by y^{-1}), that is $ay^{-1} = y^{-1}a$.

Now $(xy^{-1})a = x(y^{-1}a) = x(ay^{-1}) = (xa)y^{-1} = (ax)y^{-1} = a(xy^{-1})$.

Thus $xy^{-1} \in C(a)$. Hence by the one step test $C(a)$ is a subgroup of G . \square

The next theorem tells us that the centralizer of an element and its inverse are the same set.

Theorem 7.10. If G be a group and $a \in G$, then $C(a) = C(a^{-1})$.

Proof: We know that $C(a) = \{g \in G | ag = ga\}$. Let $x \in C(a)$. Then $ax = xa$. Pre- and post-multiplying by a^{-1} , we get $a^{-1}axa^{-1} = a^{-1}xaa^{-1}$ that is $xa^{-1} = a^{-1}x$. Hence $x \in C(a^{-1})$, so that $C(a) \subseteq C(a^{-1})$. Let $y \in C(a^{-1})$ then $a^{-1}y = ya^{-1}$. Pre- and post-multiplying by a , we get $a(a^{-1}y)a = aya^{-1}a$, that is $ya = ay$. Hence $y \in C(a)$. Thus $C(a^{-1}) \subseteq C(a)$. Thus, we have $C(a^{-1}) = C(a)$. \square

Problem 7.6. If G is a group and $a, x \in g$, then $C(x^{-1}ax) = x^{-1}C(a)x$.

Solution: Let $y \in C(x^{-1}ax)$, then

$$\begin{aligned}
(x^{-1}ax)y &= y(x^{-1}ax) \\
\Rightarrow y^{-1}x^{-1}axy &= x^{-1}ax && \text{Pre-multiplying by } y^{-1} \\
\Rightarrow xy^{-1}x^{-1}ax &= axy^{-1} && \begin{array}{l} \text{Pre-multiplying by } x \\ \text{Post-multiplying by } y^{-1} \end{array} \\
\Rightarrow xy^{-1}x^{-1}a &= axy^{-1}x^{-1} && \text{Post-multiplying by } x^{-1} \\
\Rightarrow xy^{-1}x^{-1} &\in C(a) \\
\Rightarrow (xy^{-1}x^{-1})^{-1} &\in C(a) && \because C(a) \text{ is a subgroup} \\
\Rightarrow xyx^{-1} &\in C(a) \\
\Rightarrow y &\in x^{-1}C(a)x \\
\Rightarrow C(x^{-1}ax) &\subseteq x^{-1}C(a)x && (7.6)
\end{aligned}$$

Let $z \in x^{-1}C(a)x$, so that $z = x^{-1}cx$ for some $c \in C(a)$. Thus $ca = ac$. We shall prove that $z(x^{-1}ax) = (x^{-1}ax)z$

Now

$$\begin{aligned}
z(x^{-1}ax) &= (x^{-1}cx)(x^{-1}ax) \\
&= x^{-1}cax \\
&= x^{-1}acx && \because ac = ca \\
&= x^{-1}axx^{-1}cx \\
&= (x^{-1}ax)(x^{-1}cx) \\
&= (x^{-1}ax)z
\end{aligned}$$

Thus $z \in C(x^{-1}ax)$ so that

$$x^{-1}C(a)x \subseteq C(x^{-1}ax) \quad (7.7)$$

7.6 and 7.7 $\Rightarrow C(x^{-1}ax) = x^{-1}C(a)x$.

Centralizer of a Subset

The concept of the centralizer of an element can be extended to that of a subset.

Definition 7.3. If A is a subset of a group G , then by the centralizer of A we mean the set $\{x \in G \mid xa = ax \ \forall a \in A\}$. It is denoted by $C_G(A)$. when the group G is understood, we simply denote it by $C(A)$.

From the definition we have the following important result.

Theorem 7.11. If A is a subset of a group G , then $C(A) = \bigcap_{a \in A} C(a)$

Proof:

$$\begin{aligned}
y &\in C(A) \\
\Leftrightarrow ya &= ay \quad \forall a \in A \\
\Leftrightarrow y &\in C(a) \quad \forall a \in A \\
\Leftrightarrow y &\in \bigcap_{a \in A} C(a) \\
\Leftrightarrow C(A) &= \bigcap_{a \in A} C(a)
\end{aligned}$$

Hence the result. □

Example 7.6. Let us find $C(A)$, where $A = \{R_1, H\}$ and $A \subseteq D_8$, the dihedral group of order 8. We know that $C(A) = \cap_{a \in A} C(a)$

$$\begin{aligned} \therefore C(A) &= C(R_1) \cap C(H) \\ &= \{R_0, R_2\}. \end{aligned}$$

Theorem 7.12. If A is a subset of a group G , then $C(A)$ is a subgroup of G .

Proof: Clearly $ea = ae \quad \forall a \in A$

$\therefore e \in C(A)$ so that $C(A)$ is non-empty.

Let $a, b \in C(A)$. Then

$$ax = xa \quad \forall x \in A \quad (7.8)$$

$$\text{and } bx = xb \quad \forall x \in A \quad (7.9)$$

Now, for all $x \in A$

$$\begin{aligned} (ab)x &= a(bx) \quad \text{associativity} \\ &= a(xb) \quad \text{by (7.9)} \\ &= (ax)b \quad \text{associativity} \\ &= (xa)b \quad \text{by (7.8)} \\ &= x(ab) \quad \text{associativity} \end{aligned}$$

so that $ab \in C(A)$. Hence $C(A)$ is closed.

If $a \in C(A)$ then $ax = xa \quad \forall x \in A$

$$\Rightarrow a^{-1}axa^{-1} = a^{-1}xaa^{-1} \quad \text{Pre- and post-multiplying by } a^{-1}$$

$$\Rightarrow xa^{-1} = a^{-1}x$$

$$\Rightarrow a^{-1} \in C(A)$$

$$\therefore a \in C(A) \Rightarrow a^{-1} \in C(A), \text{ so that } C(A) \text{ is a subgroup of } G. \quad \square$$

Centre of a Group

In a group G , the identity element e of G occupies a very special position in the sense that it commutes with every element of the group. There may be other elements which commute with every element of the group. In case one such element a exists, then all powers of a will also commute with every element of G . For example all scalar matrices commute with every matrix. Thus if we consider the group $G = GL(2, \mathbb{R})$ under multiplication, then every non-zero scalar matrix commutes with every element of G . This motivates the following definition.

Definition 7.4. (Centre of a group): The centre of a group G is the set of elements of G which commute with every element of G , that is, the set $\{a \in G \mid ax = xa \quad \forall x \in G\}$. It is denoted by $Z(G)$.

From the definition it follows that

(i) In term of centralizer, $Z(G) = C(G)$

(ii) $Z(G) \subseteq C(a) \quad \forall a \in G$

$$(iii) Z(G) = \bigcap_{a \in G} C(a)$$

Example 7.7. Consider the group $G = GL(2, \mathbb{R})$. Since a scalar matrix commutes with every element of G , therefore every scalar matrix lies in the centre of G . Thus $\left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \mid a \in \mathbb{R} \right\} \subseteq Z(G)$.

Example 7.8. Consider the group $G = (\mathbb{Z}, +)$. Since G is Abelian, therefore every element of G commutes with all other elements of G . Hence $Z(G) = G$.

Example 7.9. Let us find the centre of D_8 , the dihedral group of order 8.

$$\begin{aligned} Z(D_8) &= \bigcap_{a \in D_8} C(a) \\ &= \{R_0, R_2\} \end{aligned}$$

$$\therefore Z(D_8) = \{R_0, R_2\}.$$

Theorem 7.13. The centre of a group G is a subgroup of G .

Proof: Since $Z(G) = C(G)$ and $C(G)$ is a subgroup, therefore $Z(G)$ is a subgroup. \square

Remark 7.1.

1. The centre of a group is an Abelian subgroup. If G is any group, then $Z(G) = \{g \in G : xg = gx \ \forall \ x \in G\}$. Let $g_1, g_2 \in Z(G)$. Since g_1 commutes with every element of G , in particular it commutes with g_2 , so that $g_1g_2 = g_2g_1$. Hence $Z(G)$ is Abelian.

2. The centre of a group G is $G \Leftrightarrow G$ is Abelian.
Now, G is Abelian

$$\begin{aligned} \Leftrightarrow ab &= ba \quad \forall \ a, b \in G \\ \Leftrightarrow b &\in C_G(a) \quad \forall \ a, b \in G \\ \Leftrightarrow b &\in \bigcap_{a \in G} C_G(a) \quad \forall \ b \in G \\ \Leftrightarrow G &= Z(G) \end{aligned}$$

Problem 7.7. Find the centre $Z(G)$ for $G = D_6$ the dihedral group of order 6.

Solution:

$$D_6 = \{R_0, R_1, R_2, M_1, M_2, M_3\}$$

$$Z(D_6) = \{x \in D_6 : xd = dx \text{ for all } d \in D_6\}$$

Now $M_3M_2 = R_2$

$$M_2M_3 = R_1$$

Thus

$$M_2M_3 \neq M_3M_2$$

so that $M_2, M_3 \notin Z(D_6)$

$$M_1R_2 = M_3$$

$$R_2M_1 = M_2$$

so that

$$\begin{aligned}M_1, R_2 &\notin Z(D_6) \\ R_1 M_3 &= M_2 \\ M_3 R_1 &= M_1\end{aligned}$$

so that

$$M_3, R_1 \notin Z(D_6)$$

Hence $M_1, M_2, M_3, R_1, R_2 \notin Z(D_6)$. Also $R_0 \in Z(D_6)$, being the identity element. Hence $Z(D_6) = \{R_0\}$.

Problem 7.8. Prove that centre of $GL(2, \mathbb{R})$ is the set of all scalar matrices.

Solution: Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Z(G)$. Then A commutes with every member of G . In particular it must commute with $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $Y = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, where $X, Y \in G$.

$$\begin{aligned}AX &= XA \\ \Rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ \Rightarrow \begin{pmatrix} b & a \\ d & c \end{pmatrix} &= \begin{pmatrix} c & d \\ a & b \end{pmatrix} \\ \Rightarrow b = c \text{ and } a = d\end{aligned}$$

Also $AY = YA$

$$\Rightarrow b = c = 0$$

Thus $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, so that $Z(G) \subseteq \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \mid a \in \mathbb{R} \right\}$.

Also if $a \in \mathbb{R}$, $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \in Z(G)$.

$$\text{Thus } Z(G) = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \mid a \in \mathbb{R} \right\}.$$

Hence the centre of $GL(2, \mathbb{R})$ is the set of all scalar matrices.

Normalizer of a subset

Let G be a group and A a non-empty subset of G . For any $x \in G$,

$$xA = \{xa : a \in A\}; \quad Ax = \{ax : a \in A\}.$$

$$\text{If } xa = ax \quad \forall a \in A \tag{7.10}$$

$$\text{then obviously } xA = Ax \tag{7.11}$$

But the sets xA and Ax may be equal without the condition (7.10) being satisfied. This leads us to the following definition.

Definition 7.5. Let A be a subset of group G . For any $x \in G$, let $xAx^{-1} = \{xax^{-1} | a \in A\}$. The set $\{x \in G | xAx^{-1} = A\}$ is called the normalizer of A in G . It is denoted by $N_G(A)$ or simply $N(A)$, when G is understood.

Remark 7.2.
$$xAx^{-1} = A$$

$$\Leftrightarrow xA = Ax$$

Thus the normalizer of a set A is that of all elements which commute with A .

Theorem 7.14. If H is a subgroup of a group G , then $N(H)$ is a subgroup of G .

Proof: Since $eHe^{-1} = \{ehe^{-1} : h \in H\} = \{h : h \in H\} = H$, therefore $e \in N(H)$. Hence $N(H)$ is non-empty. Let $x, y \in N(H)$. Then $xHx^{-1} = H$ and $yHy^{-1} = H$. Now $(xy)H(xy)^{-1} = xyHy^{-1}x^{-1} = x(yHy^{-1})x^{-1} = xHx^{-1} = H$, so that $xy \in N(H)$. Let $x \in N(H)$. Then $xHx^{-1} = H$. Pre- and post-multiplying by x^{-1} and x respectively, we get $x^{-1}xHx^{-1}x = x^{-1}Hx$, that is, $eHe = x^{-1}H(x^{-1})^{-1}$. Hence $x^{-1}H(x^{-1})^{-1} = H$, so that $x^{-1} \in N(H)$. Thus, $N(H)$ is a subgroup of G . \square

7.5 Exercise

1. If a is an element of a group G , prove that (i) $a^{-1} \in C(a)$ (ii) $a^n \in C(a)$ for all $n \in \mathbb{Z}$.
2. For any element a of a group G , prove that $x \in C(a) \Rightarrow x \in C(a^n)$ for all $n \in \mathbb{Z}$.
3. Find $C_G(A)$ where G is the group of quaternions under multiplication and $A = \{i\}$. Also find the centre of G .
4. If A and B are subsets of a group G , does $C(A) = C(B)$ necessarily imply $A = B$?
5. A and B are subsets of a group G . Prove that (i) $C(A \cup B) = C(A) \cup C(B)$ (ii) $C(A) \cup C(B) \subseteq C(A \cap B)$.
6. Let $G = GL(2, \mathbb{R})$. If $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, find $C(A)$ and $C(B)$.
7. Find the centre of D_8 , the dihedral group of order 8.
8. For any group G , prove that $Z(G) = \bigcap_{a \in G} C(a)$.
9. Let G be the group of all 2×2 diagonal matrices under multiplication. Find the centre of G .
10. If A is a subset of a group G , prove that the centralizer of A is a subset of the normalizer of A .
11. If A is a subset of a group G , prove that (i) the centre of G is a subset of the centralizer of A . (ii) the centre of G is a subset of the normalizer of A .
12. For any subset A of a group G , obtain a relationship between centre of G , centralizer of A and the normalizer of A .

7.6 Order of an Element

Consider the group $\mathcal{U}(20)$ under multiplication $\mathcal{U}(20) = \{1, 3, 7, 9, 11, 13, 17, 19\}$. Then 1 is the identity element. Let us compute the powers of elements of $\mathcal{U}(20)$.

Clearly

$$\begin{aligned} 1^1 &= 1 \\ 3^2 &= 9, \quad 3^3 = 7, \quad 3^4 = 1, \quad 3^5 = 3, \quad 3^6 = 9 \quad \text{etc...} \\ 7^2 &= 9, \quad 7^3 = 3, \quad 7^4 = 1; \\ 9^2 &= 1; \\ 11^2 &= 1; \\ 13^2 &= 9, \quad 13^3 = 17, \quad 13^4 = 1; \\ 17^2 &= 9, \quad 17^3 = 13, \quad 17^4 = 1; \\ 19^2 &= 1. \end{aligned}$$

Thus we observe that some power of each element becomes the identity element. Each of the elements 3, 7, 13 and 17 have 4 distinct powers whereas each of the elements 9, 11 and 19 have 2 distinct powers. The identity element 1 has only one distinct power. The number of distinct powers of an element is of great importance in the study of groups and it motivates us to define the following:

Definition 7.6. (Order of an element): Let G be a group and $g \in G$. If there exists a positive integer n such that $g^n = e$, then g is said to be of finite order. If no such integer exists, then g is said to be of infinite order. If g is of finite order, then the least positive integer n such that $g^n = e$ is called the order of g .

If G is an additive group and $g \in G$, we replace g^n by ng and e by o to find the order of g .

Notation: The order of an element g is denoted by $o(g)$ or $|g|$.

To find the order of an element g of a group, we compute the sequence of powers of g , namely g, g^2, g^3, \dots until we reach the identity element for the first time. Suppose $g^n = e$ for the first time. Then n is the order of g . If the identity never appears in the sequence, then g has infinite order. Note that the order of the identity element is always 1. In fact identity is the only element of order 1.

Example 7.10.

1. As explained above, in $\mathcal{U}(20)$ $o(3) = o(7) = o(13) = o(17) = 4$. $o(9) = o(11) = o(19) = 2$.
2. Consider the group $(\mathbb{Z}, +)$. If $0 \neq a \in \mathbb{Z}$, then $na \neq 0$ for every positive integer n . Hence the order of a is infinite.
3. Consider the group $GL(2, \mathbb{R})$. If $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in GL(2, \mathbb{R})$, then $A^2 = I$ so that $o(A) = 2$. If $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ then $B \in GL(2, \mathbb{R})$ and 4 is the smallest positive integer such that $B^4 = I$. Hence $o(B) = 4$.

Theorem 7.15. Let G be a group and $a \in G$. Then

- (i) a and a^{-1} have the same order.
- (ii) a and $x^{-1}ax$ have the same order for all $x \in G$.

Proof: Let G be a group and $a \in G$. Then

$$(a^{-1})^n = (a^n)^{-1} \quad (7.12)$$

(i) Two cases arise

Case 1. $o(a)$ is finite and $o(a) = n$. Then n is the smallest positive integer such that

$$a^n = e \quad (7.13)$$

If m is any positive integer

$$\begin{aligned} a^m &= e \\ \Leftrightarrow (a^m)^{-1} &= e^{-1} \\ \Leftrightarrow (a^{-1})^m &= e \\ \Leftrightarrow o(a) &= o(a^{-1}) \quad \text{using (7.13)} \end{aligned} \quad (7.14)$$

Case 2. $o(a)$ is infinite. Then $a^m \neq e$ for any positive integer m . Using (7.14) we get $(a^{-1})^m \neq e$ for every positive integer m , so that $o(a^{-1})$ is infinite.

(ii) Let x be any element of G and let $b = x^{-1}ax$. Then $b^m = x^{-1}a^m x$ for every positive integer m (see problem 6.1). Thus $b^m = e \Leftrightarrow a^m = e$ so that $o(b) = o(a)$. Hence $o(x^{-1}ax) = o(a)$ for all $x \in G$. \square

The following theorem gives a criterion for two powers of an element of a group to be equal in terms of its order.

Theorem 7.16. *If G is a group and $a \in G$ of order n , then*

(i) $a^k = e \Leftrightarrow n$ divides k .

(ii) $a^i = a^j \Leftrightarrow n$ divides $i - j$.

Proof: $o(a) = n \Leftrightarrow n$ is the smallest positive integer such that

$$a^n = e \quad (7.15)$$

(i) Applying division algorithm to n and k , we can find integers q and r such that $k = nq + r$, $0 \leq r < n$.
Now $a^k = a^{nq+r} = (a^n)^q a^r \Rightarrow a^k = a^r$ using (7.15).

$$\begin{aligned} \text{Thus } a^k = e &\Leftrightarrow a^r = e \\ \Leftrightarrow r = 0 &\text{ (since } o(a) = n \text{ and } 0 < r < n \Rightarrow a^r \neq e) \end{aligned}$$

Hence $k = nq$. Thus n divides k .

(ii) $a^i = a^j \Leftrightarrow a^{i-j} = e \Leftrightarrow n$ divides $i - j$ using (i). \square

The next theorem gives an upper bound on the order of an element of a finite group.

Theorem 7.17. *In a finite group G each element is of finite order. In fact, the order of an element is at most $o(G)$.*

Proof: Let G be a finite group of order n , and let $a \in G$. Then a, a^2, a^3, a^4, \dots are all elements of G . Since G is finite, therefore all powers of a cannot be distinct, so that for some integers r and s , $a^r = a^s$. Without any loss of generality we may assume $r > s$. Thus $a^{r-s} = e$. Let $k = r - s$. Then $k \in \mathbb{Z}^+$ such that $a^k = e$, so that a has finite order $\leq k$. If $o(a) = m$, then $a, a^2, \dots, a^{m-1}, a^m (= e)$ are distinct elements of G and hence $m \leq o(G)$. This gives that the order of any element of G is at most $o(G)$. \square

Example 7.11. Consider the group $(\mathbb{Z}_7^*, \odot_7)$. Let us find the order of each element of \mathbb{Z}_7^* .

$$\begin{aligned} 2^2 = 4, \quad 2^3 = 1 &\Rightarrow o(2) = 3. \\ 3^2 = 2, \quad 3^3 = 6, \quad 3^4 = 4, \quad 3^5 = 5, \quad 3^6 = 1 &\Rightarrow o(3) = 6. \\ 4^2 = 2, \quad 4^3 = 1 &\Rightarrow o(4) = 3. \\ 5^2 = 4, \quad 5^3 = 6, \quad 5^4 = 2, \quad 5^5 = 3, \quad 5^6 = 1 &\Rightarrow o(5) = 6. \\ 6^2 = 1 &\Rightarrow o(6) = 2. \end{aligned}$$

Thus $o(6) = 2$, $o(2) = o(4) = 3$, $o(3) = o(5) = 6$ and the identity 1 is of order 1.

Problem 7.9. If in a group G , $x \in G$ such that $o(x) = 6$, find $o(x^2), o(x^3), o(x^4)$ and $o(x^5)$.

Solution: Since $o(x) = 6$, therefore 6 is the smallest positive integer such that $x^6 = e$. Let us find the powers of x^2 . Now $(x^2)^2 = x^4 \neq e$, $(x^2)^3 = x^6 = e$. Hence $o(x^2) = 3$.

Similarly $(x^3)^2 = x^6 = e$ so that $o(x^3) = 2$; $(x^4)^2 = x^8 = x^2 \neq e$, $(x^4)^3 = x^{12} = e \Rightarrow o(x^4) = 3$; $(x^5)^2 = x^{10} \neq e$, $(x^5)^3 = x^{15} = x^3 \neq e$, $(x^5)^4 = x^{20} = x^2 \neq e$, $(x^5)^5 = x^{25} = x \neq e$, $(x^5)^6 = x^{30} = e \Rightarrow o(x^5) = 6$.

Thus, we have obtained $o(x^2) = 3, o(x^3) = 2, o(x^4) = 3$ and $o(x^5) = 6$.

Practically we do not need to find all the powers of x^k in order to find the order of x^k . Knowing the order of an element, can we find the order of any of its power? This is answered in the following theorem.

Theorem 7.18. Let G be a group.

(a) If a is an element of G of finite order n , then

- (i) $o(a^m) = \frac{n}{m}$, if m divides n .
- (ii) $o(a^m) = n$, if m and n are coprime.
- (iii) For any integer m such that $0 < m < n$,

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

(b) If a is an element of infinite order then for all $m \in \mathbb{Z} \setminus \{0\}$ the element a^m is also of infinite order.

Proof:

(a) $o(a) = n \Rightarrow n$ is the smallest positive integer such that

$$a^n = e \tag{7.16}$$

- (i) If m divides n , then $n = mk$, for some $k \in \mathbb{Z}^+$. Thus $k = \frac{n}{m}$. Now $(a^m)^k = a^{mk} = a^n = e$. If $o(a^m) = p$, then $(a^m)^p = e$, so that $a^{mp} = e$. By Theorem 7.16(i), n divides mp , i.e., mk divides mp so that k divides p . Thus k is the smallest positive integer such that $(a^m)^k = e$. Hence $o(a^m) = k = \frac{n}{m}$.

- (ii) $(a^m)^n = a^{mn} = (a^n)^m = e^m = e$. Suppose that t is any positive integer such that $(a^m)^t = e$. Then $a^{mt} = e$. Using Theorem 7.16(i), n divides mt . Now, n divides mt and m, n are coprime $\Rightarrow n$ divides t . Thus n is the smallest positive integer such that $(a^m)^n = e$, so that $o(a^m) = n$.
- (iii) Let $l = \text{lcm}(m, n)$. Then m and n both divide l so that $l = mu, l = nv$ for some integer u and v . Now $(a^m)^u = a^{mu} = a^l = a^{nv} = (a^n)^v = (e)^v = e$. Hence $(a^m)^u = e$. Let k be a positive integer such that $(a^m)^k = e$. Then $a^{mk} = e$. By Theorem 7.16(i) n divides mk , i.e. mk is a multiple of n consequently mk is a multiple of l . Thus l divides mk i.e. mu divides mk . i.e. u divides k . This gives that u is the least positive integer such that $(a^m)^u = e$, so that $o(a^m) = u = \frac{l}{m} = \frac{\text{lcm}(m, n)}{m}$. Since $mn = \text{lcm}(m, n)\text{gcd}(m, n)$.
Therefore $\frac{\text{lcm}(m, n)}{m} = \frac{n}{\text{gcd}(m, n)}$, so that $o(a^m) = \frac{\text{lcm}(m, n)}{m} = \frac{n}{\text{gcd}(m, n)}$.

(b) If a is of infinite order then

$$a^k \neq e \quad (7.17)$$

for any positive integer k . Let, if possible, a^m be of finite order, say t . Then $(a^m)^t = e \Rightarrow a^{mt} = e$ which is a contradiction to (7.17). Hence our assumption is wrong so that a^m is also of infinite order for every $m \in \mathbb{Z}^*$.

□

Remark 7.3. In the above theorem (i) and (ii) are special case of (iii) when m divides n then $\text{gcd}(m, n) = m$ so (iii) \Rightarrow (i) when m, n are coprime $\text{gcd}(m, n) = 1$ so (iii) \Rightarrow (ii).

The above theorem is used often, so we give the equivalent version when the group operation is addition.

Theorem 7.19. Let $(G, +)$ be a group.

(a) If $a \in G$ is of finite order n , then

- (i) $ma = 0$ if and only if n divides m .
(ii) For $l, m \in \mathbb{Z}$ $la = ma$ if and only if n divides $l - m$.
(iii) $o(ma) = \frac{n}{m}$, if m divides n .
(iv) $o(ma) = n$, if m and n are coprime.
(v) For any integer m such that $0 < m < n$,

$$o(ma) = \frac{\text{lcm}(m, n)}{m} = \frac{n}{\text{gcd}(m, n)}.$$

(b) If a is an element of infinite order, then for all $m \in \mathbb{Z} \setminus \{0\}$, ma is also of infinite order.

The use of the Theorem (7.18) is illustrated in the following examples.

Example 7.12. If G is a group and $a \in G$ such that $a^{12} = e$, what can you say about the order of a ?

By the above theorem $o(a)$ divides 12. Thus the possible order of a is 1, 2, 3, 4, 6 or 12.

Example 7.13. Let G be a group and $a \in G$ such that $o(a) = 12$. Let us find the orders of $a^3, a^5, a^6, a^7, a^8, a^9, a^{10}$ and a^{11} .

Let $n = o(a) \therefore n = 12$ and $a^{12} = e$. Let $m = 3$. Since m divides n
 $\therefore o(a^m) = \frac{n}{m} = \frac{12}{3} = 4$. i.e, $o(a^3) = 4$

Also 6 divides $n \Rightarrow o(a^6) = \frac{12}{6} = 2$.

Since 5 is coprime to $12(=n)$,

\therefore By (ii) of the above theorem $o(a^5) = n = 12$. Similarly $o(a^7) = 12$,
 $o(a^{11}) = 12$

Now $o(a^8) = \frac{12}{\gcd(8,12)}$ by (iii) of the theorem thus $o(a^8) = \frac{12}{4} = 3$

$\therefore o(a^8) = 3$.

Similarly $o(a^9) = 4$, $o(a^{10}) = 6$. Thus $aa^{11} = a^{11}a = e \Rightarrow a^{11} = a^{-1}$.
 But $o(a^{-1}) = o(a)$ so that $o(a^{-1}) = 12$ i.e $o(a^{11}) = 12$.

Example 7.14. In $(\mathbb{Z}_{30}, \oplus_{30})$, $30.1=0$ and $m.1 \neq 0$ for $1 \leq m < 30$.

So $o(1) = 30$. We find the order of a given element using Theorem 7.18.

$o(ma) = \frac{o(a)}{\gcd(m, o(a))} = o(-ma)$. So for any $m \in \mathbb{Z}_{30}$

$o(m) = \frac{o(1)}{\gcd(m, o(1))} = \frac{30}{\gcd(m, 30)} = o(-m)$.

For instance $o(2) = \frac{30}{\gcd(2, 30)} = \frac{30}{2} = 15 = o(28)$.

Other elements of order 15 are $k.2$ where $(k, o(2))=1$. Thus $k = 1, 2, 4, 7, 8, 11, 13, 14$. Hence order of $2, 4, 8, 14, 16, 22, 26, 28$ is also 15.

Using Theorem 7.18 we can summarize the order of the elements of \mathbb{Z}_{30} as below:

Element	Order
0	1
1, 7, 11, 13 29, 23, 19, 17	30
2, 4, 8, 14 28, 26, 22, 16	15
3, 9, 27, 21	10
5, 25	6
6, 12, 24, 18	5
10, 20	3
15	2

7.7 Solved Problems

Problem 7.10. If G is a group and $a \in G$ such that $a^{24} = e, a^{12} \neq e, a^8 \neq e$ find the $o(a)$.

Solution: Let $o(a) = n$. Since $a^{24} = e$, therefore n divides 24, by Theorem 7.16. Hence $n = 1, 2, 3, 4, 6, 8, 12, 24$, again by Theorem 7.16 Since $a^8 \neq e$, therefore $n \neq$ divisors of 8 i.e $n \neq 1, 2, 4, 8$. Similarly $a^{12} \neq e$ so that $n \neq 1, 2, 3, 4, 6, 12$. Thus the only possible value of n is 24 so that $o(a) = 24$.

Problem 7.11. Prove that an Abelian group with two elements of order 2 must have a subgroup of order 4.

Solution: Let G be an Abelian group with two elements a, b of order 2. Then $a^2 = e = b^2$ and $a \neq e$ and $b \neq e$. Let H be a subgroup of G containing both

a and b . Then H must contain ab and e . We assert that $S = \{e, a, b, ab\}$ is a subgroup of G of order 4. Clearly $a \neq e, b \neq e$. We prove that $ab \neq e$. For $ab = e \Rightarrow a^2b = ae$ (pre-multiplying by a) $\Rightarrow eb = a \Rightarrow b = a$, a contradiction. Similarly $ab \neq a$ and $ab \neq b$, by using cancellation law. Thus S has 4 distinct element. Since G is Abelian, so $ab = ba$. Multiplication table of S is

	e	a	b	ab
e	e	a	b	ab
a	a	e	ab	b
b	b	ab	e	a
ab	ab	b	a	e

Thus S is closed and every element of S has its inverse in S . Hence S is the required subgroup of G of order 4.

Problem 7.12. Let G be a group and $x, y \in G$ such that $x \neq e$, $o(y) = 2$, and $xyx^{-1} = x^2$. Find $o(x)$.

Solution: Since $o(y) = 2$, $\therefore y^2 = e$ so that $y = y^{-1}$. Also $xyx^{-1} = x^2 \Rightarrow yx = x^2y$. Now, $x^3 = x^2x = x^2ex = x^2y^2x = (x^2y)yx = (yx)yx = (yx)(x^2y) = yx^3y = yx^3y^{-1} = (yxy^{-1})^3 = (x^2)^3 = x^6$. Thus $x^3 = x^6$, so that $x^3 = e$. Hence $o(x) = 1$ or 3. Since $x \neq e$, therefore $o(x) \neq 1$, so that $o(x) = 3$.

Problem 7.13. Let G be an Abelian group and let $T = \{a \in G | o(a) \text{ is finite}\}$. Then T is a subgroup of G .

Solution: Clearly T is nonempty, as $e \in T$. Let $a, b \in T$. Let $o(a) = m$ and $o(b) = n$. Then $a^m = e$ and $b^n = e$. Now $(ab)^{mn} = a^{mn}b^{mn}$ (as G is Abelian) $= (a^m)^n(b^n)^m = e$. Thus ab is of finite order so that $ab \in T$. If $a \in T$ and $o(a) = m$ then since $o(a) = o(a^{-1})$, therefore $o(a^{-1}) = m$ so that $a^{-1} \in T$. Hence T is a subgroup of G .

The above subgroup T is a well known subgroup of G , called the Torsion subgroup of G .

Remark 7.4. The above subset T fails to be a subgroup when G is non Abelian. This is seen by the following example.

Example 7.15. Consider the group $G = SL(2, \mathbb{R})$. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$. Then $A, B \in G$ and $o(A) = 4, o(B) = 3$. Then $A, B \in T$. Also $AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $(AB)^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ so that $(AB)^n \neq I$ for any n . Hence AB is not of finite order so that $AB \notin T$. Thus T is not a subgroup.

Problem 7.14. Find a group that contains elements a and b such that $o(a) = o(b) = 2$ and

(i) $o(ab) = 3$

(ii) $o(ab) = 4$

(iii) $o(ab) = 5$

Can you see any relationship between $o(a)$, $o(b)$, and $o(ab)$?

Solution:

- (i) In D_6 let $a = M_2, b = M_1$. Then $ab = R_1$. Also $M_1^2 = R_0 = M_2^2$, so that $o(a) = o(b) = 2$. $R_1^2 = R_2, R_1^3 = R_0$, so that $o(R_1) = 3$. That is, $o(ab) = 3$.
- (ii) In D_8 let $a = H, b = D$. Then $ab = R_1$ $o(a) = o(b) = 2$. $o(ab) = 4$.
- (iii) In D_{10} , consider $a = H_3, b = H_5$ (where H_i denotes the reflection about the line of symmetry through the i^{th} vertex of the regular pentagon). Then $ab = H_3H_5 = R_1$. Now $o(a) = 2, o(b) = 2, o(ab) = 5$.

In general there is no relationship between the orders of a, b and ab . In fact in D_{2n} we can find elements a and b such that $o(a) = o(b) = 2$ and $o(ab) = n$.

Remark 7.5. *The above situation arises only in non-Abelian groups. Can you predict what happens in an Abelian group?*

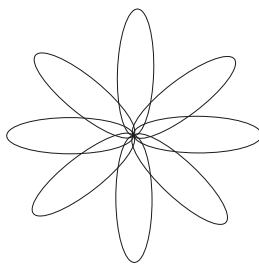
Problem 7.15. *If G is a group and $a \in G$ such that $o(a) = 5$, then prove that $C(a) = C(a^3)$. Find an element a from some group such that $o(a) = 6$ and $C(a) \neq C(a^3)$.*

Solution: Since $o(a) = 5$, therefore $a^5 = e$, so that $a^3 = a^{-2}$ clearly $C(a) \subseteq C(a^3)$ (see above remark). Let $x \in C(a^3)$. Then $xa^3 = a^3x$, so that $xa^{-2} = a^{-2}x$, that is

$$a^2x = xa^2 \quad (7.18)$$

Now $xa^3 = a^3x = a(a^2x) = a(xa^2)$ using (7.18). Thus $xa^3 = axa^2$, so that $xa = ax$, i.e $x \in C(a)$. Hence $C(a^3) \subseteq C(a)$, so that $C(a) = C(a^3)$.

Consider the group $G = D_{12}$. $o(R_1) = 6$. Let σ be the reflection about the perpendicular bisector of the side joining the vertices 1 to 6 and the opposite side. Then $R_1^3 = R_3$ $\sigma R_3 = R_3\sigma$ and $\sigma R_1 \neq R_1\sigma$. So $\sigma \in C(R_3)$ whereas $\sigma \notin C(R_1)$. Hence $C(R_1) \neq C(R_1^3)$.



7.8 Exercise

1. Prove that the identity is the only element of order 1.
2. If a, b are two elements of a group G , prove that $o(ab) = o(ba)$.

3. Without actually computing the orders, explain why the two elements in each of the following pairs of elements from $(\mathbb{Z}_{30}, \oplus_{30})$, must have the same order: $\{3, 27\}$, $\{12, 18\}$. Do the same for the following pairs of elements from $U(15)$: $\{2, 8\}$ and $\{7, 13\}$.
4. For each group in the following list, find the order of the group and the order of each elements in the group. In each case how are the orders of each of the elements related to the order of the group?
- (i) $(\mathbb{Z}_{12}, \oplus_{12})$
- (ii) $(U(20), \odot_{20})$
- (iii) $\mathbb{Z}_3^* \times \mathbb{Z}_5^*$ with respect to componentwise multiplication, where $\mathbb{Z}_3^* = \{1, 2\}$, $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$.
5. Let G be a group and $x \in G$. If $x^2 \neq e$ and $x^6 = e$, prove that $x^4 \neq e$ and $x^5 \neq e$. What can you say about the order of x ?
6. If G is a group and $x \in G$ such that $o(x) = 9$, find $o(x^k)$ for $k = 2, 3, \dots, 8$.
7. a is an element of the group G .
- (i) If $o(a^5) = 12$, what are the possibilities for $o(a)$?
- (ii) If $o(a^4) = 12$, what are the possibilities for $o(a)$?
8. For any positive integer n and any angle θ , show that in the group $SL(2, \mathbb{R})$

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}^n = \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}.$$

Use this formula to find the order of $A = \begin{pmatrix} \cos 60^\circ & -\sin 60^\circ \\ \sin 60^\circ & \cos 60^\circ \end{pmatrix}$ and

$$B = \begin{pmatrix} \cos(\sqrt{2}^\circ) & -\sin(\sqrt{2}^\circ) \\ \sin(\sqrt{2}^\circ) & \cos(\sqrt{2}^\circ) \end{pmatrix}.$$

9. Consider the group $SL(2, \mathbb{R})$ and $A, B \in SL(2, \mathbb{R})$, where $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$. Find $o(A)$, $o(B)$ and $o(AB)$. Does this answer surprise you? Justify.
10. Consider $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = A \in SL(2, \mathbb{R})$. what is the order of A ? If we view A as a member of $SL(2, \mathbb{Z}_p)$ (where p is prime), what is the order of A ?
11. If G is a group of finite order, prove that there exists a fixed positive integer n such that $a^n = e$ for all $a \in G$.

7.9 Cyclic Subgroups

Consider the group $(\mathbb{Z}_{12}, \oplus_{12})$. If $S = \{0, 3\}$ then S is not a subgroup of \mathbb{Z}_{12} , because $3 \oplus_{12} 3 = 6 \notin S$. Let us see how big a subgroup H of \mathbb{Z}_{12} would it have to be if it contained 3. H should contain the identity 0 and the inverse of 3, which is 9. Also it must contain $3 \oplus_{12} 3 = 6$. Thus, in addition to 3, 0, 6 and 9 should all belong to H . Given below is the multiplication table of $\{0, 3, 6, 9\}$:

\oplus_{12}	0	3	6	9
0	0	3	6	9
3	3	6	9	0
6	6	9	0	3
9	9	0	3	6

From the table we see that $H = \{0, 3, 6, 9\}$ forms a subgroup of \mathbb{Z}_{12} . Hence the smallest subgroup of \mathbb{Z}_{12} containing 3 is H .

We now generalize this concept. Let G be a group and $a \in G$. Any subgroup H of G containing a must contain aa , that is, a^2 ; a^2a that is a^3 etc.... In general, it must contain all positive, integral powers of a , that is, a^n for every positive integer n . Also, a subgroup containing a must also contain a^{-1} , and hence by the above argument it must contain all powers of a^{-1} , that is, a^{-m} for all positive integers m . In addition, it must contain $aa^{-1} = e = a^0$, that is, the identity element. Summarizing, we have shown that a subgroup of G containing a must contain $\{a^n | n \in \mathbb{Z}\}$. Thus we have the following result.

Theorem 7.20. *Let G be a group and $a \in G$. Then $H = \{a^n | n \in \mathbb{Z}\}$ is a subgroup of G and is the smallest subgroup of G which contains a . Moreover H is Abelian.*

Proof: *Step 1* To prove that H is a subgroup of G we shall use the two step test. Clearly $a \in H$, so that H is non-empty. Let $a^r, a^s \in H$ for some $r, s \in \mathbb{Z}$. Then $a^r a^s = a^{r+s} \in H$, $\because r + s \in \mathbb{Z}$. Thus, the product of two elements of H is an element of H , so that H is closed under the group operation on G . Let $a^r \in H$, for some $r \in \mathbb{Z}$, then $-r \in \mathbb{Z}$, so that $a^{-r} \in H$. Also $a^r a^{-r} = a^0 = e$, so that $(a^r)^{-1} = a^{-r} \in H$. Hence H is a subgroup of G .

Step 2 We shall now prove that H is the smallest subgroup of G containing $a \in H$. Let H_1 be a subgroup of G such that $a \in H_1$. Then every power of a belongs to H_1 , as H_1 is a group in its own right, i.e., $a^n \in H_1$ for all $n \in \mathbb{Z}$. Hence $H \subseteq H_1$. Thus H is the smallest subgroup of G containing a .

Step 3 Let $x, y \in H$ then $x = a^m, y = a^n$ for some $m, n \in \mathbb{Z}$. This implies that $xy = a^m a^n = a^{m+n} = a^{n+m} = a^n a^m = yx$. Thus H is Abelian. \square

In view of the above result we have the following definition of a subgroup generated by an element of the group.

Definition 7.7. (Cyclic subgroup): *Let G be a group and $a \in G$. Then subgroup $\{a^n | n \in \mathbb{Z}\}$ of G is called the cyclic subgroup of G generated by a .*

Cyclic subgroup generated by a is denoted by $\langle a \rangle$.

Remark 7.6. *In view of the above definition and Theorem 7.20 we see that $\langle a \rangle$ is the smallest subgroup of G containing a . Thus $\langle a \rangle = \{a^n | n \in \mathbb{Z}\}$. In case of additive notation $\langle a \rangle = \{na | n \in \mathbb{Z}\}$.*

Remark 7.7. From the definition it is clear that $\langle a \rangle = \langle a^{-1} \rangle$ as follows $\langle a \rangle = \{a^n | n \in \mathbb{Z}\} = \{a^{-n} | n \in \mathbb{Z}\} = \{(a^{-1})^n | n \in \mathbb{Z}\} = \langle a^{-1} \rangle$.

Example 7.16. Consider the group (\mathbb{Z}_4, \oplus_4) . Thus $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. If $a \in \mathbb{Z}_4$ then $\langle a \rangle = \{na | n \in \mathbb{Z}_4\}$. Let us find the cyclic subgroups generated by the non zero elements of \mathbb{Z}_4 .

$$\begin{aligned} \langle 1 \rangle &= \{n \oplus_4 1 | n \in \mathbb{Z}\} = \{0, 1, 1 \oplus_4 1, \dots\} \\ &= \{0, 1, 2, 3, 0, 1, \dots\} = \{0, 1, 2, 3\} \text{ using congruence modulo 4.} \\ \langle 2 \rangle &= \{n \oplus_4 2 | n \in \mathbb{Z}\} = \{0, 2, 0, 2, \dots\} = \{0, 2\} \\ \langle 3 \rangle &= \{n \oplus_4 3 | n \in \mathbb{Z}\} = \{0, 3, 2, 1, 0, 3, 2, 1, \dots\} = \{0, 1, 2, 3\} = \mathbb{Z}_4. \end{aligned}$$

Thus $\langle 1 \rangle = \langle 3 \rangle = \mathbb{Z}_4$.

Example 7.17. Consider the group $(\mathcal{U}(30), \odot_{30})$, where $\mathcal{U}(30) = \{1, 7, 11, 13, 17, 19\}$.

Then $\mathcal{U}(30)$ is a group of order 8. find the subgroup generated by 7.

$$\begin{aligned} \langle a \rangle &= \{a^n | n \in \mathbb{Z}\} \\ \langle 7 \rangle &= \{7^n | n \in \mathbb{Z}\} \\ &= \{7^0, 7^1, 7^2, \dots, 7^{-1}, 7^{-2}, \dots\} \\ &= \{1, 7, 7^2, \dots, (7^{-1}), (7^{-1})^2, \dots\} \\ &= \{1, 7, 19, 13, \dots, 13, (13)^2, \dots\} \\ &= \{1, 7, 19, 13, \dots, 13, 19, 7, 1, \dots\} \\ &= \{1, 7, 19, 13, \dots, 13, 19, 7, 1, \dots\} \\ &= \{1, 7, 13, 19\}. \end{aligned}$$

Example 7.18. In $(\mathbb{Z}, +)$, find $\langle 5 \rangle$, the cyclic subgroup of \mathbb{Z} generated by 5.

$$\begin{aligned} \langle 5 \rangle &= \{n \cdot 5 | n \in \mathbb{Z}\} \\ &= \{5n | n \in \mathbb{Z}\} \\ &= \text{all multiples of 5} \\ &= 5\mathbb{Z}. \end{aligned}$$

Remark 7.8. In $(\mathbb{Z}, +)$ the cyclic subgroup generated by n , i.e. $\langle n \rangle$ is $n\mathbb{Z}$.

The following Theorem shows that if an element a of a group G is of finite order n then $\langle a \rangle$ is a finite subgroup of G of order n .

Theorem 7.21. If G is any group and $a \in G$ of order n then $\langle a \rangle$ is a subgroup of order n . Moreover $\langle a \rangle = \{e, a, a^2, \dots, a^{n-1}\}$.

Proof: Let $a \in G$ be of order n . Then $a^n = e$. We know that $\langle a \rangle = \{a^m | m \in \mathbb{Z}\}$. Let $S = \{e, a, a^2, \dots, a^{n-1}\}$. Clearly $S \subseteq \langle a \rangle$. Let $x \in \langle a \rangle$. Then $x = a^m$ for some $m \in \mathbb{Z}$. By division algorithm, there exists $q, r \in \mathbb{Z}$ such that $m = nq + r$, $0 \leq r < n$.

$$\begin{aligned} \text{Now } a^m &= a^{nq+r}, \quad 0 \leq r < n \\ &= a^{nq} a^r \\ &= (a^n)^q a^r \\ &= e^q a^r \\ &= ea^r \\ &= a^r \in S \quad (\because 0 \leq r < n) \end{aligned}$$

Hence $x = a^m \in S$ so that $\langle a \rangle \subseteq S$. Hence $\langle a \rangle = S$. Clearly $\langle a \rangle$ is of order n , as $o(a) = n$ implies that all elements of S are distinct. □

Remark 7.9. If G is an additive group and $a \in G$ is of order m then $\langle a \rangle = \{na | 0 \leq n < m\} = \{0, a, 2a, \dots, (m-1)a\}$.

Using the above theorem, we find all cyclic subgroups of some groups.

Example 7.19. Consider the group $U(20)$. We find the order of each of its elements and hence obtain the cyclic subgroups generated by them.

$$U(20) = \{1, 3, 7, 9, 11, 13, 17, 19\}.$$

$$3^2 = 9, 3^3 = 7, 3^4 = 1 \quad \therefore o(3) = 4 \text{ and so } \langle 3 \rangle = \{1, 3, 3^2, 3^3\} = \{1, 3, 9, 7\} = \{1, 3, 7, 9\}.$$

$$7^2 = 9, 7^3 = 3, 7^4 = 1 \quad \therefore o(7) = 4 \text{ and so } \langle 7 \rangle = \{1, 7, 7^2, 7^3\} = \{1, 7, 9, 3\} = \{1, 3, 7, 9\}.$$

$$9^2 = 1 \quad \therefore o(9) = 2 \text{ and so } \langle 9 \rangle = \{1, 9\}.$$

$$\text{Similarly } o(11) = 2 \text{ and so } \langle 11 \rangle = \{1, 11\}.$$

$$o(13) = 4, \text{ and } \langle 13 \rangle = \{1, 9, 13, i7\}.$$

$$o(17) = 4, \text{ and } \langle 17 \rangle = \{1, 9, 13, 17\}.$$

$$o(19) = 2, \text{ and } \langle 19 \rangle = \{1, 19\}$$

Thus $U(20)$ has 4 cyclic subgroups of order 4 and 3 cyclic subgroups of order 2.

Example 7.20. Find all cyclic subgroups of $(\mathbb{Z}_{10}, \oplus_{10})$.

Note that $\langle a \rangle = \{na \mid n \in \mathbb{Z}\} =$ set of all integral multiples of a . Now $\mathbb{Z}_{10} = \{0, 1, 2, 3, \dots, 9\}$ is an additive group.

2 times 2 is $2 \oplus_{10} 2 = 4$, 3 times 2 is $2 \oplus_{10} 2 \oplus_{10} 2 = 6$, 4 times 2 is $2 \oplus_{10} 2 \oplus_{10} 2 \oplus_{10} 2 = 8$, 5 times 2 is $2 \oplus_{10} 2 \oplus_{10} 2 \oplus_{10} 2 \oplus_{10} 2 = 0$. Thus $o(2) = 5$ and $\langle 2 \rangle = \{0, 2, 4, 6, 8\}$. Similarly $o(3) = 10 \Rightarrow \langle 3 \rangle$ has 10 elements $\Rightarrow \langle 3 \rangle = \mathbb{Z}_{10}$.

$$o(4) = 5, \quad \langle 4 \rangle = \{4n(\text{mod}10) \mid 0 \leq n \leq 4\} = \{0, 2, 4, 6, 8\}$$

$$o(5) = 2, \quad \langle 5 \rangle = \{5n(\text{mod}10) \mid n = 0, 1\} = \{0, 5\}$$

$$o(6) = 5, \quad \langle 6 \rangle = \{0, 2, 4, 6, 8\}$$

$$o(7) = 10, \quad \langle 7 \rangle = \mathbb{Z}_{10}$$

$$o(8) = 5, \quad \langle 8 \rangle = \{0, 2, 4, 6, 8\}$$

$$o(9) = 10, \quad \langle 9 \rangle = \mathbb{Z}_{10}.$$

Example 7.21. In $U(15)$ what are the orders of $\langle 2 \rangle$ and $\langle 11 \rangle$. Also write their elements.

We know that $\langle a \rangle = \{a, a^2, a^3, \dots, a^{n-1}, a^n = e\}$, when $o(a) = n$. In $U(15)$, $o(2) = 4 \therefore o(2) = 4$ and $\langle 2 \rangle = \{2^0 = 1, 2, 2^2, 2^3\} = \{1, 2, 4, 8\}$ and $o(11) = 2$ ($\because 11^2 \equiv 1 \pmod{15}$) and $\langle 11 \rangle = \{1, 11\}$.

Example 7.22. We find the cyclic subgroup generated by $A \in SL(2, \mathbb{R})$ where $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We know that $\langle A \rangle = \{A^n \mid n \in \mathbb{Z}\}$. By induction it can be

proved that $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ for $n \in \mathbb{N}$. Since $A^n \neq I$ for any $n \in \mathbb{N}$, so that

A is not of finite order. $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, so that $A^{-n} = \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix}$, so that

$\langle A \rangle = \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z} \right\}$. Hence $\langle A \rangle$ is an infinite cyclic group.

7.10 Solved Problems

Problem 7.16. Let \mathbb{Q} be the group of rational numbers under addition and let \mathbb{Q}^* be the group of nonzero rational numbers under multiplication.

(i) In \mathbb{Q} list the elements in $\langle \frac{1}{2} \rangle$.

- (ii) In \mathbb{Q}^* list the elements in $\langle \frac{1}{2} \rangle$.
- (iii) Find the order of each element of \mathbb{Q} .
- (iv) Find the order of each element of \mathbb{Q}^* .

Solution:

- (i) In $(\mathbb{Q}, +)$ $\langle \frac{1}{2} \rangle = \{n \cdot \frac{1}{2} | n \in \mathbb{Z}\} = \{\dots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots\}$.
- (ii) In (\mathbb{Q}^*, \cdot) , $\langle \frac{1}{2} \rangle = \{(\frac{1}{2})^n : n \in \mathbb{Z}\} = \{1, \frac{1}{2}, 2, \frac{1}{4}, 4, \dots\}$.
- (iii) 0 being the identity, order of 0 is 1. If $a \in \mathbb{Q}, a \neq 0$ then there does not exist any positive integer n , such that $na = 0$, so that $o(a)$ is infinite.
- (iv) $o(1) = 1$, 1 being the identity element in \mathbb{Q}^* . Also $(-1)^2 = 1$ so that $o(-1) = 2$. If $a \in \mathbb{Q}^*, a \neq \pm 1, a^n \neq 1$ for any positive integer n . Hence a is of infinite order. Thus -1 is of order 2 and all other elements different from identity are of infinite order.

Problem 7.17. List the cyclic subgroups of $U(30)$.

Solution: $U(30) = \{1, 7, 11, 13, 17, 19, 23, 29\}$

$$\begin{aligned} \langle 1 \rangle &= \{1\} \\ \langle 7 \rangle &= \{7, 19, 13, 1\} = \{1, 7, 13, 19\} \\ \langle 11 \rangle &= \{11, 1\} = \{1, 11\} \\ \langle 13 \rangle &= \{13, 19, 7, 1\} = \{1, 7, 13, 19\} \\ \langle 17 \rangle &= \{17, 19, 23, 1\} = \{1, 17, 19, 23\} \\ \langle 19 \rangle &= \{19, 1\} = \{1, 19\} \\ \langle 23 \rangle &= \{23, 19, 17, 1\} = \{1, 17, 19, 23\} \\ \langle 29 \rangle &= \{29, 1\} = \{1, 29\}. \end{aligned}$$

Thus the cyclic subgroups of $U(30)$ are — $\langle 1 \rangle, \langle 11 \rangle, \langle 19 \rangle, \langle 29 \rangle, \langle 7 \rangle = \langle 13 \rangle, \langle 17 \rangle = \langle 23 \rangle$ as described above. There are 6 distinct cyclic subgroups. Note that $U(30)$ has 4 elements of order 4 but only 2 subgroups of order 4.

Problem 7.18. Suppose G has exactly 8 element of order 3. How many subgroup of order 3 does G have?

Solution: We assert that every subgroup H of order 3 contains exactly two elements of order 3. For, let $e \neq a \in H$. If $b = a^2$. Then $b \neq e$. Also $b^2 = a^4 = a \neq e$. $b^3 = a^6 = e$. Thus $o(b) = 3$. Also $\langle a \rangle = \{e, a, a^2\} = \{e, a, b\}$. Now $\langle a \rangle$ is a subgroup of order 3 containing exactly 2 elements of order 3. Since $\langle a^2 \rangle = \{e, a, a^2\} \therefore \langle a \rangle = \langle a^2 \rangle$. Thus two distinct elements a and a^2 of order 3 generate the same subgroup of order 3. Hence if there are 8 elements of order 3 then there are exactly 4 subgroups of order 3.

Problem 7.19. Find the smallest subgroup of \mathbb{Z} containing both 8 and 12.

Solution: Let H and K be the smallest subgroups of \mathbb{Z} containing 8, 12 respectively. Thus $H = \langle 8 \rangle, K = \langle 12 \rangle$. So $H = \{0, \pm 8, \pm 16, \pm 24, \pm 36, \dots\}$ and $K = \{0, \pm 12, \pm 24, \pm 36, \dots\}$. Let L be the smallest subgroup of \mathbb{Z} containing both H and K . Thus L must contain $12 - 8 = 4$. This is the smallest positive element of L .

$$\therefore \langle 4 \rangle \subseteq L \tag{7.19}$$

Also $8 \in \langle 4 \rangle \Rightarrow \langle 8 \rangle \subseteq \langle 4 \rangle$ and $12 \in \langle 4 \rangle \Rightarrow \langle 12 \rangle \subseteq \langle 4 \rangle$ and consequently $\langle 4 \rangle$ is a subgroup containing both H and K and so

$$L \subseteq \langle 4 \rangle \quad (7.20)$$

From (7.19) and (7.20) we get $L = \langle 4 \rangle$.

Problem 7.20. D_8 has 7 cyclic subgroups. List them. Find a subgroup of D_8 of order 4 which is not cyclic.

Solution: $D_8 = \{R_0, R_1, R_2, R_3, H, V, D, D'\}$

Clearly $o(R_0) = 1$, $o(R_2) = o(H) = o(V) = o(D) = o(D') = 2$ $o(R_1) = o(R_3) = 4$, $R_1^2 = R_2$, $R_1^3 = R_3$, $R_1^4 = R_0$

$\therefore \langle R_1 \rangle = \{R_0, R_1, R_2, R_3\}$ similarly $\langle R_3 \rangle = \{R_0, R_1, R_2, R_3\}$.

$\langle R_2 \rangle = \{R_0, R_2\}$

$\langle H \rangle = \{R_0, H\}$

$\langle V \rangle = \{R_0, V\}$

$\langle D \rangle = \{R_0, D\}$

$\langle D' \rangle = \{R_0, D'\}$

Also $\langle R_1 \rangle = \langle R_3 \rangle$, $\langle R_0 \rangle = \{R_0\}$. Thus there are seven cyclic subgroups.

Also $V_4 = \{R_0, H, V, R_2\}$ is a non-cyclic subgroup of order 4. V_4 (Viergruppe) is called Klein 4-group.

7.11 Exercise

1. List the elements of the subgroups $\langle 3 \rangle$ and $\langle 15 \rangle$ in $(\mathbb{Z}_{18}, \oplus_{18})$.
2. List the elements of $\langle 3 \rangle$ and $\langle 7 \rangle$ in $\mathcal{U}(20)$.
3. $\mathcal{U}(15)$ has six cyclic subgroups. List them.
4. In $SL(2, \mathbb{R})$, find $\langle A \rangle$ where $A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.

7.12 Lattice of Subgroups

We know that, in general, a group has subgroups, different from the trivial subgroups. We will describe a diagram associated with a group representing the relationship between its subgroups. This diagram, called the lattice of subgroups of a group (or subgroup lattice) is a very good way of visualizing a group. The structure of a group can certainly be seen in a better way than from the multiplication table of the group. In a sense we can say that the subgroup lattice gives a 'family photo' of the group.

The lattice of subgroups of a finite group G is constructed as follows:

1. Plot subgroups of G starting with $\langle e \rangle$ at the bottom and ending with G at the top. Subgroup of larger order may be positioned higher on the page than those of smaller order.
2. Draw a path upwards between subgroup using the following rule:

There will be a line segment upward from H to K if $H < K$ and there are no subgroups L such that $H < L < K$. Thus if $H < K$ there is a

many path) upward from H to K passing through a chain of subgroups. The initial positioning of the subgroups on the diagram is somewhat arbitrary. With some adjustment we can produce a diagram which is pleasing to the eye.

A unique way of drawing a diagram of a subgroup lattice is to draw a diagram of a subgroup lattice procedure outlined

to draw a subgroup lattice:

1. List the subgroups of the given finite group.

2. Draw chains of subgroups starting from $\langle e \rangle$ and ending at G .

3. Place the chains with $\langle e \rangle$ at the bottommost position and G at the top.

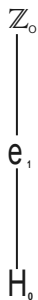
4. Indicate the common elements of the chain (i.e. those shared by 2 or more chains) only once.

5. Adjust the subgroups in the subgroup lattice of G . If needed minor positioning adjustments are done to get a beautiful look.

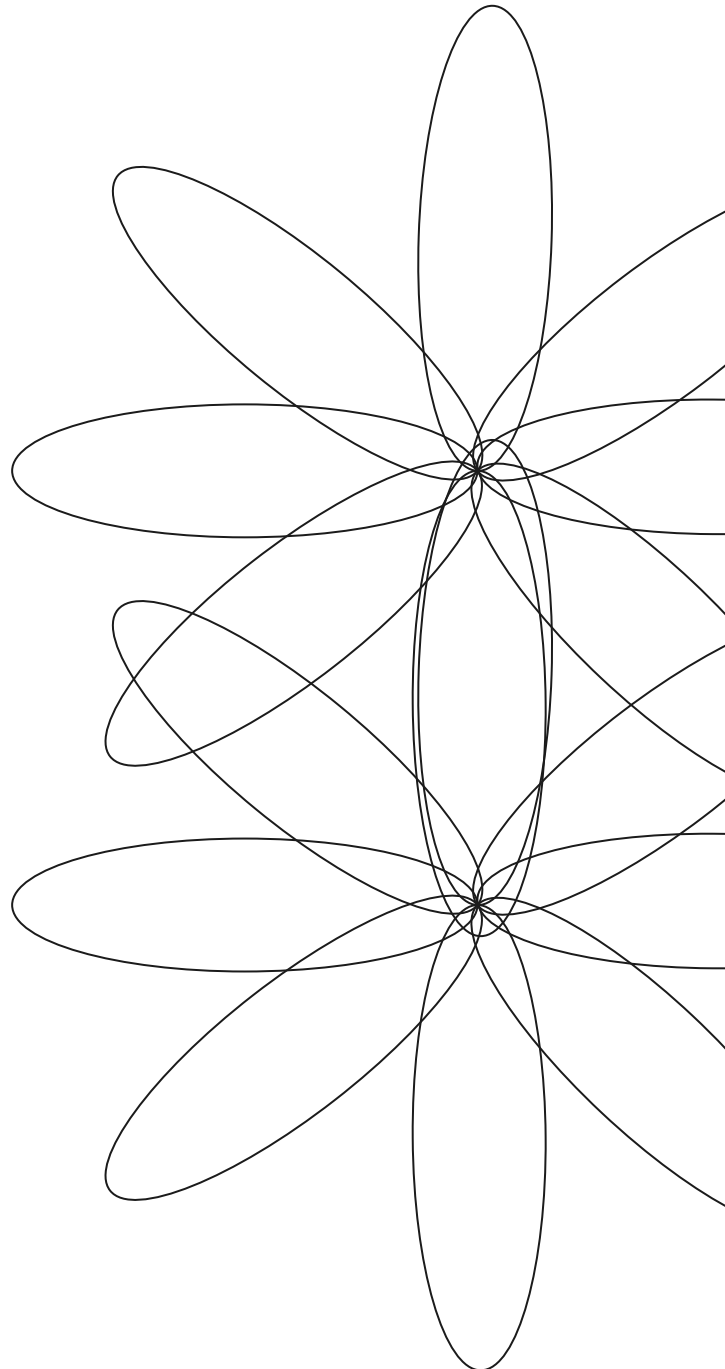
Example: Draw the subgroup lattice of (\mathbb{Z}_2, \oplus_2) , $\mathbb{Z}_2 = \{0, 1\}$. It has subgroups $\langle 0 \rangle$ and \mathbb{Z}_2 . Its subgroup lattice is

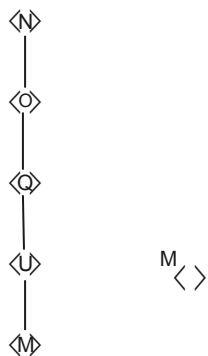


Example: Draw the subgroup lattice of (\mathbb{Z}_4, \oplus_4) , $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. The subgroups are $H_0 = \{0\}$, $H_1 = \{0, 2\} = \langle 2 \rangle$, \mathbb{Z}_4 . The only chain is $\langle 0 \rangle < \langle 2 \rangle < \mathbb{Z}_4$. Its subgroup lattice is:

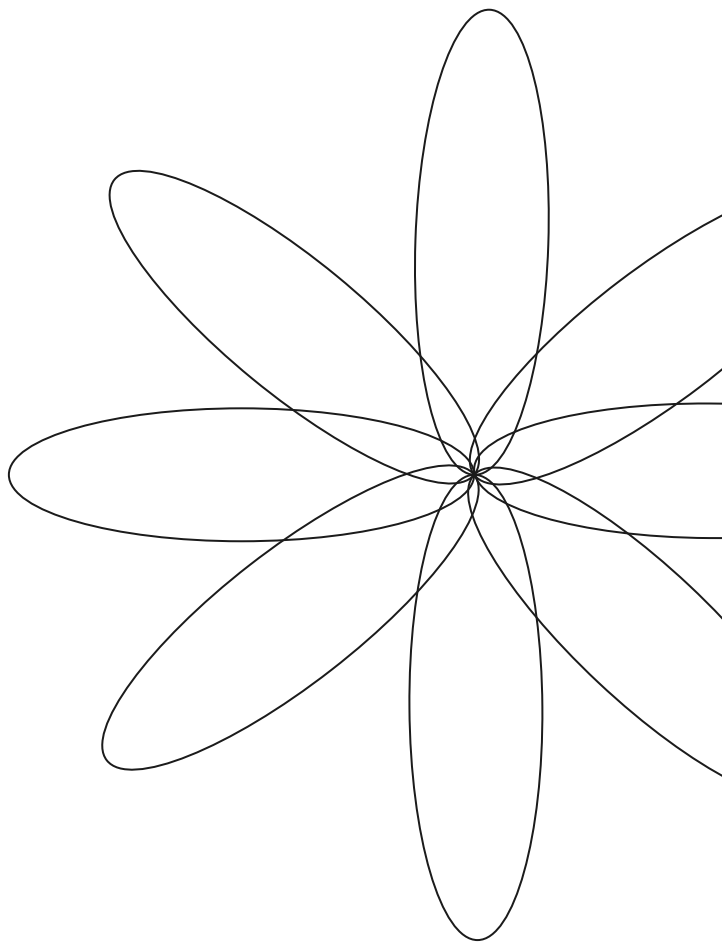


Example: Draw the subgroup lattice of $(\mathbb{Z}_{16}, \oplus_{16})$, $\mathbb{Z}_{16} = \{0, 1, 2, \dots, 15\}$. The subgroups of \mathbb{Z}_{16} are $H_0 = \langle 0 \rangle$, $H_2 = \langle 2 \rangle$, $H_4 = \langle 4 \rangle$, $H_8 = \langle 8 \rangle$, \mathbb{Z}_{16} . The only chain is $\langle 0 \rangle < \langle 8 \rangle < \langle 4 \rangle < \langle 2 \rangle < \langle 1 \rangle$. Its subgroup lattice is:





Draw the subgroup lattice of $(p^n\mathbb{Z}, \oplus_{p^n})$ where p is a prime. The subgroups of \mathbb{Z} are $\langle 1 \rangle, \langle p \rangle, \langle p^2 \rangle, \dots, \langle p^{n-1} \rangle, \langle p^n \rangle = \langle 0 \rangle$. The only chain is $\dots < \langle p^2 \rangle < \langle p \rangle < \langle 1 \rangle = p^n\mathbb{Z}$. The subgroup lattice is:

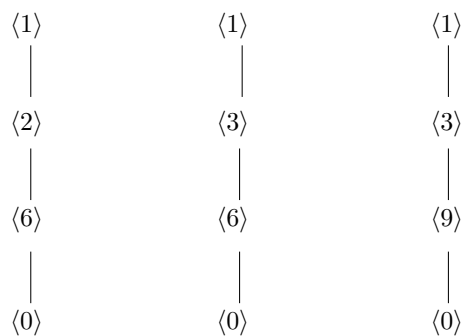


>

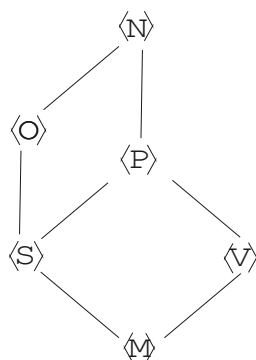
Draw the subgroup lattice of $(\mathbb{Z}_{18}, \oplus_{18})$ where $\mathbb{Z}_{18} = \{0, 1, \dots, 17\}$. The subgroups of \mathbb{Z}_{18} are obtained as follows: $\langle 0 \rangle, \langle 9 \rangle, \langle 6 \rangle, \langle 3 \rangle, \langle 2 \rangle, \langle 1 \rangle$ respectively. The chains are:

- $\langle 2 \rangle < \langle 1 \rangle$
- $\langle 3 \rangle < \langle 1 \rangle$
- $\langle 6 \rangle < \langle 1 \rangle$

vertically as follows:



Writing the common elements, $\langle 0 \rangle$, $\langle 6 \rangle$, $\langle 3 \rangle$ and $\langle 1 \rangle$ only once we get the subgroup lattice as shown.



Problem 7.21. Let $S = \{a, b, c, 1\}$ with the following multiplication table:

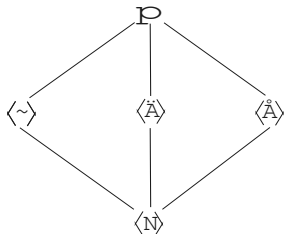
	1	a	b	c
1	1	a	b	c
a	a	1	c	b
b	b	c	1	a
c	c	b	a	1

Draw the lattice diagram of S .

Solution: From the table we see that 1 is the identity element, and $a^2 = b^2 = c^2 = 1$. Thus subgroups of S are $\langle 1 \rangle$, $\langle a \rangle$, $\langle b \rangle$, $\langle c \rangle$, S . The chains are:

$$\begin{array}{l}
 \langle 1 \rangle < \langle a \rangle < S \\
 \langle 1 \rangle < \langle b \rangle < S \\
 \langle 1 \rangle < \langle c \rangle < S
 \end{array}$$

The lattice diagram is



Problem 7.22. Draw the subgroup lattice of (i) S_3 (ii) D_8 .

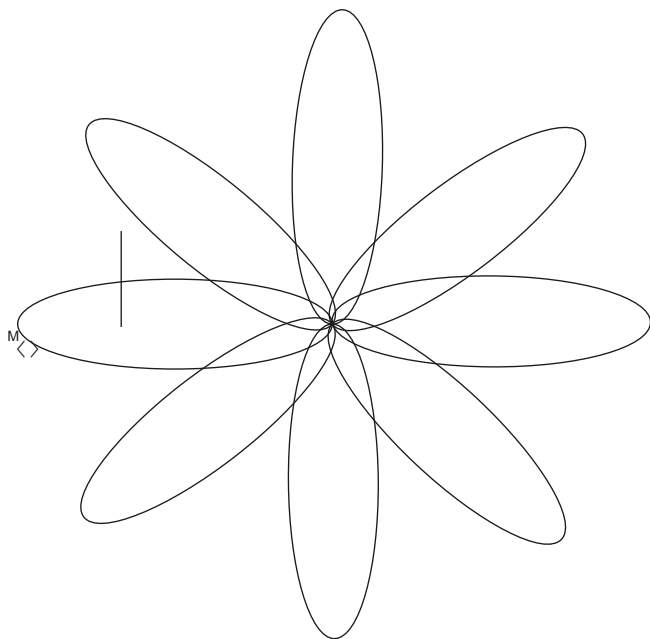
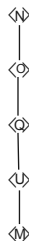
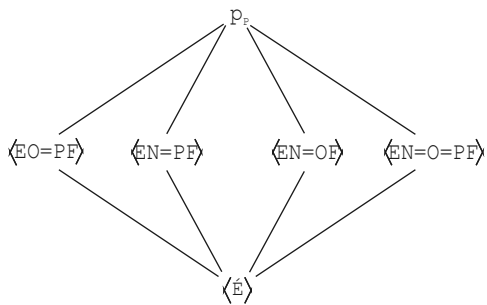
Solution:

(i) $S = \{1, 2, 3\}$, $S_3 = \{e, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$. Subgroups of S_3 are: $\langle(1\ 2)\rangle = \{e, (1\ 2)\}$, $\langle(1\ 3)\rangle = \{e, (1\ 3)\}$, $\langle(2\ 3)\rangle = \{e, (2\ 3)\}$, $\langle(1\ 2\ 3)\rangle = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$

The chains are:

$$\begin{aligned} \langle e \rangle &< \langle(1\ 2)\rangle < S_3 \\ \langle e \rangle &< \langle(1\ 3)\rangle < S_3 \\ \langle e \rangle &< \langle(2\ 3)\rangle < S_3 \\ \langle e \rangle &< \langle(1\ 2\ 3)\rangle < S_3 \end{aligned}$$

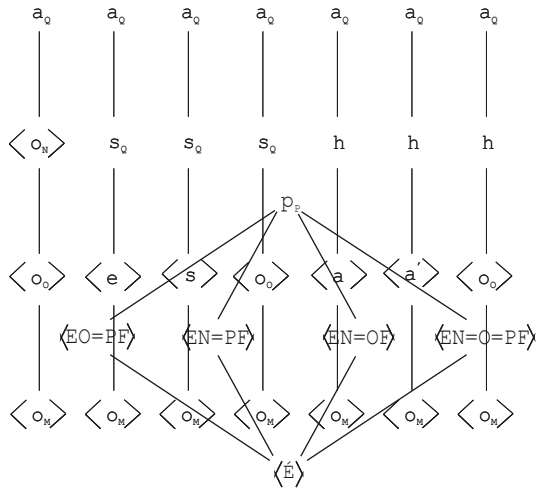
The lattice diagram is:



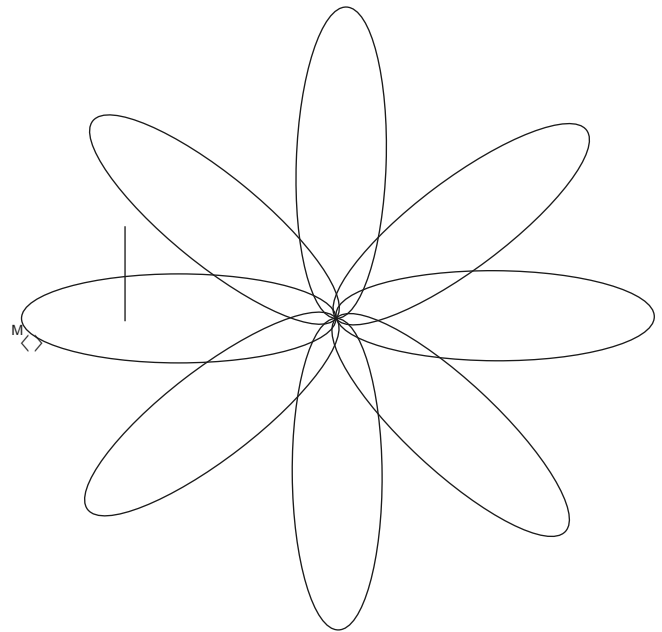
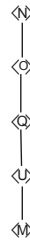
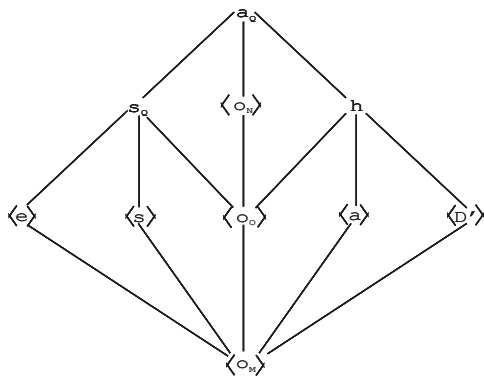
(ii) $D_8 = \{R_0, R_1, R_2, R_3, H, V, D, D'\}$ R_0 is the identity element. The subgroup of D_8 are : $\langle R_0 \rangle < \langle R_1 \rangle, \langle R_2 \rangle, \langle H \rangle, \langle R_2 \rangle, \langle H \rangle, \langle V \rangle, \langle D \rangle, \langle D' \rangle$ and see problem p75. Also chain are

$$\begin{aligned} \langle R_0 \rangle &< \langle R_2 \rangle < \langle R_1 \rangle < D_8 \\ \langle R_0 \rangle &< \langle H \rangle < V_4 < D_8 \\ \langle R_0 \rangle &< \langle V \rangle < V_4 < D_8 \\ \langle R_0 \rangle &< \langle R_2 \rangle < V_4 < D_8 \\ \langle R_0 \rangle &< \langle D \rangle < K < D_8 \\ \langle R_0 \rangle &< \langle D' \rangle < K < D_8 \\ \langle R_0 \rangle &< \langle R_2 \rangle < K < D_8 \end{aligned}$$

writing the chains vertically,



The subgroup lattice is:



7.13 Exercise

1. Draw the subgroup lattice of the following groups:
 - (i) \mathbb{Q}_8 , the group of quaternions
 - (ii) (\mathbb{Z}_3, \oplus_3)
 - (iii) (\mathbb{Z}_9, \oplus_9)
 - (iv) $(\mathbb{Z}_{12}, \oplus_{12})$
 - (v) D_6 , the dihedral group of order 6
 - (vi) $(\mathbb{Z}_{30}, \oplus_{30})$
 - (vii) $(\mathcal{U}(12), \odot_{12})$
 - (viii) $(\mathcal{U}(8), \odot_8)$.

7.14 Supplementary Exercises

1. State whether the following statements are true or false. Justify your answer. Also correct the false statements.
 - (i) Every subset H of a group G is a subgroup under the binary operation restricted to H .
 - (ii) Every group is a subgroup of itself.
 - (iii) Every set of numbers which is a group under addition is also a group under multiplication.
 - (iv) There are groups in which cancellation laws do not hold.
 - (v) The identity element of a subgroup can be different from the identity element of a group.
 - (vi) If H is a subgroup of G and $a \in H$ then the inverses of a as elements of H and G can be different.
 - (vii) The group of even integers, under addition is cyclic.
 - (viii) The cyclic group \mathbb{Z} has a unique generator.
 - (ix) The set of all purely imaginary complex numbers is a subgroup of the set of all non-zero complex numbers under multiplication.
 - (x) Every subgroup of an Abelian group is Abelian.
 - (xi) Every subgroup of a non Abelian group is non-Abelian.
 - (xii) Every element of a group generates a cyclic subgroup of the group.
 - (xiii) Every non-Abelian group has at least one non-trivial Abelian subgroup.
 - (xiv) If a and b are elements of finite order in a group G such that $ab = ba$, then ab is also of finite order.
 - (xv) In a group G if a and b are elements of G which commute, such that, $o(a) = 3$, $o(b) = 4$ then the order of ab is 12.
 - (xvi) In a group G if a and b are elements of G which commute, such that $o(a) = m$, $o(b) = n$ then $o(ab)$ is mn .
 - (xvii) An element of a group of finite order may have infinite order.
 - (xviii) A subset H of a finite group G is a subgroup if H is closed.
 - (xix) The set of all complex numbers which lie on the circumference of a circle centered at the origin and radius 2 is a subgroup of multiplicative group \mathbb{C}^* .
 - (xx) A group of order 8 cannot have a subgroup of order 6.
 - (xxi) The dihedral group D_n of symmetries of a regular polygon of n sides has order n , for $n \geq 3$.
 - (xxii) Every proper subgroup of the group of quaternions is Abelian.
 - (xxiii) Every dihedral group D_{2n} for $n \geq 3$, is non-Abelian.
 - (xxiv) Every dihedral group D_{2n} , for $n \geq 3$ has a cyclic subgroup of order n .
 - (xxv) D_8 has 4 cyclic subgroups of order 2 and one cyclic subgroup of order 4.
2. Let G be a group and let S be the set of all subgroups of G . On S , define a relation \sim as follows: $A \sim B$ if A is a subgroup of B . Is this relation an equivalence relation? Justify you answer.
3. Let G be a group and n be a fixed positive integer. Let $G^n = \{g^n : g \in G\}$.

Prove or disprove that G^n is a subgroup of G .

4. A and B are subsets of a group G . Prove that
 - (i) $C(A \cup B) = C(A) \cap C(B)$.
 - (ii) $C(A) \cup C(B) \subseteq C(A \cap B)$.
5. For any subset A of a group G , obtain a containment relationship between the centre of G , centralizer of A and the normalizer of A .
6. Find the centre of the quaternion group Q_8 .
7. Find a cyclic subgroup of order 4 in $U(40)$.
8. Find a non-cyclic subgroup of order 4 in $U(40)$.
9. Let G be an Abelian group and p any prime number. Show that the set of all elements of G whose orders are powers of p , is a subgroup of G .
10. Give an example of a group which is not cyclic, but its every proper subgroup is cyclic.
11. Let $G = GL(3, \mathbb{Q})$. Let $H = \{A \in G : |A| \text{ is an integral power of } 3\}$. Show that H is a subgroup of G .
12. List the elements of the subgroups $\langle 20 \rangle$ and $\langle 10 \rangle$ in $(\mathbb{Z}_{30}, \oplus_{30})$.
13. If a group has exactly 4 elements of order 4, then how many subgroups of order 4 are there?
14. Find all cyclic subgroups of the group G . Is there a proper subgroup of G which is not cyclic, where
 - (i) $G = D_6$?
 - (ii) $G = D_8$?
 - (iii) $G = Q_8$?
15. Let G be a finite group with at least 2 elements. Show that G has an element of prime order.
16. For any element a in a group G , prove that $\langle a \rangle$ is a subgroup of $C(a)$.
17. In a group G , if a is the only element of order 2, then prove that a lies in the center of G .
18. Prove that every non-Abelian group has at least two non-trivial Abelian subgroups.
19. G is a group and $a, b \in G$ such that $ab = ba$. Prove that $\langle b \rangle \subseteq C(a)$.
20. In a group, for any $x \in G$, prove that $\langle x \rangle \leq N_G(\langle x \rangle)$. Further show that equality need not hold.

7.15 Answers to Exercises

Exercise - 7.3

1. If $H = \phi$, no such a exists.
2. (i), (iii), (v), (vi) are Yes; (ii), (iv) are No.
4. H_1 is a subgroup. H_2 is not closed so not a subgroup.
5. (i) V_4 is an Abelian subgroup of non-Abelian group D_8 . $\{\pm 1, \pm i\}$ is a finite subgroup of (C^*, \cdot) .
9. H is not a subgroup of $\mathcal{U}(20)$ $\because 7 \in H$ but $7 \odot_{20} 7 = 9 \notin H$.
10. H_1, H_3 are subgroups.
 H_2 is not a subgroup as $\left\{ \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \right\}$ does not have an inverse.
 H_4 is not a subgroup as $\left\{ \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \right\}$ does not have an inverse.
11. H is not a subgroup as $2 + 3i, -3 - 2i \in H$ but their sum $-1 + i$ is not in H .

Exercise - 7.5

3. $\{\pm 1, \pm i\}$, $Z(G) = \{\pm 1\}$
4. No, In the group of Quaternions Q_8 , if $A = \{i, j\}$, $B = \{i, k\}$ then $C(A) = C(B)$ but $A \neq B$.
6. $C(A) = \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$, $C(B) = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$.
7. $\{R_0, R_2\}$
9. G
12. $Z(G) \subseteq C(A) \subseteq N(A)$

Exercise - 7.8

2. *Hint:* $ab = b^{-1}(ba)b$.
3. The elements in the pair are inverses of each other.
5. $o(x) = 3$ or 6 .
6. $o(x^2) = o(x^4) = o(x^5) = o(x^7) = o(x^8) = 9$, $o(x^3) = o(x^6) = 3$.
7. *Hint:* if $o(a) = n$, then $o(a^m) = \frac{n}{\gcd(m, n)}$
 - (i) $o(a) = 12$ or 60 .
 - (ii) $o(a) = 48$.

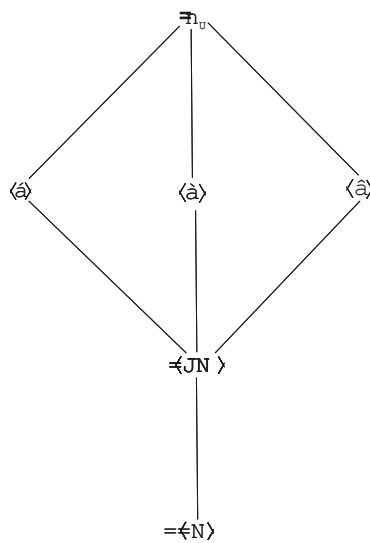
8. $o(A) = 6$ $B^n = \begin{pmatrix} \cos(n\sqrt{2}) & -\sin(n\sqrt{2}) \\ \sin(n\sqrt{2}) & \cos(n\sqrt{2}) \end{pmatrix} = I$ when $n\sqrt{2} = 360^\circ$. But no such n exists. Hence $o(B)$ is infinite.
9. $o(A) = 4, o(B) = 3, (AB)^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ so that AB is of infinite order. A, B are of finite orders but AB is of infinite order.
10. $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$: In $SL(2, \mathbb{R})$ $o(A)$ is infinite. In $SL(2, \mathbb{Z}_p)$, $o(A) = p$.
11. $n =$ product of the orders of the elements of G .

Exercise - 7.11

1. $\langle 3 \rangle = \{0, 3, 6, 9, 12, 15\} = \langle 15 \rangle$.
2. $\langle 3 \rangle = \{1, 3, 9, 7\} = \langle 7 \rangle$.
3. $\langle 1 \rangle; \langle 4 \rangle; \langle 11 \rangle; \langle 14 \rangle; \langle 2 \rangle = \langle 8 \rangle; \langle 7 \rangle = \langle 13 \rangle$.
5. $\langle A \rangle = \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}$.

Exercise - 7.13

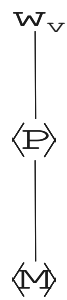
(i)



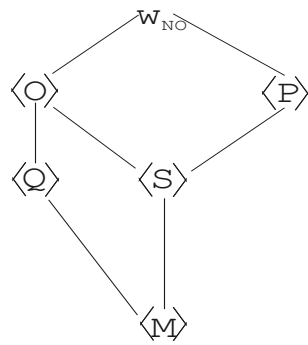
(ii)



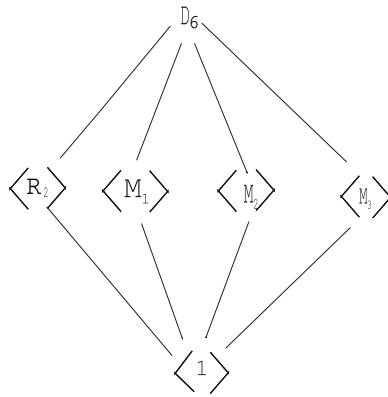
(iii)



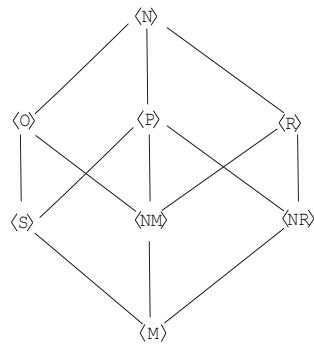
(iv)



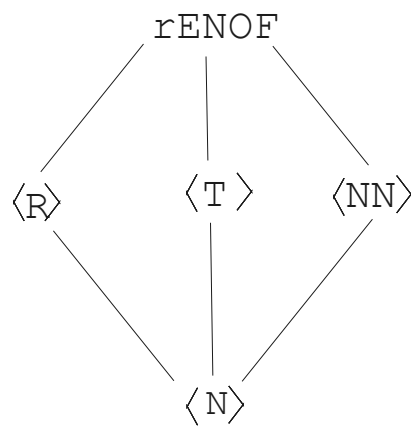
(v)



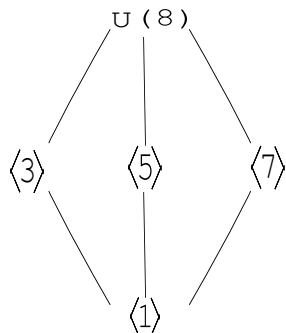
(vi)



(vii)



(viii)



Supplementary Exercise - 7.14

1.
 - (i) H should be a group.
 - (ii) T
 - (iii) F, $(\mathbb{Z}, +)$ is a group but (\mathbb{Z}, \cdot) is not.
 - (iv) F, cancellation laws always hold in a group.
 - (v) F, they are same.
 - (vi) F, they are same.
 - (vii) T
 - (viii) F, it has 2 generators ± 1 .
 - (ix) F, i^2 is not pure imaginary.
 - (x) T
 - (xi) F, subgroup of a non-Abelian group may be Abelian. Set of all 2×2 scalar matrices over \mathbb{R} is an Abelian subgroup of non-Abelian group $GL(2, \mathbb{R})$.
 - (xii) T
 - (xiii) T
 - (xiv) T
 - (xv) T
 - (xvi) F, $o(ab) = \text{lcm of } m \text{ and } n$.
 - (xvii) F, every element is of finite order.
 - (xviii) F, H should be non-empty.
 - (xix) F, $|Z_1| = 2$, $|Z_2| = 2$, then $|Z_1 Z_2| = 4 \neq 2$.
 - (xx) T
 - (xxi) T
 - (xxii) T
 - (xxiii) T
 - (xxiv) T
 - (xxv) F, 5 of order 2.
2. No, it is not symmetric. It is reflexive and transitive.
3. G^n is not a group, D_6 is a group but D_6^3 is not a subgroup of D_6 . If G is Abelian G^n in a group.
5. $Z(G) \subseteq C(A) \subseteq N(A)$

6. $\{1, -1\}$
7. $\langle 3 \rangle$
8. *Hint:* A non-cyclic subgroup of order 4 in $\mathcal{U}(40)$ must be of the form $\{e, a, b, ab\}$ such that $o(a) = o(b) = 2$. It is $\{1, 9, 11, 19\}$.
10. V_4, Q_8 .
12. $\{0, 10, 20\} = \langle 20 \rangle = \langle 10 \rangle$
13. *Hint:* If $o(a) = 4$, then $o(a^3) = 4 \quad \therefore \langle a \rangle = \langle a^3 \rangle$. Number of subgroups of order 4 = $\frac{1}{2}(\text{Number of elements of order 4}) = 2$.
14. (i) $\langle R_0 \rangle, \langle R_1 \rangle, \langle M_1 \rangle, \langle M_2 \rangle, \langle M_3 \rangle$. No proper subgroup is non-cyclic.
 (ii) $\langle R_0 \rangle, \langle R_1 \rangle, \langle R_2 \rangle, \langle H \rangle, \langle V \rangle, \langle D \rangle, \langle D' \rangle, V_4 = \{R_0, H, V, R_2\}$,
 $K = \{R_0, D, D', R_2\}$ are non-cyclic subgroups of D_8 .
 (iii) $\langle 1 \rangle, \langle -1 \rangle, \langle i \rangle, \langle j \rangle, \langle k \rangle$. No.
15. *Hint:* Let $e \neq a \in G$ and $o(a) = n$. Then there exists a prime p which divides n . Let $m = \frac{n}{p}$. Then $o(a^m) = p$.
17. *Hint:* $o(xax^{-1}) = o(a) \quad \forall x \in G$.
18. *Hint:* $e \neq a \in G$, $\langle a \rangle$ is Abelian, G non-Abelian $\Rightarrow \langle a \rangle \neq G \quad \exists b \in G \sim D \sim \langle a \rangle$. Then $\langle b \rangle$ is also Abelian.
20. $G = Q_8, \quad x = i$.

Chapter 8

Cyclic Groups

In the previous chapter we have defined a cyclic subgroup of a group. Recall that if a is an element of a group G , then $\{a^n | n \in \mathbb{Z}\}$ is a subgroup of G , called the cyclic subgroup of G generated by a and is written as $\langle a \rangle$. In this chapter we shall study cyclic groups and their properties.

8.1 Definition and Examples

Definition 8.1. A group G is said to be cyclic if there exists some $a \in G$ such that $\langle a \rangle$, the subgroup generated by a is whole of G . The element a is called a generator of G or G is said to be generated by a .

Thus $G = \langle a \rangle = \{a^n | n \in \mathbb{Z}\}$. If the binary operation is addition, then $G = \langle a \rangle = \{na | n \in \mathbb{Z}\}$.

Remark 8.1. If G is a finite cyclic group of order n , generated by a , then $G = \{a, a^2, a^3, a^4, \dots, a^{n-1}, a^n = e\}$.

An immediate consequences of the definition.

Theorem 8.1. Every cyclic group is Abelian.

Proof: Follows from Theorem 7.20. □

Example 8.1. Let $G = \{1, -1\}$, then G is a group with respect to multiplication.

$$\text{Since } (-1)^n = \begin{cases} -1 & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{ is even} \end{cases}$$

Therefore $\langle -1 \rangle = \{(-1)^n | n \in \mathbb{Z}\} = \{-1, 1\} = G$.

Hence G is a finite cyclic group of order 2.

Example 8.2. Consider the group $(\mathbb{Z}, +)$. We show that \mathbb{Z} is an infinite cyclic group. $\langle 1 \rangle = \{n1 | n \in \mathbb{Z}\} = \mathbb{Z}$. Since $n1 \neq 0$ for any $n \in \mathbb{Z}$, therefore $\langle 1 \rangle$ is infinite. Since additive inverse of 1 is -1 , therefore $\langle -1 \rangle = \langle 1 \rangle$. Hence $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$. Thus \mathbb{Z} is a cyclic group having at least two generators 1 and -1 . It is an infinite cyclic group.

Let us recall. "If G is a group and $a \in G$ such that $o(a) = n$, then $\langle a \rangle$ is a finite subgroup of G of order n ." Thus we have the following theorem:

Theorem 8.2. *Let G be a finite group of order n , then $G = \langle a \rangle$ for some $a \in G$ if and only if a is of order n . Further, if $o(a) = n$ then $G = \langle a \rangle$.*

Proof: Given $o(G) = n$. Let G be a cyclic group and $G = \langle a \rangle$ for some $a \in G$. Then, $o(G) = o(\langle a \rangle) = o(a)$. Hence $o(a) = n$.

Conversely, suppose that $o(a) = n$, and $H = \langle a \rangle$. Then H is a cyclic subgroup of G of order n . Also $H \subseteq G$, since G is finite and $o(H) = o(G)$, we get $H = G$. Hence $G = \langle a \rangle$, so that G is cyclic. \square

The above theorem tells us that in a finite group of order n , every element of order n is a generator of the group. But this is not the case if the group is not finite, i.e, in an infinite cyclic group, every element of infinite order may not be its generator. This is shown by the following example.

Example 8.3. *Consider the group $(\mathbb{Z}, +)$. Then $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$. Thus \mathbb{Z} is an infinite cyclic group. Observe that $2 \in \mathbb{Z}$ is of infinite order, as $2n \neq 0$ for any $n \in \mathbb{N}$. Further,*

$$\begin{aligned} \langle 2 \rangle &= \{2z : z \in \mathbb{Z}\} \\ &= \text{Set of even integers} \\ &\neq \mathbb{Z} \end{aligned}$$

Thus 2 is not a generator of \mathbb{Z} .

The condition of finiteness is also important in another sense. If G is an infinite group and G has an element of infinite order, still G may fail to be cyclic. This is shown in the following example.

Example 8.4. *Consider the group $(\mathbb{Q}, +)$. \mathbb{Q} is an infinite group. We assert that it is not cyclic. Let if possible \mathbb{Q} be cyclic and be generated by $\frac{p}{q}$, where $(p, q) = 1$. Without any loss of generality, we can take $\frac{p}{q}$ to be positive, such that $\mathbb{Q} = \langle \frac{p}{q} \rangle = \{n(\frac{p}{q}) : n \in \mathbb{Z}\}$. Now $\frac{3}{2}(\frac{p}{q}) \in \mathbb{Q}$. $\frac{3}{2}(\frac{p}{q}) \neq k(\frac{p}{q})$ for any $k \in \mathbb{Z}$. Therefore $\frac{3}{2}(\frac{p}{q}) \notin \langle \frac{p}{q} \rangle$. Hence \mathbb{Q} is not a cyclic group.*

We shall now give some examples of cyclic group.

Example 8.5. *Consider the group (\mathbb{Z}_6, \oplus_6) . Here $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$. $1 \in \mathbb{Z}_6$ is such that, $o(1) = 6$.*

$$\begin{aligned} \therefore \langle 1 \rangle &= \{n.1(\text{mod}6) | n \in \mathbb{Z}\} \\ &= \{n1 | n = 0, 1, 2, 3, 4, 5\} \\ &= \{1, 2, 3, 4, 5, 0\} \\ &= \mathbb{Z}_6. \end{aligned}$$

Thus \mathbb{Z}_6 is a cyclic group.

Example 8.6. *Consider the group (\mathbb{Z}_n, \oplus_n) . Here $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$. Since $1 + 1 + 1 + \dots + n$ times $= 0$, therefore order of 1 in $\mathbb{Z}_n = n$. Hence $\mathbb{Z}_n = \langle 1 \rangle$, so that \mathbb{Z}_n is cyclic.*

Example 8.7. Group of cube roots of unity.

In particular, for $n = 3$, $G = \{1, \omega, \omega^2\}$ where $\omega = \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3}$, is a cyclic group generated by ω . In fact it is also generated by ω^2 , because $\omega^{-1} = \omega^2$. Thus $G = \langle \omega \rangle = \langle \omega^2 \rangle$

Example 8.8. Group of 4th roots of unity. For $n = 4$, $G = \{\pm 1, \pm i\}$ and $G = \langle i \rangle = \langle -i \rangle = \{i, i^2 = -1, i^3 = -i, i^4 = 1\} = G$.

Example 8.9. Group of n^{th} roots of unity is a cyclic group.

If $G = \{\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} | k = 0, 1, 2, 3, \dots, n-1\}$, then G is a multiplication group of n^{th} roots of unity. We shall show that G is cyclic.

Let $a = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$, then $a \in G$. Also, by De Moivre's Theorem:

$$a^k = (\cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n})^k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \text{ for } k = 1, 2, \dots, n-1.$$

Thus $G = \{a, a^2, a^3, \dots, a^{n-1}, a^n = 1\} = \langle a \rangle$. Hence G is a cyclic group generated by a .

Example 8.10. Consider the group $U(14)$ under multiplication modulo 14, then $U(14) = \{1, 3, 5, 9, 11, 13\}$. Observe that

$$3^2 = 9, 3^3 = 13, 3^4 = 11, 3^5 = 5, 3^6 = 1.$$

Therefore order of 3 is 6. Since $6 = o(U(14))$ we get $U(14) = \langle 3 \rangle$. Thus $U(14)$ is a cyclic group.

However, in general $U(n)$ is not cyclic as can be seen from the following example.

Example 8.11. Consider the group $U(8)$ under multiplication modulo 8. We know that

$$U(8) = \{1, 3, 5, 7\}. \text{ Thus } o(U(8)) = 4.$$

If $U(8)$ were to be a cyclic group, it would have an element of order 4. But none of the elements is of order 4 as $o(1) = 1, o(3) = o(5) = o(7) = 2$. Thus $U(8)$ is not a cyclic group.

Example 8.12. Show that $U(20) \neq \langle k \rangle$ for any k in $U(20)$ and hence deduce that $U(20)$ is not cyclic. We find that

$$U(20) = \{1, 3, 7, 9, 11, 13, 17, 19\}.$$

$$o(U(20)) = 8.$$

$$o(1) = 1, o(3) = 4, o(7) = 4, o(9) = 2, o(11) = 2, o(13) = 4, o(17) = 4, o(19) = 2.$$

Thus none of the elements of $U(20)$ is of order 8. Hence $U(20)$ is not a cyclic group. Thus $U(20) \neq \langle k \rangle$ for any $k \in U(20)$.

8.2 Description of Cyclic Groups

We now describe cyclic groups with regards to the number of elements it has, i.e. whether it is finite or infinite. The following theorem gives a complete description of cyclic groups, in terms of the order of the group or the order of its generator.

Theorem 8.3. Let G be a cyclic group generated by a , then

(i) G is infinite if and only if a is of infinite order, then $G = \{a^n | n \in \mathbb{Z}\}$

(ii) G is of finite order n if and only if $o(a) = n$, and

$$G = \{a, a^2, \dots, a^{n-1}, a^n = e\}.$$

Proof: Let $G = \langle a \rangle$. Then

- (i) G is infinite \Leftrightarrow For every positive integer n , $a^n \neq e \Leftrightarrow a$ is of infinite order.
(ii) Follows from Theorem 7.21. □

Remark 8.2. If G is a cyclic group of order n , and b is any element of G of order n , then $G = \langle b \rangle$.

If a has infinite order, then multiplication in $\langle a \rangle$ works in the same way as addition in \mathbb{Z} , because $a^i \cdot a^j = a^{i+j}$ for all $i, j \in \mathbb{Z}$. If a has finite order n , then the elements of $\langle a \rangle$ are multiplied by adding the powers of a modulo n , that is, $a^i a^j = a^{i \oplus_n j}$.

For these reasons, there are essentially two cyclic groups \mathbb{Z} and \mathbb{Z}_n . What is meant by this is that, although there may be different sets $\{a^n | n \in \mathbb{Z}\}$, there is only one way to operate on these sets, depending upon the order of a . Algebraists do not really care what the elements of a set are; they care only about the way the elements of the set can be combined.

Example 8.13. Consider the set G of rotations of a regular n -gon, then $G = \{R_0, R_1, R_2, \dots, R_{n-1}\}$ where R_k is the rotation through an angle of $\frac{2k\pi}{n}$. Then $R_k = R_1^k$ so that $G = \{R_1, R_1^2, \dots, R_1^{n-1}, R_1^n = R_0\}$. Hence G is a cyclic group of order n . In fact, $R_k R_l = R_m$, where $m = k \oplus_n l$. Essentially we can say that G is nothing but the group (\mathbb{Z}_n, \oplus_n) .

Example 8.14. If G is the cyclic group of n^{th} roots of unity. Then $G = \langle a \rangle$, where $a = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$ and $a^r a^s = a^{r \oplus_n s}$. Thus G is also essentially (\mathbb{Z}_n, \oplus_n) .

Example 8.15. Let m be a fixed positive integer, then $(m\mathbb{Z}, +)$ is a group. $m\mathbb{Z} = \{mz | z \in \mathbb{Z}\} = \langle m \rangle$. Hence it is a cyclic group. If $x, y \in m\mathbb{Z}$, then $x = mi, y = mj$ for some $i, j \in \mathbb{Z}$. We get $x + y = mi + mj = m(i + j)$, so the operation in $m\mathbb{Z}$ works in the same way as the operation in \mathbb{Z} .

Example 8.16. Let $H = \langle A \rangle$ where $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Then $H = \{A^0, A^{\pm 1}, A^{\pm 2}, \dots\} = \{I, A^{\pm 1}, A^{\pm 2}, \dots\}$. Since $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$,

$$\therefore H = \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z} \right\}.$$

If $M, N \in H$, where $M = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$, $N = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$, then $M = A^m, N = A^n$ and

$$MN = \begin{pmatrix} 1 & m+n \\ 0 & 1 \end{pmatrix} = A^{m+n}.$$

Thus the multiplication operation in H is similar to addition in \mathbb{Z} . Hence we can say that H and \mathbb{Z} behave in the same way in some sense.

Problem 8.1. Consider the set $S = \{4, 8, 12, 16\}$. Show that (S, \odot_{20}) is a group by constructing its multiplication table. What is the identity element? Is this group cyclic?

Solution: The multiplication table is

\odot_{20}	4	8	12	16
4	16	12	8	4
8	12	4	16	8
12	8	16	4	12
16	4	8	12	16

The table shows that 16 is the identity element, and (S, \odot_{20}) is a group. Observe that $4^2 = 16$, $8^2 = 4$, $8^3 = 12$, $8^4 = 16$. Since 8 is an element of order $4 = o(S)$, therefore $S = \langle 8 \rangle$, so that S is cyclic of order 4 and 8 is its generator.

Problem 8.2. Prove that $\mathcal{U}(2^n)$ for $n \geq 3$ is not cyclic.

Solution: Step 1

$$\begin{aligned} \mathcal{U}(2^n) &= \{k \in \mathbb{N} \mid k \text{ is odd, } k < 2^n\} \\ &= \{1, 3, 5, \dots, 2^n - 1\} \\ \text{Therefore } o(\mathcal{U}(2^n)) &= \frac{1}{2} \times 2^n = 2^{n-1}. \end{aligned}$$

If $\mathcal{U}(2^n)$ is cyclic, then it must contain an element of order 2^{n-1} . Suppose $x \in \mathcal{U}(2^n)$ is an element of order 2^{n-1} , then $x^k \neq 1 \pmod{2^n}$ for any $k < 2^{n-1}$.

Step 2 We prove that for every odd integer a , $a^{2^{n-2}} \equiv 1 \pmod{2^n}$ for $n \geq 3$. This result will be proved by induction on n . For $n = 3$, we need to prove that $a^2 \equiv 1 \pmod{8}$. Since a is odd, say $a = 2k - 1$ for some $k \in \mathbb{N}$, then

$$\begin{aligned} a^2 - 1 &= (2k - 1)^2 - 1 \\ &= 4k(k - 1) \\ &= 8 \frac{k(k - 1)}{2} \\ &\equiv 0 \pmod{8} \quad (\text{since } \frac{k(k - 1)}{2} \in \mathbb{Z} \ \forall k \in \mathbb{N}) \\ \therefore a^2 &\equiv 1 \pmod{8}. \end{aligned} \tag{8.1}$$

Thus the result holds true for $n = 3$.

Now let us assume that the result holds for $n = k \geq 3$, that is $a^{2^{k-2}} \equiv 1 \pmod{2^k}$. Hence, observe that

$$\begin{aligned} a^{2^{k-2}} - 1 &= 2^k m \quad \text{for some } m \in \mathbb{Z}. \\ a^{2^{(k+1)-2}} - 1 &= a^{2^{k-1}} - 1 \\ &= (a^{2^{k-2}} - 1)(a^{2^{k-2}} + 1) \\ &= 2^k m (a^{2^{k-2}} + 1) \\ &= 2^k m 2l \end{aligned} \tag{8.2}$$

Since a odd $\Rightarrow a^{2^{k-2}} + 1$ is even $\Rightarrow a^{2^{k-2}} + 1 = 2l$ for some $l \in \mathbb{Z}$. Thus

$$\begin{aligned} a^{2^{(k+1)-2}} - 1 &= 2^{k+1} lm \\ a^{2^{(k+1)-2}} &\equiv 1 \pmod{2^{k+1}}. \end{aligned}$$

Hence the result holds for $n = k + 1$. By the principle of induction we have

$$a^{2^{n-2}} \equiv 1 \pmod{2^n} \quad \text{for } n \geq 3.$$

$\therefore o(a) | 2^{n-2} \Rightarrow o(a) < 2^{n-1}$. $\mathcal{U}(2^n)$ has no elements of order 2^{n-1} . Thus $\mathcal{U}(2^n)$ is not cyclic for any $n \geq 3$.

8.3 Exercise

1. Prove that a non Abelian group cannot be cyclic.
2. Let $S = \{3, 6, 9, 12\}$. Show that (S, \odot_{15}) is a group by constructing its multiplication table. Is it cyclic? What are the generators?
3. Let $S = \{7, 35, 49, 77\}$. Show that (S, \odot_{84}) is a group by constructing the multiplication table. What is the identity element? Is the group cyclic? If Yes, find its generators.
4. If a cyclic group has an element of infinite order, how many elements (other than identity) of finite order does it have?
5. Show that the group of positive rational numbers under multiplication is not cyclic.
6. Which of the groups $\mathcal{U}(n)$ for $n = 7, 10, 13, 14, 15, 16$ are cyclic?
7. Prove that $\left\{ \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} : m \in \mathbb{Z} \right\}$ is a cyclic group.
8. Prove that $V_4 = \{e, a, b, ab\}$ where $a^2 = b^2 = e$, $ab = ba$ is not cyclic.

8.4 Generators of a Cyclic Group

It was observed in the previous section, that a cyclic group can be generated by more than one of its elements. Can we find all the generators of a cyclic group without finding the order of the elements? Moreover without finding the generators as such, is it possible to know how many generators a given cyclic group can have? This is precisely our object of study in this chapter.

Theorem 8.4. (*Generators of an infinite cyclic group*) *An infinite cyclic group generated by a has precisely two generators namely a and a^{-1} .*

Proof: Let $G = \langle a \rangle$ be an infinite cyclic group, then a is of infinite order by Theorem 8.3. If $G = \langle a^k \rangle$ for some $k \in \mathbb{Z}$, then $a \in G$
 $\Rightarrow a = (a^k)^m$, for some $m \in \mathbb{Z}$.
 $\Rightarrow a = a^{km}$
 $\Rightarrow km = 1$
 $\Rightarrow k = m = 1$ or $k = m = -1$. (Since $k, m \in \mathbb{Z}$)
 Thus the only generators of G are a and a^{-1} . □

If G is an additive infinite cyclic group, then the above theorem reads as:

Theorem 8.5. *An infinite additive cyclic group generated by a has precisely two generators a and $-a$.*

Theorem 8.6. *(Generators of a finite cyclic group) Let G be a cyclic group of order n generated by $a \in G$, then $a^k \in G$ is a generator of G if and only if k and n are coprime.*

Proof: $G = \langle a \rangle$ and $o(G) = n$. Suppose that a^k is a generator of G , then $G = \langle a^k \rangle = \{(a^k)^0 = e, a^k, (a^k)^2, \dots, (a^k)^{n-1}\}$. Since $a \in G$, $a = (a^k)^m$ for some $m = 0, 1, 2, \dots, n-1$. But then $a = a^{km}$ implies that n divides $km - 1$, i.e., $km - 1 = nq$ for some $q \in \mathbb{Z}$ or that $km + (-q)n = 1$.

Hence k and n are coprime. We shall prove that $\langle a^k \rangle = G$. Since k and n are coprime, therefore there exists integers m and t such that $kt + mn = 1$, then $a = a^1 = a^{kt+mn} = a^{kt}(a^n)^m = (a^{kt})e = (a^k)^t$. Thus a and so every power of a can be expressed as a power of a^k . Hence every element of G can be expressed as a power of a^k . Thus $G = \langle a^k \rangle$. \square

Let us recall the definition of Euler ϕ function.

Definition 8.2. *If n is a natural number, then we define $\phi(1) = 1$, and for $n > 1$,*

$\phi(n)$ = *number of positive integers less than n and coprime to n .*

Thus we see

$$\phi(2) = o(\{1\}) = 1, \phi(3) = o(\{1, 2\}) = 2, \phi(4) = o(\{1, 3\}) = 2.$$

(i) $\phi(p) = p - 1$, if p is prime.

(ii) $o(\mathcal{U}(n)) = \phi(n)$, for $n > 1$.

Thus the above theorem can be restated as:

“The number of generators of a finite cyclic group $\langle a \rangle$ of order n is $\phi(n)$ and the generators are a^k , where $k \in \mathcal{U}(n)$.” The beauty of this result lies in the fact that by knowing the order of a finite cyclic group, we can find the number of generators.

Theorem 8.7. *Let $(G, +)$ be a cyclic group of order n generated by $a \in G$, then $ka \in G$ is a generator of G if and only if k and n are coprime.*

This theorem helps us to find all the generators of \mathbb{Z}_n . Using the fact that \mathbb{Z}_n is a cyclic group of order n , generated by 1 under addition modulo n , we get

Corollary 8.8. *An integer $k \in \mathbb{Z}_n$ is a generator of \mathbb{Z}_n if and only if $\gcd(k, n) = 1$.*

Remark 8.3. *The generators of \mathbb{Z}_n are precisely the elements of $\mathcal{U}(n)$.*

Example 8.17. *What are all the generators of (\mathbb{Z}_8, \oplus_8) . If k is a generator of \mathbb{Z}_8 , then $\gcd(k, 8) = 1$. Thus the generators are 1, 3, 5 and 7.*

Note that these precisely are all the elements of $\mathcal{U}(8)$. Similarly, the generators of $(\mathbb{Z}_{20}, \oplus_{20})$ precisely are also all the elements of $\mathcal{U}(20)$ namely 1, 3, 7, 9, 11, 13, 17, 19.

Example 8.18. *Let G be a cyclic group of order 12 generated by a . What are all the generators of G ? Here $G = \langle a \rangle$, $o(G) = 12$. Thus $o(a) = 12$. Now a^k is a generator of G iff $(k, 12) = 1$, $1 \leq k < 12$, therefore $k = 1, 5, 7, 11$. Hence the generators of G are a, a^5, a^7, a^{11} .*

Problem 8.3. Let $S = \{4, 8, 12, 16\}$. Prove that (S, \odot_{20}) is a cyclic group and find all its generators.

Solution: We know that (S, \odot_{20}) is a cyclic group and $S = \langle 8 \rangle$ has been proved in an earlier problem. We shall find all the generators of S . Since $o(S) = 4$, by Theorem 8.6 8^k is a generator of S , iff $(k, 4) = 1$. Hence $k = 1, 3$. Hence the generators of S are $8^1, 8^3$, i.e. 8, 12.

Problem 8.4. Consider the group $(\langle a \rangle, \oplus_{12})$ with $a = 2$. Then $\langle 2 \rangle = \{0, 2, 4, 6, 8, 10\}$, so that $\langle 2 \rangle$ is a group of order 6. What are all the generators of $\langle 2 \rangle$?

Solution: The other generators are ka where $\gcd(k, 6) = 1$. Thus $k = 5$. $\langle 5a \rangle = \langle 10 \rangle = \{0, 10, 8, 6, 4, 2\} = \{0, 2, 4, 6, 8, 10\} = \langle 2 \rangle$. Thus 2 and 10 are the only generators of $\langle 2 \rangle$.

Problem 8.5. If G is a finite cyclic group with more than 1 element then G must have an element of prime order.

Solution: Let $G = \langle a \rangle$ and let $o(a) = n$. Let p be a prime such that p divides n , then $n = pk$ for some positive integer k . Let $x = a^k$. Then $x \in G$ and $x^p = e$. $o(x)|p \Rightarrow o(x) = 1$ or p . But $o(x) = 1 \Rightarrow a^k = e$. But $a^k \neq e$ for $k < n$. $\therefore o(x) = p$. Hence G has an element of order p .

Problem 8.6. Find the number of generators and all the generators of $(9\mathbb{Z}_{24}, \oplus_{24})$.

Solution: Observe that $9\mathbb{Z}_{24} = \{0, 9, 18, 3, 12, 21, 6, 15\} = \{0, 3, 6, 9, 12, 15, 18, 21\}$. Thus $(9\mathbb{Z}_{24}, \oplus_{24})$ is a group of order 8, and $o(9) = 8$. Thus $(9\mathbb{Z}_{24}) = \langle 9 \rangle$. The number of generators $= o(\mathcal{U}(8)) = \phi(8) = o(\{1, 3, 5, 7\}) = 4$. The generators are $9k \pmod{24} | k \in \mathcal{U}(8)$ i.e. $9, 27 \pmod{24}, 45 \pmod{24}, 63 \pmod{24}$ i.e. 9, 3, 21, 15.

Problem 8.7. Show that $\mathcal{U}(14)$ is cyclic. Find all the generators.

Solution: $\mathcal{U}(14) = \{1, 3, 5, 9, 11, 13\}$, $o(\mathcal{U}(14)) = 6$. Verify that $\mathcal{U}(14) = \langle 3 \rangle$. Thus $\mathcal{U}(14)$ is a cyclic group of order 6, with 3 as a generator. The number of generators $= \phi(6) = o(\mathcal{U}(6)) = o(\{1, 5\}) = 2$. The generators are 3^k where $k \in \mathcal{U}(6)$ i.e. $3^1, 3^5 \pmod{14}$, i.e. 3, 5.

8.5 Exercise

1. Find all the generators of \mathbb{Z} .
2. Find all the generators of $(\mathbb{Z}_{10}, \oplus_{10})$.
3. If a is an element of infinite order of a group G , then how many generators does $\langle a \rangle$ have? What are they?
4. Find the number of generators and all the generators of the following cyclic groups.
 - (i) $(4\mathbb{Z}_{10}, \oplus_{10})$
 - (ii) $(2\mathbb{Z}_{12}, \oplus_{12})$
 - (iii) $(6\mathbb{Z}_{20}, \oplus_{20})$
 - (iv) $(3\mathbb{Z}_{24}, \oplus_{24})$
 - (v) $(5\mathbb{Z}_{35}, \oplus_{35})$

5. Find the number of generators and all the generators of the following cyclic groups.
- (i) $\mathcal{U}(5)$
 - (ii) $\mathcal{U}(9)$
 - (iii) $\mathcal{U}(10)$
 - (iv) $\mathcal{U}(18)$
 - (v) $\mathcal{U}(22)$
 - (vi) $\mathcal{U}(25)$
6. If $G = \langle x \rangle$ is a cyclic group of order n . Find the number of generators and all the generators of G , when
- (i) $x = a, n = 8$
 - (ii) $x = b, n = 20$
7. If G is a finite group of order > 1 , then G has an element of prime order.
8. Prove that \mathbb{Z}_n has an even number of generators if $n > 2$.
9. Let G be a cyclic group of order 105. Find all generators of subgroups of order
- (i) 15 (ii) 21 (iii) 35.
10. On a circular track there are 20 stations numbered 1 to 20 on which trains run in one direction only. All trains start from station number 20. There are 3 types of trains.
- Fast train: It stops at every alternate station ie. at 20, 2, 4, ... It stops for 4 minutes at every station and takes 8 minutes to travel from one stoppage to another.
- Express train: It stops at every third station ie. at 20, 3, 6, ... It stops for 3 minutes at every station and takes 10 minutes to travel from one stoppage to another.
- Super fast train: It stops at every sixth station ie. at 20, 6, 12, ... It stops for 2 minutes at every station and takes 15 minutes to travel from one stoppage to another.
- Now answers the following:
- (i) Swati wants to go to station 17. She boards the express train from station 20. Will she be able to reach her destination? If yes, after how long?
 - (ii) Keerti has to go to station 10 and she boards a super fast train from station 20. After how long will she reach her destination?
 - (iii) Was it better for Swati to board the super fast train? Why or why not?
 - (iv) Shruti boarded the super fast train from station 2 and has to go to station 10. How long did it take for her?
 - (v) Would it have been better to catch the fast train for Shruti? If yes, why?
11. Can you solve the above problem using the concept of cyclic groups?

8.6 Subgroups of Cyclic Groups

While dealing with subgroups of a cyclic groups, the following natural questions arise:

- (i) Are the subgroups of a cyclic group necessarily cyclic?
- (ii) Does there exists a subgroup of a given order?
- (iii) If the answer to (ii) is yes, how many subgroups of a given order are there?
- (iv) How many distinct subgroups are there?

We shall answer these questions one by one. The answer to (i) is in the affirmative as given in the following theorem.

Theorem 8.9. *A subgroup of a cyclic group is cyclic.*

Proof: Let G be a cyclic group, then $G = \langle a \rangle = \{a^n | n \in \mathbb{Z}\}$ for some $a \in G$. Let H be a subgroup of G . If $G = \{e\}$ then $H = \{e\} = \langle e \rangle$. If $G \neq \{e\}$, then $G = \langle a \rangle$ for some $e \neq a \in G$. Two cases arise:

Case 1. $H = \{e\}$. In this case $H = \langle e \rangle$, hence that H is cyclic.

Case 2. $H \neq \{e\}$, then $a^n \in H$ for some $0 \neq n \in \mathbb{Z}$. Since H is a subgroup, $\therefore a^{-n} \in H$. Of n and $-n$, one of them is positive. Hence we have $a^k \in H$ for some $k \in \mathbb{N}$. Let m be the least positive integer such that $a^m \in H$. We claim that $H = \langle a^m \rangle$. Clearly $\langle a^m \rangle \subseteq H$ as $a^m \in H$, and $\langle a^m \rangle$ is the smallest subgroup of G containing a^m .

Conversely if $b \in H$, then $b \in G$ and $b = a^n$ for some $n \in \mathbb{Z}$. By division algorithm, there exist integers $q, r \in \mathbb{Z}$ such that $n = mq + r$, $0 \leq r < m$. Then $a^n = a^{mq+r} = (a^m)^q a^r$.

So $a^r = (a^n)(a^m)^{-q}$. Since $a^n, a^m \in H$ and H is a subgroup, therefore $(a^n)(a^m)^{-q} \in H$. Hence, $a^r \in H$. this is not possible for $0 < r < m$ since m is the least positive integer such that $a^m \in H$. $\therefore r = 0$. Thus $n = mq$ and $b = a^n = (a^m)^q \in \langle a^m \rangle$. Hence $H \subseteq \langle a^m \rangle$. Combining we get $H = \langle a^m \rangle$. \square

The above theorem not only tells us that every subgroups of a cyclic group is cyclic but it also gives us a method to obtain a generator of that cyclic subgroup.

Corollary 8.10. *The subgroups of the group of integers \mathbb{Z} (w.r.t. addition) are precisely the groups $(n\mathbb{Z}, +)$, where $n \in \mathbb{Z}^+$.*

Proof: \mathbb{Z} is a cyclic group under addition generated by 1. If H is a subgroup of \mathbb{Z} , then H is a cyclic subgroup of \mathbb{Z} . If $H = \{0\}$, then it is of the form $n\mathbb{Z}$, where $n = 0$. If $H \neq \{0\}$, and let n be the least positive integer in H . Then H is generated by n . For, if $m \in H$, then by Euclidean algorithm, there exist integers q and r such that $m = nq + r$, $0 \leq r < n$.

If $r > 0$, then $r = m - nq \in H$, a contradiction.

Hence $r = 0$. But then $m = nq \in \langle n \rangle$ or that $H \subseteq \langle n \rangle$. Since $n \in H$ we get $\langle n \rangle \subseteq H$. Thus $H = \langle n \rangle$. Thus $H = n\mathbb{Z}$. Moreover, the set $n\mathbb{Z}$ of all multiples of n is a subgroup of \mathbb{Z} . Hence the subgroups of \mathbb{Z} are precisely $n\mathbb{Z}$, for $n \in \mathbb{Z}^+$. \square

8.7 Subgroups of Infinite Cyclic Groups

Theorem 8.11. *Every subgroup of an infinite cyclic group is infinite cyclic.*

Proof: Let $G = \langle a \rangle$ be an infinite cyclic group. Then a is of infinite order. Further, let H be a subgroup of G . Since every subgroup of a cyclic group is cyclic, we get $H = \langle a^m \rangle$, for some $a^m \in H$. Since, a is of infinite order, we get a^m to be of infinite order and thus $\langle a^m \rangle$ is an infinite cyclic group. Hence H is an infinite cyclic subgroup of G . \square

Theorem 8.12. *If $G = \langle a \rangle$ is an infinite cyclic group, then*

- (i) $\langle a^i \rangle \subseteq \langle a^j \rangle$ if and only if j divides i .
(ii) $\langle a^i \rangle = \langle a^j \rangle$ if and only if $j = \pm i$.

Proof: Since $\langle a \rangle$ is infinite cyclic, therefore $o(a)$ is infinite.

(i) Let $\langle a^i \rangle \subseteq \langle a^j \rangle$, then $a^i \in \langle a^j \rangle \subseteq \langle a^j \rangle$. Hence $a^i = (a^j)^k$ for some $k \in \mathbb{Z}$, i.e., $a^i = a^{jk}$. Since $o(a)$ is infinite, so no two distinct powers of a are equal. Thus $i = jk$ which implies that j divides i .

Conversely, let j divide i . Then $i = tj$ for some $t \in \mathbb{Z}$. This gives that $a^i = a^{tj} = (a^j)^t \in \langle a^j \rangle$.

Hence $a^i \in \langle a^j \rangle$, so that $\langle a^i \rangle \subseteq \langle a^j \rangle$.

(ii) Suppose, $\langle a^i \rangle = \langle a^j \rangle$. Then $\langle a^i \rangle \subseteq \langle a^j \rangle$ and $\langle a^j \rangle \subseteq \langle a^i \rangle$. By Part (i), we get that j divides i and i divides j . But then $j = \pm i$.

Conversely, let $j = \pm i$. Then j divides i and i divides j . Again by part (i) this implies that by $\langle a^i \rangle \subseteq \langle a^j \rangle$ and $\langle a^j \rangle \subseteq \langle a^i \rangle$. Hence $\langle a^i \rangle = \langle a^j \rangle$. \square

Remark 8.4. *From the Theorem 8.11, for the infinite group $G = \langle a \rangle$, it follows that*

i for every positive integer n , $\langle a^n \rangle$ is an infinite cyclic subgroup of G .

ii if $m, n \in \mathbb{Z}^+$, $m \neq n$ then $\langle a^m \rangle \neq \langle a^n \rangle$.

iii if $m \in \mathbb{Z}^+$ then $\langle a^m \rangle \supseteq \langle a^{2m} \rangle \supseteq \langle a^{4m} \rangle \supseteq \langle a^{8m} \rangle \supseteq \dots$

Example 8.19. *Suppose a is an element of infinite order, then what are all the generators of $\langle a^3 \rangle$?*

Since $o(a)$ is infinite therefore $o(a^3)$ is also infinite. Hence the generators of $G = \langle a^3 \rangle$ are a^3 and $(a^3)^{-1}$ i.e. a^3 and a^{-3} .

Example 8.20. *$(\mathbb{Z}, +)$ is an infinite cyclic group then any subgroup H of \mathbb{Z} is of the form $H = m\mathbb{Z}$ for some $m \in \mathbb{Z}$. That is H is generated by m or $-m$.*

8.8 Subgroups of Finite Cyclic Groups

Before coming to the results of this section, let us find out all subgroups of a finite cyclic group. For example, let $G = \langle a \rangle$ is cyclic group of order 20. Since subgroup of a cyclic group is cyclic, therefore all the subgroups of G will be of the form $\langle a^k \rangle$ for some non negative integer k .

Observe the following: $H_0 = \langle e \rangle = \{e\}$

$$H_1 = \langle a \rangle = \{e, a, \dots, a^{19}\} = \langle a^{-1} \rangle = \langle a^{19} \rangle = G$$

$$H_2 = \langle a^2 \rangle = \{e, a^2, a^4, \dots, a^{18}\} = \langle (a^2)^{-1} \rangle = \langle a^{18} \rangle$$

$$H_3 = \langle a^3 \rangle = \{e, a^3, a^6, \dots, (a^3)^{19} = a^{17}\} = G = \langle (a^3)^{-1} \rangle = \langle a^{17} \rangle$$

$$H_4 = \langle a^4 \rangle = \{e, a^4, a^8, a^{12}, a^{16}\} = \langle (a^4)^{-1} \rangle = \langle a^{16} \rangle$$

$$H_5 = \langle a^5 \rangle = \{e, a^5, a^{10}, a^{15}\} = \langle (a^5)^{-1} \rangle = \langle a^{15} \rangle$$

$$\begin{aligned}
H_6 = \langle a^6 \rangle &= \{e, a^6, a^{12}, a^{18}, a^4, a^{10}, a^{16}, a^2, a^8, a^{14}\} = \{e, a^2, a^4, \dots, a^{18}\} \\
&= \langle (a^6)^{-1} \rangle = \langle a^{14} \rangle \\
H_7 = \langle a^7 \rangle &= \{e, a^7, \dots, (a^7)^{19} = a^{13}\} = G = \langle (a^7)^{-1} \rangle = \langle a^{13} \rangle \\
H_8 = \langle a^8 \rangle &= \{e, a^8, a^{16}, a^4, a^{12}\} = \{e, a^4, a^8, a^{12}, a^{16}\} = \langle (a^8)^{-1} \rangle = \langle a^{12} \rangle \\
H_9 = \langle a^9 \rangle &= \{e, a^9, a^{18}, \dots, (a^9)^{19} = a^{11}\} = G = \langle (a^9)^{-1} \rangle = \langle a^{11} \rangle \\
H_{10} = \langle a^{10} \rangle &= \{e, a^{10}\} = \langle (a^{10})^{-1} \rangle = \langle a^{10} \rangle.
\end{aligned}$$

Since subgroups of a cyclic group are cyclic, therefore these are the only subgroups of G .

We observe that $o(H_0) = 1$.

Also $H_1 = H_3 = H_7 = H_9 = G$ and each is of order 20.

$H_2 = H_6$ and each is of order 10.

$H_4 = H_8$ and each is of order 5.

$o(H_5) = 4$.

Note that the orders of the distinct subgroups are 1, 2, 4, 5, 10 and 20. There is a unique subgroup of each of these orders. These orders are precisely the divisors of 20. This leads us to believe that if G is a finite cyclic group of order n , then for each divisor m of n , there exists a unique cyclic subgroup of order m . Moreover, it is generated by $a^{\frac{n}{m}}$. The next theorem confirms this.

Theorem 8.13. *Let G be a cyclic group of order n .*

- (i) *If H is a subgroup of G , then $o(H)$ divides $o(G)$.*
- (ii) *Conversely, if m is a divisor of n , then G has exactly one subgroup of order m .*

Proof: Let $G = \langle a \rangle$ and $o(G) = n$. $\therefore o(a) = n$ and $a^n = e$.

- (i) Let H be a subgroup of G . Since a subgroup of a cyclic group is cyclic. Let $H = \langle a^m \rangle$ for some m , $0 \leq m \leq n-1$. Then $o(H) = o(a^m)$. Since $(a^m)^n = a^{mn} = (a^n)^m = e$. Thus by Theorem 7.16 $o(a^m) \mid n$ i.e. $o(H) \mid o(G)$.
- (ii) Let $m \mid n$. Then $n = mk$ for some $k \in \mathbb{Z}$. Consider $H = \langle a^{n/m} \rangle = \langle a^k \rangle$. Now $(a^k)^t = a^{kt} \neq e$ for any $t < m$. This proves that $o(a^k) = m$. Hence H is a subgroup of order m . The subgroup H is unique. For, let T be another subgroup of order m , then T is cyclic. Let $T = \langle a^l \rangle$, where l is the least positive integer such that $a^l \in T$. By division algorithm, there exist integers q and r such that $n = lq + r$, $0 \leq r < l$. Now $e = a^n = a^{lq+r} = a^{lq}a^r$. Thus $a^r = a^{-lq} = (a^l)^{-q} \in T$. The choice of l forces r to be 0. So $n = lq$. Thus $o(T) = o(a^l) = \frac{n}{l} = q$. But $o(T) = m$, So $q = m$. Hence $l = \frac{n}{q} = \frac{n}{m}$ and therefore $T = \langle a^{\frac{n}{m}} \rangle = H$.

Thus we have proved that if $m \mid n$, then there exists a unique subgroup of order m generated by $a^{\frac{n}{m}}$. \square

The above theorem can be applied to obtain the subgroups of (\mathbb{Z}_n, \oplus_n) .

Corollary 8.14. *If m is a divisor of n , then there exists a unique subgroup of \mathbb{Z}_n of order m , generated by $\frac{n}{m}$ namely $(\frac{n}{m})\mathbb{Z}_n$.*

It will be proved that part (i) of the Theorem 8.13 proved above for cyclic groups, also holds for finite groups. That is, if H is a subgroup of a finite group G then $o(H)$ divides $o(G)$. This result is known as “Lagrange’s theorem”.

8.9 Number of Generators

In the examples discussed so far, we have seen that some cyclic groups have exactly one generator, whereas others have two or more generators. Is there a way to find all the generators of a subgroup of a cyclic group? This is answered by the following Theorems.

Theorem 8.15. *Let G be a cyclic group of order n generated by a and let $d \mid n$. If H is a subgroup of G of order d , then*

- (i) *the number of generators of H is $\phi(d) = o(\mathcal{U}(d))$.*
- (ii) *every generator of H is a^{km} , $k \in \mathcal{U}(d)$ and $m = \frac{n}{d}$.*

Proof: Given $o(G) = n$, and $G = \langle a \rangle$. Since $d \mid n$, so $n = md$ for some $m \in \mathbb{Z}$. If H is a subgroup of G of order d , then H is cyclic of order d . Moreover, $H = \langle a^m \rangle$. Further, a^m is of order d . Let $b = a^m$. Then $H = \langle b \rangle$, and H is of order d .

Now, the number of generators of $H = \phi(o(H)) = \phi(d) = o(\mathcal{U}(d))$. This proves part (i).

Further, the generators of H are precisely $b^k = (a^m)^k = a^{mk}$, $k \in \mathcal{U}(d)$, and $m = \frac{n}{d}$. This proves part (ii). Hence the theorem. \square

Remark 8.5. *If G is a cyclic group of order n generated by a . If $x \in G$, then x is an element of order $d \Leftrightarrow x$ is a generator of a subgroup of order d in G . Therefore, the number of elements of order d is $\phi(d)$. Moreover, these elements are x^{km} , where $k \in \mathcal{U}(d)$ and $m = \frac{n}{d}$. Thus the above theorem gives us the elements of order d and the number of such elements.*

Example 8.21. *List all the elements of order 10 in \mathbb{Z}_{40} . Here $n = 40, d = 10, d \mid n$.*

The number of elements of order 10 is nothing but the number of generators of a subgroup of order 10, which is $\phi(10)$. But $\phi(10) = 4$. Let $H = \langle b \rangle$ be a subgroup of order 10. The generators of H are $b, 3b, 7b, 9b$ by Theorem 8.7. One such element of order 10 in \mathbb{Z}_{40} is $\frac{40}{10} = 4$. Therefore the generators are 4, 12, 28, 36.

Example 8.22. *If G is a cyclic group of order 24 generated by a , find all the generators of a subgroup H of order 8.*

Here $n = 24, m = 8$, and $d = \frac{n}{m} = 3$. Then, by Theorem 8.15(ii) a^3 is a generator of H . Thus others generators of H are a^{3k} , $1 \leq k < 8$, $k \in \mathcal{U}(8)$. Therefore $k = 1, 3, 5, 7$. Hence all the generators of H are a^3, a^9, a^{15} , and a^{21} .

Example 8.23. *Given that $\mathcal{U}(49)$ is a cyclic group having 42 elements, find the number of generators without actually finding them.*

Let $G = \mathcal{U}(49)$, $o(G) = 42$. By Theorem 8.6 the numbers of generators is $\phi(42)$. But $\phi(42) = o(\mathcal{U}(42))$.

Now, $\mathcal{U}(42) = \{1, 5, 11, 13, 17, 19, 23, 25, 29, 31, 37, 41\}$, therefore $o(\mathcal{U}(42)) = 12$. Thus there are 12 generators of $\mathcal{U}(49)$. [In fact $\mathcal{U}(49) = \langle 2 \rangle = 2^k \pmod{49}$, $k \in \mathcal{U}(42)$ are different 12 generators of $\mathcal{U}(49)$.]

Theorem 8.16. *If $G = \langle a \rangle$ is a finite cyclic group of order n , then*

- (i) *$\langle a^r \rangle = \langle a^{n-r} \rangle$.*
- (ii) *$\langle a^r \rangle \subseteq \langle a^s \rangle$ if and only if r is a multiple of $s \pmod n$.*
- (iii) *$\langle a^r \rangle = \langle a^{\gcd(n,r)} \rangle$.*
- (iv) *$\langle a^r \rangle = \langle a^s \rangle$ if and only if $\gcd(n, r) = \gcd(n, s)$.*

Proof: Since a is of order n , therefore n is the smallest positive integer such that $a^n = e$. Also, $G = \langle a \rangle$ and $o(G) = n$. Now

(i) Since a^r and a^{n-r} are the inverses of each other therefore $a^r \in \langle a^{n-r} \rangle$ and $a^{n-r} \in \langle a^r \rangle$. Hence $\langle a^r \rangle \subseteq \langle a^{n-r} \rangle$ and $\langle a^{n-r} \rangle \subseteq \langle a^r \rangle$. Thus $\langle a^r \rangle = \langle a^{n-r} \rangle$.

(ii) $\langle a^r \rangle \subseteq \langle a^s \rangle$

$$\begin{aligned} \Leftrightarrow a^r &\in \langle a^r \rangle \subseteq \langle a^s \rangle \\ \Leftrightarrow a^r &= (a^s)^m \text{ for some } m, 1 \leq m \leq n \\ \Leftrightarrow a^r &= a^{sm} \\ \Leftrightarrow a^{r-sm} &= e \\ \Leftrightarrow r-sm &\text{ is a multiple of } n \\ \Leftrightarrow r &= sm \text{ modulo } n \\ \Leftrightarrow r &\text{ is a multiple of } s \text{ modulo } n. \end{aligned}$$

(iii) Let $gcd(n, r) = d$. Then $0 < d < n$. Also $d \mid r$ and $d \mid n$. Since $d \mid r \Rightarrow \langle a^r \rangle \subseteq \langle a^d \rangle$ by (ii). Since $gcd(n, r) = d$, by Euclidean Algorithm there exists integers p, q such that $np + qr = d$.

Now $a^d = a^{np+qr} = a^{np}a^{qr} = (a^n)^p(a^r)^q = e(a^r)^q = (a^r)^q$. Therefore $a^d \in \langle a^r \rangle$ so that $\langle a^d \rangle \subseteq \langle a^r \rangle$. Hence we get $\langle a^r \rangle = \langle a^d \rangle = \langle a^{gcd(n,r)} \rangle$.

(iv) Let $d' = gcd(n, r)$ and $d = gcd(n, s)$. Then, by Theorem, 7.18 $o(\langle a^{d'} \rangle) = \frac{n}{d'}$ and $o(\langle a^d \rangle) = \frac{n}{d}$.

By part (iii) $o(\langle a^r \rangle) = o(\langle a^{d'} \rangle) = \frac{n}{d'}$. Similarly, $o(\langle a^s \rangle) = \frac{n}{d}$.

Now $\langle a^r \rangle = \langle a^s \rangle \Leftrightarrow o(\langle a^r \rangle) = o(\langle a^s \rangle) \Leftrightarrow \frac{n}{d'} = \frac{n}{d} \Leftrightarrow d = d'$. \square

Given a finite cyclic group, the above theorem helps us to find

(i) all the generators of the unique subgroup of a given order.

(ii) the order of a given subgroup $\langle a^k \rangle$ of $\langle a \rangle$. If $o(a) = n$ then

$$o(\langle a^k \rangle) = \frac{n}{gcd(n,k)} \text{ for every positive integer } k \leq n.$$

Example 8.24. Let G be a cyclic group generated by a of order 15. Compute the orders and generators of the subgroups $\langle a^3 \rangle$, $\langle a^6 \rangle$, $\langle a^8 \rangle$ and $\langle a^{10} \rangle$.

Solution: Here $n = o(G) = 15$. Observe that

$$o(\langle a^3 \rangle) = o(a^3) = \frac{15}{gcd(3,15)} = \frac{15}{3} = 5.$$

$$o(\langle a^6 \rangle) = o(a^6) = \frac{15}{gcd(6,15)} = \frac{15}{3} = 5.$$

$$o(\langle a^8 \rangle) = o(a^8) = \frac{15}{gcd(8,15)} = 15.$$

$$o(\langle a^{10} \rangle) = o(a^{10}) = \frac{15}{gcd(10,15)} = \frac{15}{5} = 3.$$

Example 8.25. In the cyclic group $(\mathbb{Z}_{24}, \oplus_{24})$, compute the orders of the subgroups $\langle 8 \rangle$, $\langle 5 \rangle$, and $\langle 9 \rangle$.

Here order of group is $n = 24$. For the element $m = 8 \in \mathbb{Z}_{24}$, we have $o(\langle 8 \rangle) = o(8) = \frac{n}{m} = \frac{24}{8} = 3 = d$ (say) number of generators of subgroup of order 3 = $\phi(3) = o(\mathcal{U}(3)) = 2$.

The generators are $k8$, $k \in \mathcal{U}(3)$.

Thus $k = 1, 2$. Therefore generators are 8, 16.

Hence $\langle 8 \rangle = \langle 16 \rangle$. Now $o(\langle 5 \rangle) = o(5) = \frac{n}{gcd(5,n)} = \frac{24}{gcd(5,24)} = 24$.

Therefore generators of $\langle 5 \rangle$ are $5k \pmod{24}$, $k \in \mathcal{U}(24)$.

Hence $k = 1, 3, 7, 11, 13, 17, 19, 23$ and generators of $\langle 5 \rangle$ are $5, 15, 11, 7, 17, 13, 23, 19$. $o(\langle 9 \rangle) = o(9) = \frac{n}{\gcd(9, n)} = \frac{24}{\gcd(9, 24)} = 8$. Therefore generators of $\langle 9 \rangle$ are $9k \pmod{24}$, $k \in \mathcal{U}(8)$ i.e. $k = 1, 3, 5, 7$.

Problem 8.8. How many subgroups does \mathbb{Z}_{20} have? List all the generators for each of these subgroups.

Solution: \mathbb{Z}_{20} is a cyclic group of order 20, generated by 1. Thus $n = o(\mathbb{Z}_{20}) = 20$, the possible divisors of 20 are 1, 2, 4, 5, 10, 20. Thus there exists a subgroup H_i of order i , for $i = 1, 2, 4, 5, 10, 20$. Subgroup of order 1 is $\{0\}$. $H_1 = \langle 0 \rangle$.

Subgroup of Order 2. Here $d = 2$, $\therefore m = \frac{n}{d} = 10$. Thus H_2 will be a subgroup generated by 10. $H_2 = \langle 10 \rangle$. The other generators of H_2 are $10k$, where $(k, 2) = 1$. The only possible value of k is 1. Hence $H_2 = \langle 10 \rangle$ is the only subgroup of \mathbb{Z}_{20} of order 2.

Subgroup of Order 4. Here $d = 4$, $\therefore m = \frac{n}{d} = \frac{20}{4} = 5$. Hence $H_4 = \langle 5 \rangle$. The generators of H_4 are $5k$, $k \in \mathcal{U}(4) = \{1, 3\}$ i.e. 5, 15. Thus $H_4 = \langle 5 \rangle = \langle 15 \rangle$.

Subgroup of Order 5. Proceeding as above $H_5 = \langle 4 \rangle$. All generators are $4k$, $k \in \mathcal{U}(5)$ i.e. 4, 8, 12, 16.

Subgroup of Order 10. $H_{10} = \langle 2 \rangle$. All generators are $2k$, $k \in \mathcal{U}(10)$ i.e. 2, 6, 14, 18.

Subgroup of Order 20. Proceeding as above $H_{20} = \langle 1 \rangle$. All generators are k , $k \in \mathcal{U}(20)$ i.e. 1, 3, 7, 9, 11, 13, 17, 19.

Summarizing:

Name of subgroup	Order of subgroup	Generators
H_1	1	0
H_2	2	10
H_4	4	5, 15
H_5	5	4, 8, 12, 16
H_{10}	10	2, 6, 14, 18
H_{20}	20	1, 3, 7, 9, 11, 13, 17, 19

Problem 8.9. If G is a cyclic group of order 24, then find a generator for $\langle a^{21} \rangle \cap \langle a^{18} \rangle$.

Solution: $o(G) = n = 24$. Hence $\langle a^{21} \rangle = \langle a^k \rangle$, where $k = \gcd(24, 21) = 3$ and $\langle a^{18} \rangle = \langle a^l \rangle$, where $l = \gcd(24, 18) = 6$. Thus $\langle a^{21} \rangle = \langle a^3 \rangle$, and $\langle a^{18} \rangle = \langle a^6 \rangle$. Further, $\langle a^{21} \rangle \cap \langle a^{18} \rangle = \langle a^3 \rangle \cap \langle a^6 \rangle = \langle a^6 \rangle$, since $\langle a^6 \rangle \subseteq \langle a^3 \rangle$.

Problem 8.10. Suppose that a cyclic group G has exactly three subgroups: G , $\{e\}$ and a subgroup of order 7. What is $o(G)$?

Solution: Let $o(G) = n$. For every divisor d of n , the group G has a subgroup of order d . But G has three subgroups of orders 1, 7 and $n \neq 7$. Since the only 3 divisors of n are 1, 7, n . Hence $n = 7^2 = 49$. Thus $o(G) = 49$.

Problem 8.11. Let x be an element of order 40 in a cyclic group G . List all the elements of $\langle x \rangle$ of order 10.

Solution: Given $x \in G$ such that $o(x) = 40$. Let $H = \langle x \rangle$, then H is a cyclic group of order 40. We have to find elements of H which are of order 10. Now, $o(x^4) = 10$ if $(x^4)^k$ is of order 10. So, $(k, 10) = 1$. Hence, $k = 1, 3, 7, 9$. Thus $x^4, (x^4)^3, (x^4)^7, (x^4)^9$ are of order 10, i.e. x^4, x^{12}, x^{28} and x^{36} are of order 10.

Problem 8.12. Determine the order of each element of D_{66} . How many elements are there of given order.

Solution: D_{66} consists of 33 rotations and 33 reflections. Since a dihedral group is not cyclic, D_{66} is not cyclic. It has a cyclic subgroup H of order 33 consisting of the 33 rotations. Each reflection is an element of order 2. Also $o(H)$ is odd, so it does not have any element of order 2.

Thus there are 33 elements of order 2. Further, the divisors of $o(H)$ are 1, 3, 11, 33. Hence,

Number of elements of order 1 = $\phi(1) = 1$

Number of elements of order 3 = $\phi(3) = 2$

Number of elements of order 11 = $\phi(11) = 10$

Number of elements of order 33 = $\phi(33) = 20$.

We can summarize as follows:

Order	Numbers of elements
1	1
2	33
3	2
11	10
33	20

8.10 Exercise

1. Suppose $G = \langle a \rangle$ and $o(a) = 20$. How many subgroups does G have? List all generators for each of these subgroups.
2. How many subgroups does $(\mathbb{Z}_{18}, \oplus_{18})$ have? List all generators for each of these subgroups.
3. Let $G = \langle a \rangle$ and let $o(a) = 28$. List all the generators of a subgroup of order 4.
4. If G is a group and $a \in G$ is of infinite order, find all the generators of $\langle a^5 \rangle$.
5. List all elements of order 8 in \mathbb{Z}_{80000} . How do you know that your list is complete?
6. Suppose G is a cyclic group of order n
 - (i) If 6 divides n , how many elements of order 6 does G have?
 - (ii) If 10 divides n , how many elements of order 10 does G have? If a is an element of order 10, what are others elements of order 10?
7. Let m and n be elements of the group \mathbb{Z} . Find a generator for the group $\langle m \rangle \cap \langle n \rangle$.

8. Let p be prime. If a group G has more than $p - 1$ elements of order p , can G be a cyclic? Justify.
9. If G is a cyclic group of order 15, find the orders of the subgroups $\langle a^9 \rangle$, $\langle a^{12} \rangle$, $\langle a^5 \rangle$, $\langle a^4 \rangle$ and $\langle a^{14} \rangle$.
10. Let a, b be elements of a group G . If $o(a) = 10$, $o(b) = 21$, show that $\langle a \rangle \cap \langle b \rangle = \{e\}$.
11. Let a, b be elements of a group G . If $o(a) = 24$, $o(b) = 10$, what are the possibilities for $o(\langle a \rangle \cap \langle b \rangle)$?

8.11 Solved Problems

Problem 8.13. Find the smallest subgroup of \mathbb{Z} containing 18, 30 and 40.

Solution: Subgroups of \mathbb{Z} are of the form $n\mathbb{Z} = \langle n \rangle = \{na \mid a \in \mathbb{Z}\}$.

Now, $\gcd(18, 30, 40) = 2$. Since $18 = 2 \times 9$, $30 = 2 \times 15$, $40 = 2 \times 20$, we get $18, 30, 40 \in \langle 2 \rangle$. Hence $\langle 2 \rangle$ is subgroup of \mathbb{Z} containing 18, 30, 40.

If $18, 30, 40 \in \langle n \rangle$, then n divides each of 18, 30 and 40 and therefore n divides $\gcd(18, 30, 40) = 2$. Thus n divides 2, so that $2 \in \langle n \rangle$ i.e. $\langle 2 \rangle \subseteq \langle n \rangle$. Thus $\langle 2 \rangle$ is the smallest subgroup of \mathbb{Z} containing 18, 30 and 40.

Remark 8.6. In \mathbb{Z} , the smallest subgroup of \mathbb{Z} containing $a_1, a_2, a_3, \dots, a_k$ is $\langle d \rangle$ where $d = \gcd(a_1, a_2, a_3, \dots, a_k)$. This subgroup is denoted by $\langle a_1, a_2, a_3, \dots, a_k \rangle$.

Problem 8.14. Every group of order 3 is cyclic.

Solution: Let G be a group such that $o(G) = 3$. Let $a \in G$, $a \neq e$. Consider $H = \langle a \rangle = \{a, a^2, \dots\}$. Two cases arise:

Case 1. $a^2 = e$. Then $H = \{a, e\}$ so that $o(H) = 2$. Thus there exists $b \in G$ such that $b \notin H$. Now $bH = \{b, ba\}$. Then $H \cap bH = \phi$, so $o(G) \geq 4$, a contradiction.

Case 2. $a^2 \neq e$. If $a^3 \neq e$ then e, a, a^2, a^3 are distinct elements of G and so again a contradiction $o(G) > 3$. So that we must have $a^3 = e$. Thus $H = \{a, a^2, e\}$, $o(H) = 3$. Now since $H \subseteq G$, G is finite and $o(H) = o(G)$ therefore $H = G$. Hence G is a cyclic group.

Problem 8.15. Prove that a group of order 4 is Abelian.

Solution: Let G be a group of order 4. If $a \in G$ and $o(a) = n > 4$ then $e, a, a^2, \dots, a^{n-1}$ are $n > 4$ distinct elements of G , which is not possible in a group of order 4. Hence order of every element of G is less than or equal to 4. Three cases arise:

Case 1. G has an element of order 4, say a , then $G = \langle a \rangle = \{a, a^2, a^3, a^4 = e\}$ and therefore G is Abelian.

Case 2. G has no element of order 4 but G has an element of order 3, say a , then a, a^2, e are 3 distinct elements of G . Let $b \in G$ be different from these 3 elements, then ab, a^2b, b are all distinct elements of G which are different from a, a^2, e . Thus G has at least 6 distinct elements which is not possible in a group of order 4. Hence G does not have an element of order 3.

Case 3. Every non-identity element is of order 2. Let $a, b \in G, a \neq e, b \neq e$, then $o(a) = o(b) = 2$. Then e, a, b, ab are 4 distinct elements of $G = \{e, a, b, ab\}$. Similarly $G = \{e, a, b, ba\}$. Hence $ab = ba$, so that G is Abelian. Hence every group of order 4 is Abelian.

Problem 8.16. *Can a group of order 4 have an element of order 3? What are the possible orders of an element of in group of order 4? Can you make a conjecture about the relation between order of a group and order of its elements?*

Solution: Let G be a group of order 4. Let if possible, G has an element say a of order 3, then $a^3 = e$. Let $H = \{e, a, a^2\}$. Then H is a subgroup of G and $H = \langle a \rangle$. Clearly $H \neq G$ (Since $o(H) = 3$ while $o(G) = 4$). Let $b \in G \setminus H$, then $bH = \{b, ba, ba^2\}$, and $bH \cap H = \phi$, as $ba^i = a^j$ for some $0 \leq i, j < 3$ implies that $b \in H$, but $b \notin H$. Also $bH \subseteq G$. Thus G has at least 6 elements namely e, a, a^2, ba, b, ba^2 which is not possible as $o(G) = 4$. This contradicts our assumption that G has an element of order 3. So that G does not have any element of order 3. If G is a group of order 4, then e is an element of order 1. G may have all non identity elements of order 2 or G may have an element of order 4, as we may have $G = \langle a \rangle$. Thus the possible orders of elements of G are 1, 2 or 4. These are all divisors of $4 = o(G)$.

We can conjecture:

Conjecture 8.1. *In a finite group G , the order of each element of G divides the order of G .*

Problem 8.17. *Let a and b be elements of a group G . If $o(a) = m, o(b) = n$ and m, n are coprime, prove that $\langle a \rangle \cap \langle b \rangle = \{e\}$.*

Solution: Let $x \in \langle a \rangle \cap \langle b \rangle$, then $x \in \langle a \rangle$ and $x \in \langle b \rangle$, so that $x = a^p$ and $x = b^q$ for some positive integers p and q . Then $o(x) = o(a^p) = \frac{m}{\gcd(p, m)}$. Thus $o(x)$ divides m . Similarly $o(x) = o(b^q) = \frac{n}{\gcd(q, n)}$. So that $o(x)$ divides n . Hence $o(x)$ divides $\gcd(m, n)$. But $\gcd(m, n) = 1$. Thus $o(x)$ divides 1 so that $o(x) = 1$. Hence $x = e$, therefore $\langle a \rangle \cap \langle b \rangle = \{e\}$.

Aliter: Let $A = \langle a \rangle, B = \langle b \rangle, A \cap B$ is a subgroup of A as well as B . Hence $o(A \cap B) \mid o(A) \Rightarrow o(A \cap B) \mid m$
Similarly $o(A \cap B) \mid n$
Hence $o(A \cap B) \mid (m, n) = 1$
 $\Rightarrow o(A \cap B) = 1$
 $\Rightarrow A \cap B = \{e\}$.

Problem 8.18. *Let G be an Abelian group of order pq , with $(p, q) = 1$. Assume that there exist $a, b \in G$, such that $o(a) = p, o(b) = q$. Show that G is cyclic.*

Solution: Since $o(a) = p, o(b) = q, \therefore a^p = e, b^q = e, (ab)^{pq} = a^{pq}b^{pq} = e$, since G is an Abelian group $\therefore o(ab) \mid pq$.
Let $o(ab) = k$, so that $(ab)^k = e$, therefore $a^k b^k = e$ (since G is Abelian)
 $\Rightarrow b^k = a^{-k}$
 $\Rightarrow b^k \in \langle a \rangle$, since $a^{-k} \in \langle a \rangle$
 $\Rightarrow (b^k)^p = e$, since $o(\langle a \rangle) = p$
 $\Rightarrow b^{kp} = e$

$$\begin{aligned} &\Rightarrow q \mid pk \\ &\Rightarrow q \mid k \text{ since } (p, q) = 1. \end{aligned}$$

Similarly $p \mid k$.

Hence, $pq \mid k$. Since $(p, q) = 1$.

$\Rightarrow o(ab) = pq$.

$\Rightarrow ab$ is an element of order pq in G .

Since order $o(G) = pq$, we get $G = \langle ab \rangle$. Hence G is cyclic.

Problem 8.19. Suppose that G is a group that has exactly one non-trivial proper subgroup. Prove that G is cyclic and $o(G) = p^2$, where p is prime.

Solution:

Step 1 We shall prove that G is cyclic. Let $e \neq a \in G$. Then $\langle a \rangle$ is a non-trivial subgroup of G . If $G = \langle a \rangle$ then G is cyclic. If $\langle a \rangle \neq G$. Let $b \in G \setminus \langle a \rangle$, then $\langle b \rangle \subsetneq G$ and $\langle b \rangle \neq \langle a \rangle$ as $b \notin \langle a \rangle$. Thus G has at least two non-trivial proper subgroups which is not possible. Hence G is cyclic.

Step 2 We shall prove that G is finite. Since G is cyclic, we can assume that $G = \langle a \rangle$. If G is infinite, then $\langle a^2 \rangle$ is a subgroup of G such that $a \notin \langle a^2 \rangle$ also $\langle a^3 \rangle$ is a subgroup of G such that $a \notin \langle a^3 \rangle$ and $a^2 \notin \langle a^3 \rangle$. Thus G has at least two non-trivial proper subgroups $\langle a^2 \rangle$ and $\langle a^3 \rangle$, which contradicts our hypothesis. Hence our assumption is wrong. Hence G must be finite. Let $o(G) = n$.

Step 3 We shall now prove that there exists only one prime p such that p divides n . Let if possible, there are two distinct primes p and q which divide n . Since $o(a) = n$, $\therefore o(a^{\frac{n}{p}}) = p < n$, and $o(a^{\frac{n}{q}}) = q < n$, so that $\langle a^{\frac{n}{p}} \rangle$ and $\langle a^{\frac{n}{q}} \rangle$ are two proper non trivial subgroups of G . This contradicts our hypothesis. Hence our assumption is wrong so there is exactly one prime p which divides n . Thus $o(G) = n = p^k$ for some k .

Step 4 We shall prove that $k = 2$. If $k = 1$ then $o(G) = p$. Let H be a proper non-trivial subgroup of G , then $H = \langle a^m \rangle$ for some, m such that $0 < m < p$. Now $o(a^m) = \frac{p}{\gcd(p, m)} = p$. Thus H is a subgroup of order p which contradicts the fact that H is a proper subgroup of G . Hence $k \neq 1$. If $k \geq 3$, then $H_1 = \langle a^p \rangle$ and $H_2 = \langle a^{p^2} \rangle$ are two non trivial proper subgroups of G . Again a contradiction. Hence $k = 2$ and $o(G) = p^2$.

Problem 8.20. Let a and b be elements of a group G . If $o(a) = 12, o(b) = 22$ and $\langle a \rangle \cap \langle b \rangle \neq \{e\}$. Prove that $a^6 = b^{11}$.

Solution: Since intersection of two subgroups is a subgroup, $\therefore H = \langle a \rangle \cap \langle b \rangle$ is a subgroup of G . Moreover, being a subgroup of $\langle a \rangle$ as well as $\langle b \rangle$, it is cyclic. Let $H = \langle c \rangle$. Since order of a subgroup of a group divides the order of the group $\therefore o(H) \mid o(\langle a \rangle)$ and $o(H) \mid o(\langle b \rangle)$ i.e. $o(\langle c \rangle) \mid 12$ and $o(\langle c \rangle) \mid 22$. Since $o(H) = o(c)$, $\therefore o(c) = 1$ or 2 . But $o(c) = o(\langle c \rangle) \neq 1$. $\therefore o(c) = 2$. Thus $c = a^6$, since $o(c) = 2$ as $c \in \langle a \rangle$. Also $c = b^{11}$, since $O(c) = 2$ and $c \in \langle b \rangle$. Hence $a^6 = b^{11}$.

Problem 8.21. Prove that an infinite group must have an infinite number of subgroups.

Solution: Let G be an infinite group. Two cases arise:

Case 1. G has an element of infinite order. Let $a \in G$ be such an element. If $H = \langle a \rangle$ then a^k for each $k > 0$, will generate a distinct subgroup of G . Hence G has an infinite number of subgroups.

Case 2. Every element of G is of finite order. Let $g_1 \in G$ and $H_1 = \langle g_1 \rangle$. Then H_1 is a finite subgroup of G . Hence $H_1 \neq G$. Let $g_2 \in G \setminus H_1$, $H_2 = \langle g_2 \rangle$. Then H_2 is a finite subgroup of G and $H_1 \cup H_2 \neq G$. Let $g_3 \in G \setminus (H_1 \cup H_2)$ and $H_3 = \langle g_3 \rangle$. We continue till $G \setminus H_1 \cup H_2 \cup \dots \cup H_k \neq \phi$, $k = 1, 2, 3, \dots$. Since G is infinite and $H_1 \cup H_2 \cup \dots \cup H_k$ is finite, for every k , therefore the process does not end. Hence G has an infinite number of subgroups.

Problem 8.22. Using the concept of cyclic groups, prove that if n is any positive integer and $d_1, d_2, d_3, \dots, d_k$ are all the distinct divisors of n , then

$$\phi(d_1) + \phi(d_2) + \dots + \phi(d_k) = n$$

where ϕ is the Euler ϕ function. Verify the result for $n = 20$.

Solution: *Step 1* Consider the cyclic group \mathbb{Z}_n of order n i.e. $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$. On \mathbb{Z}_n define a relation \sim as follows: For $r, s \in \mathbb{Z}_n$, $r \sim s \Leftrightarrow \langle r \rangle = \langle s \rangle$. Then \sim is an equivalence relation. Each equivalence class contains all the generators of the subgroup associated with it. \therefore Number of equivalence classes = Number of distinct subgroups of \mathbb{Z}_n .

Step 2 We know that in a cyclic group of order m , there exists a unique subgroup of order k , for each divisor k of m , and vice versa. Therefore, number of distinct subgroups of \mathbb{Z}_n is same as the number of distinct divisors of n . Let H_1, H_2, \dots, H_k be subgroups of \mathbb{Z}_n of orders d_1, d_2, \dots, d_k respectively. Then number of generators of H_i is $\phi(d_i)$, since every element of \mathbb{Z}_n belongs to one and only one equivalence class. $\therefore o(\mathbb{Z}_n)$ is the sum of the numbers of elements in disjoint equivalence classes, i.e. the sum of the number of generators of the distinct subgroups of \mathbb{Z}_n . Thus

$$\phi(d_1) + \phi(d_2) + \phi(d_3) + \dots + \phi(d_k) = n.$$

Now, verification for $n = 20$.

The distinct divisors of 20 are 1, 2, 4, 5, 10, 20. Hence

$\phi(1) = 1$, $\phi(2) = 1$, $\phi(4) = 2$, $\phi(5) = 4$, $\phi(10) = 4$, $\phi(20) = 8$. We find that

$$\begin{aligned} \phi(1) + \phi(2) + \phi(4) + \phi(5) + \phi(10) + \phi(20) &= 1 + 1 + 2 + 4 + 4 + 8 \\ &= 20 \\ &= n. \end{aligned}$$

8.12 Supplementary Exercise

1. State whether the following statements are true or false.
 - (i) $(\mathbb{Q}, +)$ is a cyclic group.
 - (ii) $(\mathbb{Z}, +)$ is a cyclic group.
 - (iii) Every cyclic group is Abelian.
 - (iv) Every Abelian group is cyclic.
 - (v) Every element of a cyclic group generates the group.
 - (vi) Every group of order ≤ 4 is cyclic.

- (vii) Every group of order < 4 is cyclic.
 - (viii) Every cyclic group has a unique generator.
 - (ix) Let G be a group and $a \in G$ such that $a^{12} = e$, then $o(a)$ can be 5.
 - (x) There exists a cyclic group of every order.
 - (xi) Every subgroup of a cyclic group is cyclic.
 - (xii) If every subgroup of a group is cyclic then the group is cyclic.
 - (xiii) If every proper subgroup of a group is cyclic then the group is cyclic.
 - (xiv) The number of subgroups of a cyclic group of order n is $\phi(n)$.
 - (xv) (\mathbb{R}^+, \cdot) is a cyclic group.
 - (xvi) (\mathbb{R}^*, \cdot) is a non cyclic group.
 - (xvii) If G is a finite group of order n then G has a subgroup of order m for every divisor m of n .
2. Give an example of a cyclic group of order 4.
 3. Give an example of a non-cyclic group of order 4.
 4. Give an example of a cyclic group of order 180.
 5. If G is an Abelian group and contains a pair of subgroups of order 2, show that G must contain a subgroup of order 4. Is this subgroup necessarily cyclic?
 6. If G is an Abelian group and contains cyclic subgroups of order 4 and 5. What are the other sizes of cyclic subgroups G ?
 7. If G is an Abelian group and contains cyclic subgroup of order 4 and 6. What are the sizes of other cyclic subgroups, G must contain?
 8. Give an example of a group which is not cyclic but its every proper subgroup is cyclic.
 9. Give an example of a non-Abelian group where every proper subgroup is Abelian.
 10. Suppose G is a group with exactly 8 elements of order 10. How many cyclic subgroups of order 10 does it have?
 11. Find the smallest subgroup of \mathbb{Z} containing
 - (i) 6 and 4.
 - (ii) 8 and 15.
 - (iii) 12, 15 and 18.
 - (iv) 4 and 16.
 - (v) m and n .
 12. Given $n_1, n_2, \dots, n_k \in \mathbb{Z}$. Prove that there exists $d \in \mathbb{Z}$, $d > 0$ such that the smallest subgroup of \mathbb{Z} containing n_1, n_2, \dots, n_k is $\langle d \rangle$.
 13. For each value of $n = 5, 8, 9, 10, 14, 15, 16, 18, 20, 22, 25$
 - (i) determine whether $U(n)$ is a cyclic group or not.
 - (ii) in case it is cyclic find all the generators.
 14. Suppose a, b are elements of a group G such that a has odd order and $aba^{-1} = b^{-1}$. Show that $b^2 = e$.

15. Let G be an infinite group and let $x \in G$ is of infinite order.
 - (i) Is $G = \langle x \rangle$? Prove your assertion.
 - (ii) What happens when G is cyclic?
16. Prove that a group of order 5 must be cyclic.
17. Find all subgroups of \mathbb{Z}_{45} giving a generator for each. Describe the containment between these subgroups. Also draw the subgroup lattice.
18. Draw the subgroup lattice of D_{10} .

8.13 Answers to Exercises

Exercise - 8.3

1. Do it yourself.
2. Cyclic $S = \langle 3 \rangle = \langle 12 \rangle$.
3. 49, is the identity element. Group is not cyclic.
4. None
5. *Hint:* Suppose it is cyclic. Let $\mathbb{Q}^+ = \langle x \rangle$ where $x = \frac{p}{q}$ and $(p, q) = 1$. Since $2 \in \mathbb{Q}^+$, therefore $2 = (\frac{p}{q})^n$ for some n . Clearly $n \neq 0, 1, -1$. If $n > 1$, then $p^n = 2q^n$ so that 2 divides p . But then 2 also divides q . So 2 divides (p, q) , a contradiction. Hence \mathbb{Q}^+ is not cyclic.
6. For $n = 7, 10, 13, 14$ the group $\mathcal{U}(n)$ is cyclic. Others are non cyclic.
7. Generated by $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Exercise - 8.5

1. 1 and -1 .
2. 1, 3, 7, 9
3. Two a and a^{-1}
4. (i) $4; \{4k \bmod 10 : k \in \mathcal{U}(5)\}$
 (ii) $2; \{2k \bmod 12 : k \in \mathcal{U}(6)\}$
 (iii) $4; \{6k \bmod 20 : k \in \mathcal{U}(10)\}$
 (iv) $4; \{3k \bmod 24 : k \in \mathcal{U}(8)\}$
 (v) $6; \{5k \bmod 35 : k \in \mathcal{U}(7)\}$
5. (i) $2; \{2^n \bmod 5 : n \in \mathcal{U}(4)\} = \{2, 3\}$
 (ii) $2; \{2^n \bmod 9 : n \in \mathcal{U}(6)\} = \{2, 5\}$
 (iii) $2; \{3^n \bmod 10 : n \in \mathcal{U}(4)\} = \{3, 7\}$
 (iv) $2; \{5^n \bmod 18 : n \in \mathcal{U}(6)\} = \{5, 11\}$
 (v) $4; \{7^n \bmod 22 : n \in \mathcal{U}(10)\} = \{7, 13, 17, 19\}$
 (vi) $8; \{2^n \bmod 25 : n \in \mathcal{U}(20)\} = \{1, 3, 7, 9, 11, 13, 17, 19\}$.

6.

- (i) $4; \{a^k : k \in \mathcal{U}(8)\} = \{a, a^3, a^5, a^7\}$
 (ii) $8; \{b^k : k \in \mathcal{U}(20)\} = \{b, b^3, b^7, b^9, b^{11}, b^{13}, b^{17}, b^{19}\}$

7. $o(G) = n > 1$. Let $e \neq a \in G$ and $H = \langle a \rangle$. If $o(H) = m$ then $m > 1$. If p is a prime dividing m , then $m = pk$, $k \in \mathbb{Z}^+$, then $o(a^k) = p$.

8. If $a \in \mathbb{Z}_n$ is a generator, so is $n - a$. Also $\gcd(a, n) = 1$. If $a = n - a$, then $2a = n$. Therefore, if n odd $a \neq n - a$. If n even $2a = n \Rightarrow \gcd(a, n) \neq 1$, a contradiction. Therefore generators occur in pairs. Number of generators is even.

9.

- (i) If $o(H) = 15$ then $H = \langle a^7 \rangle$ (since $105 \div 15 = 7$) $H = \langle a^{7k} \rangle$
 if $k \in \mathcal{U}(15)$, i.e. $k = 1, 2, 4, 7, 8, 11, 13, 14$.
 (ii) $o(K) = 21$, then $K = \langle a^{5k} \rangle$, $k \in \mathcal{U}(21)$
 (iii) $o(L) = 35$, then $L = \langle a^{3k} \rangle$, $k \in \mathcal{U}(35)$

10.

- (i) 4 hours and 4 minutes.
 (ii) 1 hours and 23 minutes.
 (iii) No. She would never reached her destination.
 (iv) 2 hours and 13 minutes.
 (v) Yes. She would have reached her destination in 44 minutes.

11. Yes. $S = \{1, 2, 3, \dots, 20\}$. S is the cyclic group $(\mathbb{Z}_{20}, \oplus_{20})$ generated by 1. The stoppage of the fast train is the cyclic subgroup generated by 2. The stoppage of the express train is the cyclic subgroup generated by 3. The stoppage of the super fast train is the cyclic subgroup generated by 6.

Exercise - 8.10

1. number of subgroups = number of divisors of 20 = 6. Generators are

- $H_1 = \langle e \rangle$
 $H_2 = \langle a^{10} \rangle$
 $H_3 = \langle a^5 \rangle = \langle a^{15} \rangle$
 $H_4 = \langle a^{4k} \rangle$, $k \in \mathcal{U}(5)$
 $H_5 = \langle a^{2k} \rangle$, $k \in \mathcal{U}(10)$
 $H_6 = \langle a^k \rangle$, $k \in \mathcal{U}(20)$

2. 6.

- $H_1 = \langle 0 \rangle$
 $H_2 = \langle 9 \rangle$
 $H_3 = \langle 6k \rangle$, $k \in \mathcal{U}(3)$
 $H_4 = \langle 3k \rangle$, $k \in \mathcal{U}(6)$
 $H_5 = \langle 2k \rangle$, $k \in \mathcal{U}(9)$
 $H_6 = \langle k \rangle$, $k \in \mathcal{U}(18)$

3. $\langle a^{7k} \rangle$, $k \in \mathcal{U}(4)$ 4. a^5 and a^{-5}

5. An element of order 8 is $10000k$, $k \in \mathcal{U}(8)$
6. (i) $\mathcal{U}(6)$
(ii) $\mathcal{U}(10); a^k$, $k \in \mathcal{U}(10)$
7. $\text{lcm}[m, n]$
8. If G is infinite cyclic then G has no element of order p , hence not possible. So G is finite. If it is cyclic, then an element a of order p gives rise to a subgroup $\langle a \rangle$ of order p , which has exactly $\phi(p) = p - 1 =$ number of elements of order p . Thus it is not possible to have more than $p - 1$ elements of order p .
9. 5, 5, 3, 15, 15.
10. $\langle a \rangle \cap \langle b \rangle \leq \langle a \rangle$, $\langle a \rangle \cap \langle b \rangle \leq \langle b \rangle$, $\therefore o(\langle a \rangle \cap \langle b \rangle) \mid o(\langle a \rangle)$ and $o(\langle a \rangle \cap \langle b \rangle) \mid o(\langle b \rangle)$.
11. 1 or 2.

Supplementary Exercises

1. (i) False
(ii) True
(iii) True
(iv) False, An example is (Q^+, \cdot) .
(v) False, In $(\mathbb{Z}, +)$, 2 is not a generator.
(vi) False, for example V_4 .
(vii) True
(viii) False, $(\mathbb{Z}, +)$ has 1, -1 as generators.
(ix) False
(x) True
(xi) True
(xii) True
(xiii) False, for example S_3
(xiv) False, Every subgroup of a cyclic group is cyclic.
(xv) False, It has a non cyclic subgroup (Q^+, \cdot) .
(xvi) True
(xvii) False
2. $\{1, -1, i, -i\}$, where $i^2 = -1$
3. V_4
4. 180th roots of unity.
5. If $o(a) = o(b) = 2$ then $\{e, a, b, ab\}$ is a subgroup (since $ab=ba$) of order 4. It is not cyclic as there is no element of order 4.
6. G will contain an element of order $4 \times 5 = 20$, $\therefore G$ has cyclic subgroups of all orders which are divisors of 20, i.e. 20, 10, 2 and 1.

7. Show that $o(ab) = 12$. Sizes of other cyclic subgroups are 1, 2, 3 and 12.
8. D_6
9. Do it yourself
10. Let $a \in G$ be $o(a) = 10$. Then a^3, a^7, a^9 are also of order 10. \therefore a cyclic subgroup of order 10 has exactly 4 elements of order 10. Hence G has 2 cyclic subgroups of order 10.
11. (i) $\langle 2 \rangle$
 (ii) \mathbb{Z}
 (iii) $\langle 3 \rangle$
 (iv) $\langle 4 \rangle$
 (v) $\langle d \rangle$, $d = \gcd(m, n)$
12. $\langle d \rangle$, where $d = \gcd(n_1, n_2, \dots, n_k)$.
13. $U(n)$, for $n = 8, 15, 16, 20$ are non-cyclic, all others are cyclic.
 $U(5) = \langle 2 \rangle = \langle 3 \rangle$
 $U(9) = \langle 2 \rangle = \langle 5 \rangle$
 $U(10) = \langle 3 \rangle = \langle 7 \rangle$
 $U(14) = \langle 3 \rangle = \langle 5 \rangle$
 $U(18) = \langle 5 \rangle = \langle 11 \rangle$
 $U(22) = \langle 7 \rangle = \langle 13 \rangle = \langle 17 \rangle = \langle 19 \rangle$
 $U(25) = \langle 2 \rangle = \langle 3 \rangle = \langle 8 \rangle = \langle 12 \rangle = \langle 13 \rangle = \langle 17 \rangle = \langle 22 \rangle = \langle 23 \rangle$.
14. $a^2ba^{-2} = (b^{-1})^{-1} = b$. Prove that $a^{2n}ba^{-2n} = b$. Then $o(a) = \text{odd} \Rightarrow o(a) = 2k + 1, a^{2k+1} = e$.
 Now, $a^{2k}ba^{-2k} = b \Rightarrow a^{2k+1}ba^{-2k-1} = aba^{-1} \Rightarrow b = b^{-1} \Rightarrow b^2 = e$
17. $\langle 1 \rangle, \langle 3 \rangle, \langle 5 \rangle, \langle 9 \rangle, \langle 15 \rangle, \langle 0 \rangle$ are subgroups of order 45, 15, 9, 5, 3 respectively.
 $\langle 0 \rangle \subset \langle 9 \rangle \subset \langle 3 \rangle \subset \langle 1 \rangle$
 $\langle 0 \rangle \subset \langle 15 \rangle \subset \langle 3 \rangle \subset \langle 1 \rangle$
 $\langle 0 \rangle \subset \langle 15 \rangle \subseteq \langle 5 \rangle \subseteq \langle 1 \rangle$

This page is intentionally left blank.

UNIT - 3

Chapter 9

Rings

9.1 Ring

Definition 9.1. A non-empty set R equipped with two binary operations $+$ and (usually called addition and multiplication) is called a ring if

- (i) $(R, +)$ is an Abelian group,
- (ii) (R, \cdot) is a semigroup,
- (iii) \cdot is right as well as left distributive over $+$, that is
$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \forall a, b, c \in R$$
$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \forall a, b, c \in R$$

It is denoted by $(R, +, \cdot)$.

When the operations are understood we simply say that R is a ring. Moreover we use juxtaposition instead of \cdot . Since $(R, +)$ is an Abelian group, therefore

- (i) the additive identity is unique,
- (ii) additive inverse of an element is unique,
- (iii) cancellation laws hold for addition.

The additive identity of a ring is called the zero element and is denoted by 0 . This should not be confused with the integer 0 .

If $(R, +, \cdot)$ is a ring then $(R, \cdot, +)$ need not be a ring as

- (i) either (R, \cdot) may not be an Abelian group or
- (ii) $+$ may not be distributive over \cdot .

Hence the order of the binary operations is important.

Definition 9.2. (Commutative ring):

A ring $(R, +, \cdot)$ is said to be commutative if $a \cdot b = b \cdot a, \forall a, b \in R$.

Definition 9.3. (Ring with unity):

A ring $(R, +, \cdot)$ is said to be a ring with unity if there exist an element $e \in R$ such that $a \cdot e = e \cdot a = a$, for all $a \in R$.

The element ' e ' is called the unity of R .

The unity is also called the identity or unit element of R . Generally it is denoted by 1 (not to be confused with the integer 1). The unity of a ring, if it exists, is unique.

If R is a ring and $a \in R, n \in \mathbb{Z}$ then na has its usual meaning as in additive groups. Note that $\frac{a}{2}, \frac{a}{3}$ etc... where $a \in R$, are not defined in a ring. A ring having a finite number of elements is called a finite ring. If we consider $R = \{a\}$ and define $+$ and \cdot as follows : $a + a = a, a \cdot a = a$, then R is a ring with a as the zero element. Such a ring is called the zero ring and is denoted O .

9.2 Examples of Ring

Rings of Numbers

Example 9.1. *The set of integers \mathbb{Z} is a ring under the usual addition and multiplication. It is a commutative ring with unity 1. Similarly \mathbb{Q} and \mathbb{C} are commutative rings with unity 1, under the usual addition and multiplication of numbers.*

Example 9.2. *$2\mathbb{Z}$, under the usual addition and multiplication is a commutative ring without unity. For $n \in \mathbb{N}, n > 1, n\mathbb{Z}$ is a commutative ring without unity.*

Let $Z_o =$ set of odd integers. Then $(Z_o, +)$ is not even a group, so that Z_o is not a ring.

Example 9.3. *Let $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Z}\}$. Under the usual addition and multiplication of numbers, $\mathbb{Z}[\sqrt{2}]$ is a ring. It is a commutative ring and has unity 1. Similarly $\mathbb{Q}[\sqrt{2}]$ is a commutative ring with unity 1.*

Example 9.4. *Let $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}$. Under the usual addition and multiplication of complex numbers $\mathbb{Z}[i]$ is a commutative ring with unity $1+i0=1$. This ring is called the ring of Gaussian Integers.*

Rings of Residues

Example 9.5. *Let $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. The tables for (\mathbb{Z}_4, \oplus_4) and (\mathbb{Z}_4, \odot_4) are given below*

\oplus_4	0	1	2	3	\odot_4	0	1	2	3
0	0	1	2	3	0	0	0	0	0
1	1	2	3	0	1	0	1	2	3
2	2	3	0	1	2	0	2	0	2
3	3	0	1	2	3	0	3	2	1

(\mathbb{Z}_4, \oplus_4) is a group, as shown in chapter 5. Clearly (\mathbb{Z}_4, \odot_4) is a semigroup. We know that in modular arithmetic

$$(a + b) \text{ mod } 4 = (a \text{ mod } 4 + b \text{ mod } 4) \text{ mod } 4 \dots (1)$$

$$(ab) \text{ mod } 4 = ((a \text{ mod } 4)(b \text{ mod } 4)) \text{ mod } 4 \dots (2)$$

Let $a, b, c \in \mathbb{Z}_4$, then

$$a \odot_4 (b \oplus_4 c) = (a \odot_4 b) \oplus_4 (a \odot_4 c).$$

because $RHS = (a \odot_4 b) \oplus_4 (a \odot_4 c)$

$$= ((a \odot_4 b) + (a \odot_4 c)) \text{ mod } 4 \quad \text{using definition of } \oplus_4$$

$$= (ab \text{ mod } 4 + ac \text{ mod } 4) \text{ mod } 4 \quad \text{using definition of } \odot_4$$

$$= (ab + ac) \text{ mod } 4 \quad \dots \text{using (1)}$$

$$= (a(b + c)) \text{ mod } 4$$

$$= ((a \text{ mod } 4 (b + c) \text{ mod } 4)) \text{ mod } 4 \quad \text{using (2)}$$

$$= (a(b \oplus_4 c)) \text{ mod } 4$$

$$= a \odot_4 (b \oplus_4 c)$$

= LHS

Similarly $(a \oplus_4 b) \odot_4 = (a \odot_4 c) \oplus_4 (b \odot_4 c)$ can be shown.

Hence distributive laws hold.

Thus $(\mathbb{Z}_4, \oplus_4, \odot_4)$ is a ring. It is a commutative ring and 1 is the unit element.

Hence $(\mathbb{Z}_4, \oplus_4, \odot_4)$ is a finite commutative ring with unity.

Example 9.6. Let $n > 1$ be a fixed positive integer and let $\mathbb{Z}_n = \{0, 1, 2, 3, \dots, n-1\}$.

Then (\mathbb{Z}_n, \oplus_n) is an Abelian group.

Also \mathbb{Z}_n is closed with respect to \odot_n and associative law holds. Thus (\mathbb{Z}_n, \odot_n) is a semigroup.

Also in the above example (on replacing 4 by n) we can prove that for all $a, b, c \in \mathbb{Z}_n$

$$a \odot_n (b \oplus_n c) = (a \odot_n b) \oplus_n (a \odot_n c)$$

$$(a \oplus_n b) \odot_n c = (a \odot_n c) \oplus_n (b \odot_n c).$$

Thus $(\mathbb{Z}_n, \oplus_n, \odot_n)$ is a ring. Moreover it is a commutative ring with unity 1.

Hence $(\mathbb{Z}_n, \oplus_n, \odot_n)$ is a finite commutative ring with unity.

Rings of Matrices

Example 9.7. Under the usual addition and multiplication of matrices, $M_2(\mathbb{Z})$ is a ring with unity I_2 . It is noncommutative ring. Thus, $M_2(\mathbb{Z})$ is a non-commutative ring with unity.

Similarly the set of all $n \times n$ matrices over \mathbb{Z} (\mathbb{Q} , \mathbb{R} or \mathbb{C}) are non-commutative rings with unity, for every $n \in \mathbb{N}$.

Example 9.8. $M_2(2\mathbb{Z})$ is a noncommutative ring without unity. The possible unity is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, which does not belong to $M_2(2\mathbb{Z})$.

Example 9.9. $M_2(\mathbb{Z}_2)$ is a finite non-commutative ring with unity I_2 . This ring has 2^4 elements.

$M_n(\mathbb{Z}_2)$ is a finite non-commutative ring with unity, have 2^{n^2} elements.

Ring of polynomials

Example 9.10. The sets $\mathbb{Z}[x]$ of all polynomials in x with integral coefficients under the usual addition and multiplication of polynomials, is a commutative

ring with unity. The unity is the constant polynomial 1. Similarly $\mathbb{Q}[x]$ and $\mathbb{R}[x]$ are commutative rings with unity.

Ring of Functions

Example 9.11. Let F be the set of all functions from \mathbb{R} into \mathbb{R} . On F define addition and multiplication as follows:

For $f, g \in F$, Define

$$(f + g)(x) = f(x) + g(x) \quad \forall x \in \mathbb{R}$$

$$(fg)(x) = f(x)g(x) \quad \forall x \in \mathbb{R}.$$

F is closed with respect to addition and multiplication. Addition and multiplication of real valued functions is associative and commutative.

The zero function is the zero element for addition, and inverse of f is the function $-f$. Thus $(F, +)$ is an Abelian group and (F, \cdot) is a semigroup.

Since distributive law holds in numbers, therefore it holds in F also. Thus $(F, +, \cdot)$ is a ring. It is a commutative ring. The constant function i which maps every element of \mathbb{R} to 1 is the unit element. Hence $(F, +, \cdot)$ is a commutative ring with unity.

Example 9.12. Let $C[0, 1]$ be the set of all real valued continuous functions defined on $[0, 1]$. On $C[0, 1]$ define addition and multiplication of functions pointwise. As the sum and the product of two continuous functions is continuous, therefore $C[0, 1]$ is a commutative ring with unity.

Ring of Quaternions

Let $D = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}$ and i, j, k are such that $i^2 = j^2 = k^2 = -1, ij = k = -ji, jk = i = -kj, ki = j = -ik$.

Define addition componentwise, and multiplication by using the distributive laws and the relations given above. It can be verified that D is a ring. It is called the ring of real quaternions. It is a non-commutative ring with unity D .

Elementary Properties of Ring

Theorem 9.1. Let R be a ring and $a, b \in R$. Then

(i) $0a = a0 = 0$, where 0 is the zero element of R .

(ii) $a(-b) = (-a)b = -ab$

(iii) $(-a)(-b) = ab$

(iv) if R has unity 1 , then

$$(-1)a = -a, (-1)(-1) = 1$$

Proof:

(i) Let $a \in R$. Then

$$(0 + 0)a = 0a + 0a \quad \text{By distributive law}$$

$$\Rightarrow 0a = 0a + 0a; \quad \because 0 + 0 = 0$$

$$\Rightarrow 0a + 0 = 0a + 0a \quad \text{By definition of zero element}$$

$$\Rightarrow 0 = 0a \quad \text{Using cancellation law for addition}$$

$$\text{Hence } 0a = 0, \quad \text{for all } a \in R$$

Similarly $a0 = 0$ can be proved.

(ii) Let $a, b \in R$

$$\begin{aligned}
 & \text{Since } b + (-b) = 0 \\
 & \therefore a(b + (-b)) = a0 \\
 & \Rightarrow ab + a(-b) = 0 \quad \text{Using distributive law and (i)} \\
 & \Rightarrow ab + a(-b) = ab + (-ab), \quad \text{By definition of additive inverse} \\
 & \Rightarrow a(-b) = -ab \quad \text{Using cancellation law for addition} \\
 & \text{Similarly } (-a)b = -ab.
 \end{aligned}$$

(iii) In (ii) replace a by $-a$

$$\begin{aligned}
 & \therefore (-a)(-b) = -(-a)b \\
 & \quad = -(-ab) \quad \text{using (ii)} \\
 & \quad = ab \\
 & \therefore (-a)(-b) = ab.
 \end{aligned}$$

(iv) Suppose R has unity 1.

$$\begin{aligned}
 & \text{From (ii)} \\
 & (-a)b = -ab \\
 & \text{Take } a = 1, b = a \\
 & \therefore (-1)a = -1a \\
 & \Rightarrow (-1)a = -a \\
 & \text{In (iii) take } a = 1 = b \\
 & \therefore (-1)(-1) = 1.1 = 1 \\
 & \text{Hence } (-1)(-1) = 1. \quad \square
 \end{aligned}$$

Theorem 9.2. Let R be a ring and $a, b \in R, m, n \in \mathbb{Z}$. Then

- (i) $0a = 0, 0 \in \mathbb{Z}$
- (ii) $(-n)a = -na$
- (iii) $(ma)(nb) = (mn)(ab)$.

Proof: Parts (i) and (ii) follow from the fact that $(R, +)$ is a group.

(iii) Four cases arise:

Case 1. $m = 0$ or $n = 0$

$$\text{then } (ma)(nb) = 0 \quad \text{Using (i)}$$

$$(mn)(ab) = 0(ab) = 0$$

$$\therefore (ma)(nb) = (mn)(ab)$$

Case 2. let $m, n \in \mathbb{N}$. Then $m > 0$ and $n > 0$

$$\begin{aligned}
 (ma)(nb) &= (a + a + a + \dots m \text{ times})(nb) \\
 &= a(nb) + a(nb) + \dots m \text{ times} \quad \text{using distributive law} \\
 &= a(b + b + b \dots n \text{ times}) + a(b + b + b \dots n \text{ times}) + \dots m \text{ times} \\
 &= (ab + ab + \dots n \text{ times}) + (ab + ab + \dots n \text{ times}) + \dots m \text{ times} \\
 &= ab + ab + ab + \dots mn \text{ times} \\
 &= (mn)(ab)
 \end{aligned}$$

Case 3. One of m, n is positive and the other is negative.

Suppose $m < 0, n > 0$.

Let $m = -p$, where $p > 0$

$$\begin{aligned}
 \text{Then } (ma)(nb) &= ((-p)a)(nb) \\
 &= (-pa)(nb) \quad \text{Using theorem 9.1} \\
 &= -(pa)(nb) \quad \text{Using theorem 9.1} \\
 &= -(pn)(ab) \quad \text{By case 2} \\
 &= (-pn)(ab) \quad \text{Using (ii)}
 \end{aligned}$$

$$\begin{aligned}
&= ((-p)n)(ab) \\
&= (mn)(ab) \\
\therefore (ma)(nb) &= (mn)(ab).
\end{aligned}$$

The result can be proved similarly if $m > 0$ and $n < 0$.

Case 4. $m < 0, n < 0$.

Let $m = -m_1, n = -n_1$, where $m_1, n_1 > 0$

$$\begin{aligned}
(ma)(nb) &= ((-m_1)a)((-n_1)b) \\
&= (-m_1a)(-n_1b), \quad \text{using (ii)} \\
&= (m_1a)(n_1b) \quad \text{using theorem 9.1} \\
&= (m_1n_1)(ab), \quad \text{using case 2} \\
&= ((-m)(-n))ab \\
&= (mn)ab.
\end{aligned}$$

Hence result is proved in all the cases. \square

9.3 Constructing New Rings

Having the knowledge of a basic ring (given in the examples) we can construct many more rings according to our requirement by making a few variations. How this can be done is explained below.

Let R be any ring. Then $M_n(R)$, the set of all $n \times n$ matrices over R , is a non-commutative ring such that if R has unity 1 then $M_n(R)$ has unity I_n . If R does not have unity then $M_n(R)$ also does not have unity. Furthermore, if R is a finite ring having m elements then $M_n(R)$ is also a finite ring having m^{n^2} elements. Thus $M_n(\mathbb{Z})$ is an infinite non-commutative ring with unity I_n whereas $M_2(2\mathbb{Z})$ is a non-commutative ring without unity. $M_n(\mathbb{Z}_m)$ is a finite non-commutative ring with unity having m^{n^2} elements.

$M_2(2\mathbb{Z}_8)$ is a finite non-commutative ring without unity.

This helps us to construct infinite or finite non-commutative rings with or without unity.

Analogous to the ring $\mathbb{Z}[i]$, we construct the ring $\mathbb{Z}_n[i]$ where

$\mathbb{Z}_n[i] = \{a + ib | a, b \in \mathbb{Z}_n\}$. The elements are added and multiplied as in complex numbers except that the coefficients are reduced modulo n . This is called the ring of Gaussian integers modulo n .

Consider $\mathbb{Z}_n[x]$, the set of all polynomials in x with coefficients from \mathbb{Z}_n . The elements are added and multiplied as in $\mathbb{Z}[x]$ except that the coefficients are reduced modulo n .

If R_1 and R_2 are two rings we can use them to construct a new ring as follows:

Define $R_1 \times R_2 = \{(a, b) | a \in R_1, b \in R_2\}$

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$$

$$(a_1, a_2)(b_1, b_2) = (a_1b_1, a_2b_2)$$

Then $R_1 \times R_2$ is a ring under these operations. It is called the direct sum of R_1 and R_2 , denoted by $R_1 \oplus R_2$. The zero element of $R_1 \oplus R_2$ is $(0, 0)$. If R_1 and R_2 are finite having n_1, n_2 elements respectively, then $R_1 \oplus R_2$ is also finite having n_1n_2 elements. If both R_1, R_2 are commutative so is $R_1 \oplus R_2$. If both R_1, R_2 have unity e_1 and e_2 respectively then (e_1, e_2) is the unity of $R_1 \oplus R_2$.

We can similarly define the direct sum of n rings. It appears that all the properties of R_1 and R_2 are carried over to $R_1 \oplus R_2$, but hold on, this is not the case. Wait and watch.

9.4 Special Elements of a Ring

Some of the elements of a ring have a special property, which makes them stand out from the other elements. Knowledge of these elements helps us to know the structure of the ring in a better way.

Definition 9.4. (Idempotent element):

Let R be a ring. An element $a \in R$ is said to be an idempotent if $a^2 = a$.

If $a \in R$ is an idempotent then all the powers of a are identical. The zero element and the unity if it exists are idempotents.

Definition 9.5. (Nilpotent element):

Let R be a ring. An element $a \in R$ is said to be nilpotent if $a^n = 0$ for some positive integer n .

The zero element is a nilpotent element.

Definition 9.6. (Unit):

Let R be a ring with unity 1. An element $a \in R$ is said to be a unit in R if there exists some $b \in R$ such that $ab=ba=1$

Definition 9.7. (Boolean ring):

A ring in which every element is an idempotent is said to be a Boolean ring. Such a ring will be shown to be commutative.

Remark 9.1. In groups, equations of the form $ax = b$ have a solution for all elements a and b of the group. But this is not the case in rings. For example in the ring \mathbb{Z} the equation $2x=3$, has no solution. But in $(\mathbb{Z}_5, \oplus_5, \odot_5)$ the equation $2 \odot_5 x = 3$. . . (1)

has a solution, viz. $x=4$. This is because $2 \in \mathbb{Z}_5$ is a unit and its inverse is 3. Premultiplying equation (1) by $2^{-1} = 3$ we get

$$3 \odot_5 (2 \odot_5 x) = 3 \odot_5 3$$

or $x=4$.

Thus in a ring R the equations of the form $ax=b$, have a solution if a is a unit, and the solution is $x = a^{-1}b$.

Thus units of a ring help us to solve certain equations. In a ring an equations of the form $x^2 = a^2$, does not imply that $x = \pm a$ in a ring.

For example, in $(\mathbb{Z}_8, \oplus_8, \odot_8)$, $7^2 = 3^2$ but $7 \neq \pm 3$. Thus the usual identities and rules valid in \mathbb{R} do not hold good in any arbitrary ring.

Theorem 9.3. In a ring R with unity, the set of all units form a group under multiplication.

Proof: Let 1 be the unity of R and S be the set of all units of R .

Since $1 \in R$ is a unit $\therefore 1 \in S$, so that $S \neq \phi$.

Let $a, b \in S$. Then there exists $c, d \in S$

such that $ac = ca = 1$, $bd = db = 1$.

Now $(ab)(dc) = a(bd)c = a1c = ac=1$.

Similarly $(dc)(ab) = 1$ and so $(ab)(dc) = (dc)(ab) = 1$.

	<i>Ring</i>	<i>Idempotent Element</i>	<i>Nilpotent Element</i>	<i>Units</i>
1	\mathbb{Z}	0, 1	0	± 1
2	\mathbb{Q}	0, 1	0	Every non-zero element
3	\mathbb{R}	0, 1	0	Every non-zero element
4	\mathbb{C}	0, 1	0	Every non-zero element
5	$2\mathbb{Z}$	0	0	Not applicable as $2\mathbb{Z}$ is without unity
6	\mathbb{Z}_4	0, 1	0, 2	1, 3
7	\mathbb{Z}_9	0, 1	0, 3, 6	1, 2, 4, 5, 7, 8
8	\mathbb{Z}_{18}	0, 1	0, 6	1, 5, 7, 11, 13, 17
9	\mathbb{Z}_n	0, 1	$\{0\} \cup \{\frac{n}{p} p$ is a prime and $p n\}$	$r \in \mathbb{Z}_n$ such that $(r, n) = 1$
10	$\mathbb{Z}[x]$	0, 1	0	Same as those of \mathbb{Z} , ± 1

Hence ab is a unit and. therefore, $ab \in S$.

Since (R, \cdot) is a semi-group, therefore associative law holds in S .

Since $1 \in R$ is a unit, therefore $1 \in S$ and is the identity element of S .

Let $a \in S$. Then a is the unit so that there exists $b \in R$ such that $ab = ba = 1$.

Then b is a unit in R , so that $b \in S$. But $a^{-1} = b$, so that $a^{-1} \in S$.

Hence every element of S has an inverse.

Thus S is a group under multiplication.

The group of all units of R is denoted by $U(R)$. □

9.5 Solved Problems

Problem 9.1. Let $(R, +, \cdot)$ be a ring with unity 1.

Define \oplus and \odot in R as follows:

$$a \oplus b = a + b + 1$$

$$a \odot b = a + b + ab, \text{ for all } a, b \in R.$$

(i) Prove that $R' = (R, \oplus, \odot)$ is a ring.

(ii) What is the zero element of R' ?.

(iii) Does R' has unity ?

(iv) If R is a commutative ring then prove that R' is also commutative.

Solution:

(i) Let $a, b \in R$. Then $a + b + 1 \in R$, so that $a \oplus b \in R$.

Since associative law for addition holds in R , therefore it holds for \oplus .

Let $a \in R$. Let $0_R \in R$ such that

$$a \oplus 0_R = 0_R \oplus a = a$$

$$\therefore a + 0_R + 1 = a$$

so that $0_R = -1$.

Hence -1 is the zero element of R' .

Let $a \in R$, suppose that $b \in R$ such that

$$a \oplus b = -1$$

$$\therefore a + b + 1 = -1$$

so that $b = -a - 2$.

Thus additive inverse of a is $-a - 2$, so that each element has an inverse in R' .

Let $a, b \in R$, then

$$a \oplus b = a + b + 1 = b + a + 1 = b \oplus a$$

Hence (R, \oplus) is an Abelian group.

For $a, b \in R, a + b + ab \in R$, so that $a \odot b \in R$. $\therefore R$ is closed with respect to \odot .

Let $a, b, c \in R$,

then $(a \odot b) \odot c$

$$= (a + b + ab) \odot c$$

$$= a + b + ab + c + (a + b + ab)c$$

$$= a + b + ab + c + ac + bc + (ab)c$$

$$= a + ab + ac + (b + c + bc) + a(bc)$$

$$= a + a(b + c + bc) + b \odot c$$

$$= a + a(b \odot c) + b \odot c$$

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

Hence \odot is associative.

That \odot is right and left distributive over \oplus can be easily verified.

Hence R' is a ring.

(ii) The zero element of R' is -1 as proved in (i).

(iii) If 0 denotes the zero element of R , then

for any $a \in R$,

$$a \odot 0 = 0 \odot a = a \text{ (verify!)}$$

Hence 0 is the unity of the R' .

(iv) For $a, b \in R$

$$a \odot b = a + b + ab$$

$$= b + a + ba, \quad \text{using commutativity of } R$$

$$= b \odot a$$

Hence R' is commutative.

Problem 9.2. Let S be a non-empty set and $\mathcal{P}(S)$ the power set of S . For $A, B \in \mathcal{P}(S)$. Define $'+' and $'\cdot'$ on $\mathcal{P}(S)$ as:$

$$A + B = A \triangle B \text{ and } A \cdot B = A \cap B.$$

Prove that $(\mathcal{P}(S), \triangle, \cap)$ is a Boolean ring with unity.

Solution: Clearly Δ and \cap are binary operations on $\mathcal{P}(S)$. The null set ϕ is the zero element of $\mathcal{P}(S)$ and the additive inverse of A is A . The other properties are easy to verify.

Also $A \cap S = A = S \cap A \forall A \in \mathcal{P}(S)$

Thus S is the identity element of $\mathcal{P}(S)$. Hence $\mathcal{P}(S)$ is a ring with unity.

For any $A \in \mathcal{P}(S)$ $A.A = A \cap A = A$, so that $\mathcal{P}(S)$ is a Boolean ring.

Remark: $\mathcal{P}(S)$ is a finite or infinite Boolean ring according as S is finite or infinite.

Problem 9.3. Let R be a system satisfying all the axioms for a ring with the possible exception of $a + b = b + a$. If there exists $c \in R$, such that $ac = bc \Rightarrow a = b \forall a, b \in R$, prove that R is a ring.

Solution: To prove that R is a ring we have to prove that $a + b = b + a$ for all $a, b \in R$.

Let $a, b \in R$. Then $(a + b)(c + c) = a(c + c) + b(c + c)$, using right distributive law

$$\begin{aligned} &= ac + ac + bc + bc \\ \text{again } (a + b)(c + c) &= (a + b)c + (a + b)c \\ &= ac + bc + ac + bc \end{aligned}$$

Thus we get

$$\begin{aligned} ac + ac + bc + bc &= ac + bc + ac + bc \\ \Rightarrow ac + bc &= bc + ac \\ \Rightarrow (a + b)c &= (b + a)c \\ \Rightarrow a + b &= b + a \end{aligned}$$

Hence R is a ring.

The usual properties of integers do not hold in rings. This is shown in the following problem.

Problem 9.4. Find integers a, b, c not having the following property in \mathbb{Z}_n .

- (i) $a^2 = a \Rightarrow a = 0$ or $a = 1$.
- (ii) $ab = 0 \Rightarrow a = 0$ or $b = 0$.
- (iii) $ab = ac, a \neq 0 \Rightarrow b = c$.

Solution:

(i) Consider the ring \mathbb{Z}_6 . Here $3 \in \mathbb{Z}_6, 3^2 = 3$ and $3 \neq 0, 1$.

(ii) In $\mathbb{Z}_6, 2 \odot_6 3 = 0$, but $2 \neq 0, 3 \neq 0$.

(iii) In $\mathbb{Z}_6, 3 \odot_6 2 = 3 \odot_6 4$ but $2 \neq 4$.

Problem 9.5. If a ring R has the property that $ab = ca \implies b = c$, when $a \neq 0$. Prove that R is commutative.

Solution: Let $a, b \in R$.

If $a = 0$ then $ab = 0$ also $ba = 0 \therefore ab = ba$ hold for all $b \in R$.

If $a \neq 0$, then $a(ba) = (ab)a$ by associative law

By the given condition $ba = ab$.

Hence R is commutative.

Problem 9.6. Give an example of an infinite non-commutative ring without unity.

Solution: See Example 9.8.

Problem 9.7. Give an example of a finite non-commutative ring.

Solution: See example 9.9.

Problem 9.8. In $(\mathbb{Z}_p, \oplus_p, \odot_p)$, where p is prime prove that if $a, b \in \mathbb{Z}_p$ such that $a \odot_p b = 0$, then $a = 0$ or $b = 0$.

Solution: Let $a, b \in \mathbb{Z}_p$ such that
 $a \odot_p b = 0 \Rightarrow (ab) \bmod p = 0$
 $\Rightarrow ab$ is a multiple of p .
 $\Rightarrow p|ab \Rightarrow p|a$ or $p|b$, as p is prime.
 $\Rightarrow a$ is a multiple of p or b is a multiple of p .
 $\Rightarrow a = 0$ or $b = 0$, as $a, b \in \mathbb{Z}_p = \{0, 1, \dots, p-1\}$ and so no other element is a multiple of p .

Problem 9.9. Find all the units of $M_2(\mathbb{Z})$.

Solution: Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z})$ be a unit.

Then $|A| \neq 0$ and

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \in M_2(\mathbb{Z})$$

if $|A| = k$, then

$$\frac{d}{k}, \frac{-b}{k}, \frac{-c}{k}, \frac{a}{k} \in \mathbb{Z}$$

$\Rightarrow k|a, k|b, k|c$ and $k|d$. Let $a = ka_1, b = kb_1, c = kc_1, d = kd_1$.

$$k = |A| = ad - bc$$

$$\Rightarrow k = k^2(a_1d_1 - b_1c_1)$$

$$\Rightarrow k^2|k$$

$$\Rightarrow k = \pm 1 \text{ as } k \in \mathbb{Z}.$$

Thus if A is a unit of $M_2(\mathbb{Z})$, then $|A| = \pm 1$.

On the other hand if $|A| = \pm 1$, then $A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ or $-\begin{pmatrix} d & -b \\ -a & a \end{pmatrix}$.

So $A^{-1} \in M_2(\mathbb{Z})$. Hence $A \in M_2(\mathbb{Z})$ is a unit if and only if $|A| = \pm 1$.

Problem 9.10. If R_1 and R_2 are commutative rings with unity prove that $U(R_1 \oplus R_2) = U(R_1) \oplus U(R_2)$.

Solution: Let e_1, e_2 denote the unities of R_1 and R_2 respectively.

Let $r_1 \in R_1$ and $r_2 \in R_2$. Then

$(r_1, r_2) \in R_1 \oplus R_2$ is a unit

\Leftrightarrow there exist $t_1 \in R_1, t_2 \in R_2$ such that

$$(r_1, r_2)(t_1, t_2) = (e_1, e_2)$$

$$\Leftrightarrow (r_1t_1, r_2t_2) = (e_1, e_2)$$

$$\Leftrightarrow r_1t_1 = e_1, r_2t_2 = e_2$$

$\Leftrightarrow r_1$ is a unit in R_1 and r_2 is a unit in R_2 .

$$\text{Hence } U(R_1 \oplus R_2) = U(R_1) \oplus U(R_2).$$

Problem 9.11. Prove that the units of $\mathbb{Z}[x]$ and \mathbb{Z} are the same.

Solution: Let $f(x) \in \mathbb{Z}[x]$ be a unit.

$$\text{Let } f(x) = a_0 + a_1x + \dots + a_nx^n$$

$f(x)$ is a unit in $\mathbb{Z}[x]$
 \Leftrightarrow there exists $g(x) = b_0 + b_1x + \dots + b_mx^m \in \mathbb{Z}[x]$
 such that $f(x)g(x) = g(x)f(x) = 1 \dots (1)$
 $\Leftrightarrow \deg(f(x)g(x)) = \deg 1 = 0$
 $\Leftrightarrow \deg f(x) + \deg g(x) = 0$
 $\Leftrightarrow \deg f(x) = \deg g(x) = 0$
 $\Leftrightarrow f(x)$ and $g(x)$ are the constant polynomials.
 $\Leftrightarrow f(x) = a_0$ and $g(x) = b_0$ for some $a_0, b_0 \in \mathbb{Z}$.
 $\Leftrightarrow a_0$ is a unit in \mathbb{Z} .
 $\Leftrightarrow f(x)$ is a unit in \mathbb{Z} .

Problem 9.12. Let R be a ring with unity such that $(ab)^2 = a^2b^2$, $\forall a, b \in R$. Prove that R is commutative. Prove that the result is not true when R does not have unity.

Solution: Let 1 denote the unity of R .

Step 1 Let $x, y \in R$.
 Then $(x(y+1))^2 = x^2(y+1)^2$
 $\therefore (xy+x)^2 = x^2(y^2+2y+1)$
 $\Rightarrow (xy)^2 + xyx + xxy + x^2 = x^2y^2 + 2x^2y + x^2$
 $\Rightarrow xyx = x^2y \dots (1)$

Step 2 In (1) replace x by $x+1$
 $\therefore (x+1)y(x+1) = (x+1)^2y$
 so that $(x+1)(yx+y) = (x^2+2x+1)y$
 $\Rightarrow xyx + xy + yx + y = x^2y + 2xy + y$
 $\Rightarrow yx = xy$ using (1) and cancellation law for $'+'$.
 Thus $xy = yx \forall x, y \in R$.
 Hence R is commutative.

Let $R = \left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$. Then R is a non-commutative ring without unity, such that $(AB)^2 = A^2B^2$ for all $A, B \in R$.

9.6 Exercise

1. Prove that unity of a ring, if it exists, is unique.
2. Let R be a ring and $a, b, c, d \in R$. Express the following as the sum of the products.
 - (i) $(a+b)(c+d)$
 - (ii) $(a+b)^2$
 - (iii) $(a+b)^3$
 - (iv) $(a+b)^4$
3. If R is a ring find an expression for $(a+b)^n$ in terms of powers of a and b , where $a, b \in R, n \in \mathbb{Z}^+$. What happens if R is commutative?
4. Let R be a system satisfying all conditions for a ring with unity, with the possible exception of $a+b = b+a$. Prove that R is a ring.
5. Let $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$. Prove that R is a ring under the usual addition and multiplication of real numbers.

6. If R is a ring, is $(R, \cdot, +)$ also a ring? Justify.
7. Prove that if a ring with unity has more than one elements, then zero and unity are distinct.
8. Show that \mathbb{Z} with binary composition $\#$ and $*$ defined as $a \# b = a + b + 1$, $a * b = a + b - ab$ for all $a, b \in \mathbb{Z}$ is a commutative ring with unity.
9. Show that $2\mathbb{Z}$ with binary compositions $\#$ and $*$ defined as $a \# b = a + b$, $a * b = \frac{ab}{2}$ for all $a, b \in \mathbb{Z}$ is a commutative ring with unity.
10. Prove that if a ring has a unique left unity then it also has a right unity, and therefore a unity.
11. Give an example of a ring R in which for $a, b \in R$, $ax = b$ has more than one solutions in R .
12. Prove that $M_2(2\mathbb{Z}_8)$, the set of all 2×2 matrices over $(2\mathbb{Z}_8, \oplus_8, \odot_8)$ is a finite non-commutative ring having 4^4 elements.
13. Give an example of a Boolean ring having 4 elements.
14. Let $p(x) = 2x^3 - 3x^2 + 4x - 5$, $q(x) = 7x^3 + 33x - 4$
 Compute $p(x) + q(x)$ and $p(x)q(x)$ under the assumption that the coefficient of the two given polynomial are taken from the specified ring R , when
 (i) $R = \mathbb{Z}$ (ii) $R = \mathbb{Z}_2$ (iii) $R = \mathbb{Z}_3$
15. Let S be any set. Is $(\wp(S), +, \cdot)$ a ring, where $+, \cdot$ are defined as follows for $A, B \in \wp(S)$
 $A + B = A \cup B$, $A \odot B = A \cap B$.
 Justify your answer.
16. Is $S = \left\{ \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}, |a, b \in \mathbb{Z} \right\}$ a ring with respect to usual addition and multiplication of matrices? Give reasons for your answer.
17. Let $R = \{a + bi + cj + dk | a, b, c, d \in \mathbb{Z}\}$ and i, j, k are such that $i^2 = j^2 = k^2 = -1$, $ij = k = -ji$, $jk = i = -kj$, $ki = j = -ik$. Define $a + bi + cj + dk = a' + b'i + c'j + d'k$ if and only if $a = a'$, $b = b'$, $c = c'$, $d = d'$. Define addition componentwise and multiplication by distributive law using the above relations.
 (i) Prove that R is a ring.
 (ii) Is it commutative?
 (iii) Does it have unity?
 (This ring is called the ring of integral quaternions.)
18. Prove that a ring R is commutative if and only if $(a + b)(a - b) = a^2 - b^2$, for all $a, b \in R$.
19. Prove that every Boolean ring is commutative. Is the converse is true? Justify your answer.
20. Give an example of a non-commutative ring R such that $(ab)^2 = a^2b^2$, $\forall a, b \in R$

21. In $(\mathbb{Z}_p, \oplus_p, \odot_p)$, where p is prime, prove that
- if $a \in \mathbb{Z}_p$ and $a^2 = a$ then $a = 0$ or 1
 - if $a, b, c \in \mathbb{Z}_p$ and $a \neq 0$
then $ab=ac \implies b = c$
22. Check whether the following rings have unity. If yes, find it.
- (S, \oplus_8, \odot_8) , where $S=\{0, 2, 4, 6\}$
 - (S, \oplus_6, \odot_6) , where $S=\{0, 2, 4\}$
 - $(S, \oplus_{10}, \odot_{10})$, where $S=\{0, 2, 4, 6, 8\}$
 - $(S, \oplus_{12}, \odot_{12})$, where $S=\{0, 2, 4, 6, 8, 10\}$.
23. Give an example of a non-commutative ring having exactly k elements, where
- $k = 16$
 - $k = 81$
 - $k = n^{m^2}$ for $m, n \in \mathbb{N}$.
24. In a commutative ring prove that the product of two idempotent elements is an idempotent.
25. Find all the idempotent elements of the rings
- $(\mathbb{Z}_6, \oplus_6, \odot_6)$
 - $(\mathbb{Z}_{12}, \oplus_{12}, \odot_{12})$
 - $(\mathbb{Z}_{20}, \oplus_{20}, \odot_{20})$
26. Prove that a non-zero idempotent cannot be nilpotent.
27. In a ring R , prove that the following conditions are equivalent.
- R has no non-zero nilpotent elements.
 - If $a \in R$ is such that $a^2 = 0$ then $a = 0$.
28. Prove that in a commutative ring R
- sum of two nilpotent elements is nilpotent.
 - If a is nilpotent then ar is nilpotent for all $r \in R$.
- Show by an example that the result fails to hold if the ring is non-commutative.
29. Describe all nilpotent elements of the following rings:
- \mathbb{Z}_4 (ii) \mathbb{Z}_8 (iii) \mathbb{Z}_{10} (iv) \mathbb{Z}_{12} (v) \mathbb{Z}_{16} (vi) \mathbb{Z}_{20} (vii) \mathbb{Z}_{36}
30. If a ring does not have any non-zero nilpotent element, prove that every idempotent element commutes with every element of R .
31. Find all the units of the following rings:
- \mathbb{Z} (ii) \mathbb{Z}_5 (iii) \mathbb{Z}_6 (iv) \mathbb{Z}_{15} (v) $\mathbb{Z} \oplus \mathbb{Z}$ (vi) $\mathbb{Q} \oplus \mathbb{Q}$ (vii) $M_2(\mathbb{Q})$ (viii) $\mathbb{Q}[\sqrt{2}]$.
32. Show that in a ring with unity, the sum of two units need not be a unit.
33. Find all the units of $\mathbb{Z}[i]$.
34. In a commutative ring R , let $a \in R$ is a unit and $b \in R$ is such that $b^2 = 0$, then prove that $a + b$ is a unit. Show by an example that result fails to hold when R is non-commutative.

35. Determine
 (i) $U(Q[x])$ (ii) $U(R[x])$.
36. If R is a ring in which for some $n \in \mathbb{Z}^+$, $x^n = x, \forall x \in R$, show that $xy = 0 \Rightarrow yx = 0$.
37. Let R be the set of all real valued functions defined on \mathbb{R} . Define suitable operations on R so that it is a ring.
38. Let $n > 1$ be a fixed integer.
 Let $\mathbb{Z}_n[i] = \{a + bi : a, b \in \mathbb{Z}_n\}$.
 Prove that $\mathbb{Z}_n[i]$ under the usual addition and multiplication of complex numbers, where real and imaginary parts are reduced modulo n , is a ring.
 Is it commutative? Does it have unity?

9.7 Subrings

As discussed in groups, the best way to study any algebraic structure is to study its subsets which themselves have the same structure with respect to the binary operations restricted to the subsets. Thus we study subsets of a ring which themselves are rings. Such subsets are called subrings.

Definition 9.8. (*Subring*):

Let $(R, +, \cdot)$ be a ring. A subset S of R is called a subring of R , if S is a ring under the operations of R restricted to S .

Example 9.13. If R is any ring, then $\{0\}$ and R are subrings of R . These are called the trivial subrings of R .

Example 9.14.

1. $(\mathbb{Z}, +, \cdot)$ is a subring of $(\mathbb{Q}, +, \cdot)$
2. $(\mathbb{Q}, +, \cdot)$ is a subring of $(\mathbb{R}, +, \cdot)$.
3. $(\mathbb{R}, +, \cdot)$ is a subring of $(\mathbb{C}, +, \cdot)$.
4. $(\mathbb{Z}, +, \cdot)$ is a subring of $(\mathbb{R}, +, \cdot)$.
5. $(2\mathbb{Z}, +, \cdot)$ is a sub ring of $(\mathbb{Z}, +, \cdot)$.

Is $(\mathbb{Z}_6, \oplus_6, \odot_6)$ a subring of $(\mathbb{Z}_8, \oplus_8, \odot_8)$?
 Clearly $\mathbb{Z}_6 \subseteq \mathbb{Z}_8$ but the binary operations of \mathbb{Z}_8 restricted to \mathbb{Z}_6 are not those of \mathbb{Z}_6 . So \mathbb{Z}_6 is not a subring of \mathbb{Z}_8 . In \mathbb{Z}_6 $4+2=0$ whereas $4+2=6$ in \mathbb{Z}_8 .

Criterion for a subset to be a subring

Theorem 9.4. Let R be a ring. A subset S of R is a subring of R if and only if

- (i) $0 \in S$.
- (ii) $a, b \in S \Rightarrow a - b \in S$.
- (iii) $a, b \in S \Rightarrow ab \in S$.

Proof: Let S be a subring R .

Then S is a ring in its own right, so that the conditions hold.

Conversely, let the conditions hold.

(i) $\Rightarrow S \neq \phi$.

Then (ii) $\Rightarrow (S, +)$ is a subgroup of $(R, +)$. Since associative law holds with respect to \cdot in R and $S \subseteq R$, therefore associative law holds in S . This together with (iii) $\Rightarrow (S, \cdot)$ is a semigroup. Also distributive law will hold in S as it holds in R .

Thus S itself is a ring and hence a subring of R . \square

Theorem 9.5. A non-empty subset S of a ring R is a subring if and only if

(i) $a - b \in S$, for all $a, b \in S$,

(ii) $ab \in S$, for all $a, b \in S$.

Proof: If S is a subring of R , then clearly the conditions hold. Conversely let S be a non-empty subset of R for which conditions (i) and (ii) hold.

$S \neq \phi \Rightarrow \exists a \in S$

By (i) $a - a \in S$ i.e. $0 \in S$.

By the above theorem S is a subring of R . \square

The above theorem gives to a very useful criterion to determine whether a non-empty subset of a ring is a subring.

Example 9.15. Consider the ring $M_2(\mathbb{Z})$ of all 2×2 matrices over \mathbb{Z} .

Let $S = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$

Clearly $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in S$, so that S is non-empty.

Let $A, B \in S$,

then $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, B = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$, for some $a, b, c, d \in \mathbb{Z}$.

$A - B = \begin{pmatrix} a - c & 0 \\ 0 & b - d \end{pmatrix} \in S$

$AB = \begin{pmatrix} ac & 0 \\ 0 & bd \end{pmatrix} \in S$

Hence S is a subring of $M_2(\mathbb{Z})$.

Definition 9.9. (Centre of a ring):

Let R be a ring. The centre of R , is the set of all those element of R which commute with every element of R . It is denoted by $Z(R)$. Thus

$Z(R) = \{x \in R \mid xr = rx, \forall r \in R\}$.

Theorem 9.6. The center of a ring R is a commutative subring of R .

Proof: Let R be any ring. Then the centre of R is

$Z(R) = \{x \in R \mid xr = rx, \forall r \in R\}$

Since $0r = r0, \forall r \in R$

$\therefore 0 \in Z(R)$. Hence $Z(R) \neq \phi$.

Let $x, y \in Z(R)$. Then

$xr = rx, \forall r \in R$...(i)

and $yr = ry, \forall r \in R$...(ii)

For any $r \in R$

$$\begin{aligned}(x - y)r &= xr - yr \\ &= rx - ry \quad \text{using (i) and (ii)} \\ &= r(x - y)\end{aligned}$$

$$\therefore x - y \in Z(R).$$

For any $r \in R$

$$\begin{aligned}(xy)r &= x(yr) \quad \text{by associative law} \\ &= x(ry) \quad \text{using (ii)} \\ &= (xr)y \\ &= (rx)y \quad \text{using (i)} \\ &= r(xy)\end{aligned}$$

$$\therefore (xy)r = r(xy) \quad \forall r \in R.$$

so that $xy \in Z(R)$.

Hence $Z(R)$ is a subring of R .

Let $x, y \in Z(R)$.

Then x commutes with every element of R . In particular it commutes with y

$$\therefore xy = yx.$$

Hence $Z(R)$ is commutative subring of R . □

Theorem 9.7. *The intersection of two subrings is a subring.*

Proof: Let R be a ring and S_1, S_2 two subrings of R . Let $S = S_1 \cap S_2$.

Since $0 \in S_1$ and $0 \in S_2$

$\therefore 0 \in S_1 \cap S_2 = S$, so that S is non-empty.

Let $a, b \in S$. Then $a, b \in S_1$ and $a, b \in S_2$.

Since S_1 and S_2 are subrings.

$\therefore a - b, ab \in S_1$ and $a - b, ab \in S_2$

Hence $a - b, ab \in S_1 \cap S_2 = S$, so that S is a subring of R . □

The union of two subrings need not be a subring as is seen by the following example.

Example 9.16. *Consider the ring \mathbb{Z} . Then $3\mathbb{Z}, 4\mathbb{Z}$ are subrings of \mathbb{Z} . Let $S = 3\mathbb{Z} \cup 4\mathbb{Z}$.*

Now $3 \in 3\mathbb{Z} \subseteq S, 4 \in 4\mathbb{Z} \subseteq S$.

$\therefore 3, 4 \in S$ but $3 + 4 = 7 \notin S$, as $7 \notin 3\mathbb{Z}$ and $7 \notin 4\mathbb{Z}$.

Thus S is not closed under addition so that S is not a subring of R .

Theorem 9.8. *Every subring of \mathbb{Z} is of the form $n\mathbb{Z}$ for some $n \in \mathbb{Z}$.*

Proof: Let S be a subring of \mathbb{Z} . Then $(S, +)$ is a subgroup of $(\mathbb{Z}, +)$, so that $S = n\mathbb{Z}$ for some $n \in \mathbb{N}$.

Hence proved. □

Corollary 9.9. *Every subring of \mathbb{Z}_n is of the form $k\mathbb{Z}_n$ for some $k \in \mathbb{Z}_n$.*

Problem 9.13. *\mathbb{Z} is a subring of $\mathbb{Z}[x]$.*

Solution: Clearly $\mathbb{Z} \subseteq \mathbb{Z}[x]$ as each element of \mathbb{Z} can be regarded as constant polynomial.

Since the zero polynomial namely 0 , is an element of \mathbb{Z} .

$\therefore \mathbb{Z} \neq \phi$.

Let $m, n \in \mathbb{Z}$. Then $m - n, mn \in \mathbb{Z}$, so that \mathbb{Z} is a subring of $\mathbb{Z}[x]$.

A subring inherits some of the properties of the ring. The properties which relate to the binary operations are inherited, whereas those which relate to the existence of certain types of elements are not inherited.

This is given below.

Theorem 9.10. *A subring of a commutative ring is commutative.*

Proof: Let S be a subring of a commutative ring R . Let $a, b \in S$. Then $a, b \in R$, so that $ab = ba$. So S is commutative. \square

A subring of a non-commutative ring need not be non-commutative as is seen in the following example.

Example 9.17. *Consider $M_2(\mathbb{Z})$ and let*

$$S = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$$

Then S is a commutative subring of $M_2(\mathbb{Z})$.

The existence of identity element in a ring and its subring are independent of each other as is seen by the following examples.

Example 9.18. *If a ring has unity its subring may not have unity. Consider the ring \mathbb{Z} with unity 1. $2\mathbb{Z}$ is a subring of \mathbb{Z} and it does not have unity.*

A subring may have unity but the ring may not have unity as is seen in the following example.

Example 9.19. *Let $R = \left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$*

Then R is the ring without unity.

Let $S = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$

Then S is a subring of R .

If $E = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ then $EA = AE = A$ for all $A \in S$, so that E is the unity of S .

Thus S is a subring of R with unity.

Even if both the ring and its subring have unity, they may be different, as is seen in the following example.

Example 9.20. *Consider $M_2(\mathbb{Q})$, the ring of all 2×2 matrices over \mathbb{Q} .*

Let $S = \left\{ \begin{pmatrix} a & a \\ a & a \end{pmatrix} \mid a \in \mathbb{Q} \right\}$

Then, S is a subring of $M_2(\mathbb{Q})$.

$M_2(\mathbb{Q})$ has identity, namely $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

If $E = \begin{pmatrix} e & e \\ e & e \end{pmatrix}$ is the identity of S

then $\begin{pmatrix} e & e \\ e & e \end{pmatrix} \begin{pmatrix} a & a \\ a & a \end{pmatrix} = \begin{pmatrix} a & a \\ a & a \end{pmatrix}$

so that $2ae = a \quad \forall a \in \mathbb{Q}$.

Hence $e = \frac{1}{2}$

Thus $E = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \in S$ is the identity of S .

Both S and $M_2(\mathbb{Q})$ have identity elements but they are different.

Since a unit element depends upon the identity element, we would like to know the relationship between the units of a ring and its subring. In fact we are in for a surprise. Even if the ring and subring have the same identity element, they may not have the same units.

Theorem 9.11. *Let S be a subring of a ring R with the same identity element as that of R . If an element of S is a unit in S then it is also a unit in R .*

Proof: Let 1 be the identity element of both R and S . Let $u \in S$ be a unit in S . Then there exists $v \in S$ such that $uv = vu = 1$... (i)
Also $u \in R$ and so by (i) u is a unit in R . □

If S is a subring of R and both have the same identity element 1 , then an element $r \in S$ may be a unit in R , but not in S . This is shown on the following example.

Example 9.21. \mathbb{Z} is a subring of \mathbb{Q} . Both have the same identity 1 . In Theorem 9.11, $2 \in \mathbb{Z}$ is not a unit in \mathbb{Z} , but it is a unit in \mathbb{Q} .

Moreover, if $1_S \neq 1_R$, a unit of S may be not a unit of R . For example:

$$\text{Let } R = M_2(\mathbb{Q}), S = \left\{ \begin{pmatrix} a & a \\ a & a \end{pmatrix} \mid a \in \mathbb{Q} \right\}.$$

$$\text{then } 1_R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, 1_S = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\text{Every non-zero element of } S \text{ is an unit in } S, \text{ since } \begin{pmatrix} a & a \\ a & a \end{pmatrix} \begin{pmatrix} \frac{1}{4a} & \frac{1}{4a} \\ \frac{1}{4a} & \frac{1}{4a} \end{pmatrix} = 1_S$$

whereas $\begin{pmatrix} a & a \\ a & a \end{pmatrix}$ is not a unit in R as this is a singular matrix.

9.8 Exercise

1. Verify whether S is a subring of the ring R . If not, state which condition fails to hold

(i) $R = (\mathbb{Z}, +, \cdot), S = \mathbb{N}$

(ii) $R = (M_{2 \times 2}(\mathbb{Z}), +, \cdot), S =$ set of all 2×2 non-singular matrices.

(iii) R as in (ii), $S = \left\{ \begin{pmatrix} a & a \\ a & a \end{pmatrix} \mid a \in \mathbb{Z} \right\}$

(iv) $R = (M_2(\mathbb{C}), +, \cdot) S = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a, b \in \mathbb{C} \right\}$

(v) Let F be the set of all real valued continuous functions defined on $[0, 1]$ under pointwise addition and multiplication. $R = (F, +, \cdot)$
 $S = \{f \in F \mid f(\frac{1}{2}) = 0\}$

(vi) R as in (v)
 $S = \{f \in F \mid f(\frac{1}{2}) = 1\}$

2. Which of the following are subrings of \mathbb{Q} .
- $\{\frac{p}{q} \in \mathbb{Q} \mid q \text{ is odd}\}$
 - $\{\frac{p}{q} \in \mathbb{Q} \mid q \text{ is even}\}$
 - $\{\frac{p}{q} \in \mathbb{Q} \mid p \text{ is odd}\}$
 - $\{\frac{p}{q} \in \mathbb{Q} \mid p \text{ is even}\}$
 - $\{\frac{p}{q} \in \mathbb{Q} \mid \frac{p}{q} = (\frac{r}{s})^2 \text{ for some } \frac{r}{s} \in \mathbb{Q}\}$.
3. Let R be the ring of all real valued functions defined on \mathbb{R} , under pointwise addition and multiplication. Determine whether the following subsets of R are subrings.
- S_1 = set of all continuous functions.
 - S_2 = set of all polynomial functions.
 - S_3 = set of all functions which are zero at finitely many point together with the zero function.
 - S_4 = set of all functions which are zero at infinite number of points.
 - S_5 = $\{f \in R \mid f(x) = 0 \text{ if } x \text{ is rational}\}$
- In the above question find all relations of containment within those S_i which are subrings.
4. Let $S = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$
 Prove that
- S is a subring of $M_2(\mathbb{Z})$.
 - S has a left unity but no right unity.
 - S has an infinite number of distinct left unit elements.
5. Let R be a ring and S_1, S_2, S_3 subrings of R . If $S_3 \subseteq S_1 \cup S_2$ show that $S_3 \subseteq S_1$ or $S_3 \subseteq S_2$.
6. Prove that $m\mathbb{Z}$ is a subring of $n\mathbb{Z}$ if and only if n divide m .
7. Let a be a element of ring R and let $S = \{x \in R \mid ax = 0\}$. Show that S is a subring of R .
 (S is called the right annihilator of a).
8. Let R be a ring with unity 1_R . If S is a subring of a ring R such that $1_R \in S$, then prove that S has unity 1_R .
9. Let 'a' be an element of a ring R with unity 1 such that $a^2 = 1$. If $S = \{ara \mid r \in R\}$, prove that S is a subring of R . Does $1 \in S$? Is it the unity of S ?
10. Determine the smallest subring of \mathbb{Q} containing a where
- $a = \frac{1}{3}$
 - $a = \frac{2}{5}$
11. Prove that the set of all nilpotent elements of a commutative ring form a subring.
 Prove that $S = \left\{ \begin{pmatrix} n & 0 \\ 2n & 0 \end{pmatrix} \mid n \in \mathbb{Z} \right\}$ is a commutative subring of $M_2(\mathbb{Z})$ with unity. Are the unities $M_2(\mathbb{Z})$ and S same?
12. Give an example of a subset S of a ring R such that S is a group under addition but S is not a subring of R .

13. Prove that the property of being a subring is transitive.
14. Find n such that $9\mathbb{Z} \cap 12\mathbb{Z} = n\mathbb{Z}$.

9.9 Integral Domains and Fields

Rings were introduced to put the algebraic properties of integers into an abstract setting. But many of the properties of integers, like commutativity with respect to multiplication, existence of multiplicative identity and the product of two non-zero integers being non-zero, were not taken into consideration. Now we introduce other algebraic structures which have these properties.

Definition 9.10. (Zero Divisor):

Let R be a ring. A non-zero element $a \in R$ is said to be

- (i) a left divisor of zero if there exists a non-zero element $b \in R$ such that $ab = 0$.
- (ii) a right divisor of zero if there exists a non-zero element $c \in R$ such that $ca = 0$.
- (iii) a divisor of zero, if it is either a left divisor or a right divisor of zero.

If R is a commutative ring then every left divisor of zero is a right divisor of zero and vice versa.

Definition 9.11. (Ring without zero divisors): A ring R is called a ring without zero divisors, if R has no zero divisors.

Definition 9.12. (Integral domain):

A commutative ring with unity is said to be an integral domain if it is without zero divisors.

Example 9.22. $(\mathbb{Z}, +, \cdot)$ is an integral domain.

Example 9.23. $(\mathbb{Z}_5, \oplus_5, \odot_5)$ is an integral domain.

Example 9.24. $2\mathbb{Z}$ is a ring without zero divisor but it does not have unity. So it is not an integral domain.

Example 9.25. $(\mathbb{Z}_6, \oplus_6, \odot_6)$ is not an integral domain as it is with zero divisor, since $2, 3$ are non-zero element of \mathbb{Z}_6 but $2 \odot_6 3 = 0$.

Definition 9.13. (Division ring):

A ring with unity is called a division ring (or a skew field) if every non-zero element has a multiplicative inverse.

Example 9.26. $(\mathbb{Q}, +, \cdot)$ is a division ring.

Definition 9.14. (Field):

A commutative ring with unity is called a field if every non-zero element has a multiplicative inverse.

Example 9.27. $(\mathbb{Q}, +, \cdot)$ is a field.

Example 9.28. $(\mathbb{Z}_5, \oplus_5, \odot_5)$ is a field.

From the definition it is clear that a commutative division ring is a field. From the definition it follows that:

1. A subring of a ring without zero divisors is also without zero divisors.
2. A subring of an integral domain need not to be an integral domain.
3. A subring of a division ring may not be a division ring.
4. A subring of a field may not be a field.

9.10 Examples

Example 9.29.

1. The ring \mathbb{Z} is an integral domain. Since 2 does not have a multiplicative inverse, therefore \mathbb{Z} is not a field.
2. The ring \mathbb{Q} is an integral domain. Also every non-zero element has an inverse. Therefore it is a field.
3. The ring \mathbb{R} of real numbers and the ring \mathbb{C} of complex numbers are fields.

Example 9.30. The ring $2\mathbb{Z}$ of even integers is a commutative ring without zero divisors.

Similarly $2\mathbb{Z}, 4\mathbb{Z}$ etc. are all commutative rings without zero divisors.

If $n \in \mathbb{N}, n > 1$ then $n\mathbb{Z}$ does not have unity. It is a commutative ring without zero divisors. Thus these are not integral domains.

Example 9.31. $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Z}\}$

$\mathbb{Z}[\sqrt{2}]$ is a commutative ring, with unity 1.

It is without zero divisors as $\mathbb{Z}[\sqrt{2}] \subseteq \mathbb{R}$ and \mathbb{R} is without zero divisors thus $\mathbb{Z}[\sqrt{2}]$ is without zero divisors, so that it is an integral domain. Since $2 \in \mathbb{Z}[\sqrt{2}]$ does not have an inverse, therefore $\mathbb{Z}[\sqrt{2}]$ is not a division ring.

Example 9.32. $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Q}\}$.

Since $\mathbb{Q}[\sqrt{2}]$ is a subring of \mathbb{R} , therefore it is commutative ring without zero divisors.

Also $1 = 1 + 0\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$, thus $\mathbb{Q}[\sqrt{2}]$ has unity.

If $0 \neq x = a + b\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$, then

$\frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2}\sqrt{2}$ is the inverse of x (explain why $a^2 - 2b^2 \neq 0$).

Thus, every non-zero element of $\mathbb{Q}[\sqrt{2}]$ is invertible. Hence $\mathbb{Q}[\sqrt{2}]$ is a field.

Example 9.33. Consider $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}$ that $\mathbb{Z}[i]$ is a commutative ring with unity is seen before. It is without zero divisor follows from the fact that it is a subring of \mathbb{C} . Thus $\mathbb{Z}[i]$ is an integral domain. $\mathbb{Z}[i]$ is not a field as $2 \in \mathbb{Z}[i]$ does not have an inverse in $\mathbb{Z}[i]$.

Thus $\mathbb{Z}[i]$ is an integral domain which is not a field.

Example 9.34. $\mathbb{Q}[i] = \{a + ib | a, b \in \mathbb{Q}\}$.

$\mathbb{Q}[i]$, being a subring of \mathbb{C} it is a commutative ring without zero divisors. Also $1 \in \mathbb{Q}[i]$ is the unity of the $\mathbb{Q}[i]$. If $0 \neq x = a + ib \in \mathbb{Q}[i]$ then $\frac{a}{a^2 + b^2} + i\frac{-b}{a^2 + b^2}$ is the inverse of x . Hence $\mathbb{Q}[i]$ is a field.

Examples from rings of residues

Example 9.35. Consider $(\mathbb{Z}_4, \oplus_4, \odot_4)$. It is a commutative ring with unity. Also $2 \odot_4 2 = 0$, so that \mathbb{Z}_4 is with zero divisors. Hence it is not an integral domain.

It is also not a field as $2 \in \mathbb{Z}_4$ does not have an inverse. Consequently \mathbb{Z}_4 is neither a division ring nor a field.

Example 9.36. Consider $(\mathbb{Z}_5, \oplus_5, \odot_5)$.

\mathbb{Z}_5 is a commutative ring with unity and without zero divisors. Let $a, b \in \mathbb{Z}_5$ such that $a \odot_5 b = 0$.

Then ab is a multiple of 5 i.e. $5|ab$

$\Rightarrow 5|a$ or $5|b$ as 5 is a prime.

$\Rightarrow a = 0$ or $b = 0 \quad \because 0 \leq a, b < 5$.

Thus \mathbb{Z}_5 is without zero divisors, so that \mathbb{Z}_5 is an integral domain.

Also every non-zero element is invertible, so that $(\mathbb{Z}_5, \oplus_5, \odot_5)$ is a field.

Example 9.37. If p is prime $(\mathbb{Z}_p, \oplus_p, \odot_p)$ is an integral domain.

As seen earlier $(\mathbb{Z}_n, \oplus_n, \odot_n)$ is commutative ring with unity. We prove that it is without zero divisors:

Let $a, b \in \mathbb{Z}_p$ such that $a \odot_p b = 0$.

Then ab is a multiple of p .

$\Rightarrow p|ab$

$\Rightarrow p|a$ or $p|b$ as p is prime

$\Rightarrow a = 0$ or $b = 0 \quad \because a, b \in \mathbb{Z}_p$.

Hence $a \odot_p b = 0 \implies a = 0$ or $b = 0$.

$\therefore \mathbb{Z}_p$ is without zero divisors, hence is an integral domain.

Example 9.38. Consider $(4\mathbb{Z}_{12}, \oplus_{12}, \odot_{12})$.

Then $4\mathbb{Z}_{12} = \{0, 4, 8\}$. The multiplication tables are:

\oplus_{12}	0	4	8	\odot_{12}	0	4	8
0	0	4	8	0	0	0	0
4	4	8	0	4	0	4	8
8	8	0	4	8	0	8	4

Being a subring of $(\mathbb{Z}_{12}, \oplus_{12}, \odot_{12})$, $4\mathbb{Z}_{12}$ is a commutative ring. Also from the table it is clear that 4 is the identity element and $4^{-1} = 4, 8^{-1} = 8$, so that every non-zero element has an inverse.

Thus $(4\mathbb{Z}_{12}, \oplus_{12}, \odot_{12})$ is a field. Moreover, product of two non-zero elements is non-zero, therefore it is also an integral domain.

Examples from Matrices

Example 9.39. Let $M = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{C} \right\}$

then M is a subring of $M_{2 \times 2}(\mathbb{C})$. It has unity $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

M is non-commutative, as

$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, B = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \in M$ and $AB \neq BA$.

We now prove that every non-zero element of M is invertible

Let $0 \neq A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \in M$, if $B = \frac{1}{|a|^2 + |b|^2} \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix}$

Then $AB = BA = I$, so that $A^{-1} = B$.

This shows that M is a division ring. Since M is not commutative.

$\therefore M$ is not a field.

Example 9.40. Consider $M = \left\{ \begin{pmatrix} a & a \\ a & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$

It is verified that M is a commutative subring of $M_2(\mathbb{R})$,

$E = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ is the unity of M .

If $0 \neq A \in M$, then $A = \begin{pmatrix} a & a \\ a & a \end{pmatrix}$ for some $0 \neq a \in \mathbb{R}$, and

$B = \begin{pmatrix} \frac{1}{4a} & \frac{1}{4a} \\ \frac{1}{4a} & \frac{1}{4a} \end{pmatrix}$ is the inverse of A . Thus M is a field.

Example 9.41. Let F be the set of all $n \times n$ scalar matrices with real entries.

It can be easily verified that F is a field.

Example from Quaternions

Example 9.42. Consider the ring D of real quaternions. D is a non-commutative ring with unity.

We now show that every non-zero element of D is invertible.

Let $x \in D$, then $0 \neq x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$,

for some $\alpha_0, \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ not all zero.

If $y = \frac{1}{\beta}(\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k)$, where $\beta = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \neq 0$

then $xy = 1$.

Hence every non-zero element of D is invertible, so that D is a division ring.

Since D is a non-commutative. $\therefore D$ is not a field.

Theorem 9.12. A commutative ring R with unity is an integral domain if and only if cancellation law holds in R .

Proof: Let R be a commutative ring with unity.

Suppose that R is an integral domain. We prove that cancellation law holds in R .

Let $a, b, c \in R, a \neq 0$ such that

$ab = ac$.

Then $a(b - c) = 0 \dots$ (i)

Since R is an integral domain. It is without zero divisors,

so that (i) $\Rightarrow a = 0$ or $b - c = 0$

Since $a \neq 0 \therefore b - c = 0 \Rightarrow b = c$.

Hence $ab = ac$ and $a \neq 0 \Rightarrow b = c$. Thus left cancellation law holds in R .

Commutativity in R implies that right cancellation law also holds.

Conversely, let cancellation laws hold in R .

Let $a, b \in R$ such that $ab = 0$. If $a = 0$, result is proved. Let $a \neq 0$ then $ab = 0 = a \cdot 0$. By cancellation law, $ab = 0 \Rightarrow a = 0$ or $b = 0$, so that R is an integral domain. \square

Theorem 9.13. In a field cancellation laws hold.

Proof: Let R be a field and $a, b, c \in R, a \neq 0$ such that $ab = ac$. Since $a \neq 0$, a^{-1} exists.

Pre-multiplying by a^{-1} , we get

$$\begin{aligned} a^{-1}(ab) &= a^{-1}(ac) \\ \Rightarrow (a^{-1}a)b &= (a^{-1}a)c \\ \Rightarrow 1b &= 1c \\ \Rightarrow b &= c \end{aligned}$$

Thus $ab = ac, a \neq 0 \Rightarrow b = c$. So that left cancellation law holds. Right cancellation law follows similarly. \square

Corollary 9.14. *A field is without zero divisors.*

Proof: See proof of Theorem 9.12. \square

Corollary 9.15. *Every field is an integral domain.*

Proof: Follow from corollary 9.14. \square

The converse of the above corollary is not true as is seen by the following example.

Example 9.43. *An integral domain may not be a field.*

\mathbb{Z} is an integral domain. Since 2 does not have an inverse, therefore \mathbb{Z} is not a field.

Theorem 9.16. *A finite commutative ring without zero divisors is a field.*

Proof: Let $R = \{a_1, a_2, \dots, a_n\}$ be a finite commutative ring without zero divisors. We prove that

- (i) R has unity.
- (ii) every nonzero element has a multiplicative inverse.

Step 1 let $0 \neq a \in R$, and $aR = \{aa_1, aa_2, \dots, aa_n\}$.

Clearly $aR \subseteq R$.

All the elements of aR are distinct,

for if $aa_i = aa_j$

then $a(a_i - a_j) = 0$

so that $a_i - a_j = 0$ as $a \neq 0$ and R is without zero divisors.

$\therefore a_i = a_j$ a contradiction. Hence aR has n elements.

Since $aR \subseteq R$ and aR and R are finite sets having the same number of elements,

so that $aR = R$.

Now $a \in R = aR$,

\therefore there exists $a_{i_0} \in R$

such that $aa_{i_0} = a$.

We assert that a_{i_0} is the identity element of R .

Let $x \in R$. Then $x = aa_k$ for some $a_k \in R$.

$$\begin{aligned} xa_{i_0} &= (aa_k)a_{i_0} \\ &= a(a_ka_{i_0}) \\ &= a(a_{i_0}a_k) \\ &= (aa_{i_0})a_k \quad \text{for all } x \in R. \\ &= (aa_k) \\ &= x \end{aligned}$$

$\therefore xa_{i_0} = x$,

As R is commutative $\therefore xa_{i_0} = a_{i_0}x = x \quad \forall x \in R$.
 $\therefore a_{i_0}$ is the identity of R . We denote it by e .

Step 2 Let $0 \neq a \in R$.

Considering aR as in step 1, we get

$R = aR$. Now $e \in R = aR$.

\therefore there exists some $a_i \in R$ such that

$$e = aa_i$$

$$\therefore aa_i = a_i a = e.$$

Hence a_i is the multiplicative inverse of a . Thus every non-zero element of R has a multiplicative inverse. Therefore R is a field. \square

Corollary 9.17. *A finite integral domain is a field.*

Using the above theorem we have the following result.

Corollary 9.18. *A finite ring without zero divisors is a division ring.*

Theorem 9.19. *\mathbb{Z}_n is a field if and only if n is prime.*

Proof: Let n be prime. Then \mathbb{Z}_n is a finite integral domain. By Corollary 9.17, \mathbb{Z}_n is a field. Conversely, let \mathbb{Z}_n be a field then \mathbb{Z}_n is an integral domain. If n is composite then $n = m_1 m_2$, where $0 < m_1, m_2 < n$, so that $m_1, m_2 \in \mathbb{Z}_n$ such that $m_1 \odot_n m_2 = 0$, thus \mathbb{Z}_n is with zero divisors, which contradicts the fact that \mathbb{Z}_n is an integral domain. Thus n is prime. \square

Remark 9.2. *An independent proof of the result that if p is prime then $(\mathbb{Z}_p, \oplus_p, \odot_p)$ is a field is given below:*

Example 9.44. *$(\mathbb{Z}_p, \oplus_p, \odot_p)$, where p is prime is a field.*

It has been seen before that $(\mathbb{Z}_p, \oplus_p, \odot_p)$ is a commutative ring with unity.

We prove that every non-zero element of \mathbb{Z}_p has an inverse.

Let $0 \neq a \in \mathbb{Z}_p$. Then $1 \leq a < p$, so that $(a, p) = 1$, as p is prime.

Hence by Euclid's algorithm there exists $m, n \in \mathbb{Z}$ such that

$$am + pn = 1 \Rightarrow am \equiv 1 \pmod{p} \text{ or that } am \pmod{p} = 1.$$

Let $m \pmod{p} = r$.

Then $r \in \mathbb{Z}_p$

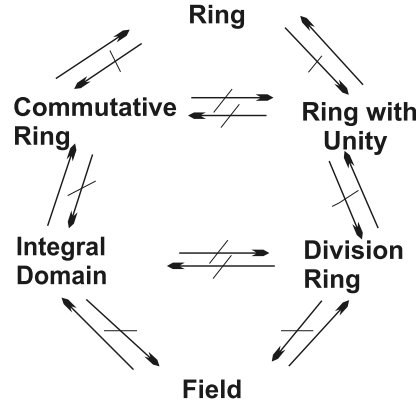
$$\text{Now } am \pmod{p} = 1$$

$$\Rightarrow a \pmod{p} \odot_p m \pmod{p} = 1$$

$$\Rightarrow a \odot_p r = 1$$

$\Rightarrow r \in \mathbb{Z}_p$ is the inverse of a .

The relationship between the different types of rings is given in the following diagram.



Problem 9.14. Let R be the ring of all real valued functions defined on $[0,1]$ under pointwise addition and pointwise multiplication.

- (i) Find all units of R .
 (ii) Find all the zero divisors of R .

Solution: R has unity, namely the function e defined by $e(x) = 1 \quad \forall x \in [0,1]$.

(i) Let $f \in R$ be a unit. Then there exists $g \in R$ such that $fg = e$.

$$\implies (fg)(x) = e(x) \quad \forall x \in [0,1]$$

$$\implies f(x)g(x) = 1 \quad \forall x \in [0,1]$$

$$\implies \text{for all } x \in [0,1], f(x) \neq 0$$

$$\text{and } g(x) = \frac{1}{f(x)}.$$

Thus if f is a unit, then $f(x) \neq 0 \quad \forall x \in [0,1]$.

On the other hand, let $f \in R$ be such that $f(x) \neq 0 \quad \forall x \in [0,1]$. Define $g(x) = \frac{1}{f(x)} \quad \forall x \in [0,1]$. Then $fg = gf = e$, so that f is a unit.

$$\therefore U(R) = \{f \in R \mid f(x) \neq 0 \quad \forall x \in [0,1]\}.$$

(ii) Let $f \in R$ be a zero divisor.

Then f is not a unit, so that $f(a) = 0$ for some $a \in [0,1]$.

$$\text{Define } g \text{ such that } g(x) = \begin{cases} 0 & \text{when } f(x) \neq 0 \\ 1 & \text{when } f(x) = 0 \end{cases}$$

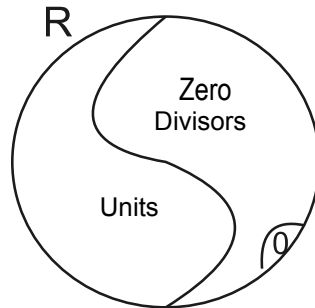
Then $g(a) = 1$, so that $g \neq 0$.

For any $x \in [0,1]$

$$(fg)(x) = f(x)g(x) = 0.$$

Thus the set of zero divisors = $\{0 \neq f \in R \mid f(a) = 0 \text{ for some } a \in [0,1]\}$.

Remark 9.3. The above problem shows that every non-zero element of R is either a unit or a zero divisor.



Can we generalize the above result to any ring? Certainly not!
 This is not the case if R is the set of all continuous real valued functions on $[0,1]$.
 In fact there are non-zero functions which are neither units nor zero divisors.

Problem 9.15. Let R be the ring of all real valued continuous functions defined on $[0,1]$ under pointwise addition and pointwise multiplication.

- (i) Find units of R .
- (ii) Find 5 zero divisors of R .
- (iii) Find 5 elements of R which are not zero divisors.

Solution:

- (i) $U(R) = \{f \in R \mid f(x) \neq 0 \forall x \in [0,1]\}$.
 Proof is similar to (i) of the above problem.

- (ii) Let $0 < a < 1$
 Define f as follows:

$$f(x) = \begin{cases} 0 & \text{if } x \in [0, a] \\ 2x & \text{if } x \in (a, 1] \end{cases}$$

Then $f \in R$. Now define

$$g(x) = \begin{cases} 3x & \text{if } x \in [0, a] \\ 0 & \text{if } x \in (a, 1] \end{cases}$$

Then $g \in R$ and
 $(fg)(x) = f(x)g(x)$
 $= 0 \quad \forall x \in [0,1]$
 $\therefore fg = 0, f \neq 0, g \neq 0$.

Thus f is a zero divisor for every a , such that $0 < a < 1$.
 Taking $a = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{2}{3}$ we get 5 zero divisors in R .

- (iii) Let $0 < a < 1$.
 Define a function f as follows:
 $f(x) = |x - a|, x \in [0,1]$
 Then $0 \neq f \in R$ and $f(a) = 0$.
 Suppose f is a zero divisor. Then there exists $0 \neq g \in R$, such that $fg = 0$
 so that $f(x)g(x) = 0 \quad \forall x \in [0,1]$
 Since $f(x) \neq 0$, when $x \neq a$, therefore $g(x) = 0$ for $x \neq a$.

Since g is a continuous, we must have that $g(a)=0$. But then $g = 0$, which is a contradiction.

Hence f is not a zero divisor.

Taking $a = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{2}{3}$, we get the five element of R which are not zero-divisors.

Remark 9.4. Our method of construction of the functions show that

(i) We can find infinitely many zero divisors in R .

(ii) A continuous function which is zero only at finitely (or countably) many points cannot be a zero divisor.

Example 9.45. We know that when n is prime $(\mathbb{Z}_n, \oplus_n, \odot_n)$ is a field. If n is composite we find a subring of \mathbb{Z}_n which is a field.

Let p be a prime such that p divides n , but p^2 does not divide n . Let $n = pm$ then $(p, m) = 1$.

$(m\mathbb{Z}_n, \oplus_n, \odot_n)$ is a subring of $(\mathbb{Z}_n, \oplus_n, \odot_n)$ having p elements.

$m\mathbb{Z}_n = \{0, m, 2m, \dots, (p-1)m\} = \{am \mid 0 \leq a \leq p-1\}$. Let us find the unity of $m\mathbb{Z}_n$.

Let, if possible, $am \in m\mathbb{Z}_n$ be the unity, then

$$am \odot_n bm = bm, \forall bm \in m\mathbb{Z}_n.$$

$\therefore (am)(bm) - bm$ is a multiple of n

$$\text{i.e. } bm(am - 1) = kn = kmp$$

$$\text{i.e. } b(am-1) = kp$$

$$\text{i.e. } p \mid b(am - 1)$$

$$\text{i.e. } p \mid (am - 1) \quad \because p \text{ is a prime and } (p, b) = 1$$

$$\text{i.e. } am \equiv 1 \pmod{p}$$

Thus the unity of $m\mathbb{Z}_n$ is am , if $am \equiv 1 \pmod{p}$.

9.11 Exercise

- Let R be a field. Is every subring of R a field? Justify.
- Let R be a ring which is not a field. Can a subring of R be a field? Justify.
- List all the zero divisors and units of
(i) \mathbb{Z}_{12} (ii) \mathbb{Z}_8 (iii) \mathbb{Z}_{20}
Do you see any relation between the set of units and the set of zero divisors of a ring?
- Let R be an integral domain. Is every subring of R an integral domain? Justify.
- Let R be a ring which is not an integral domain. Can a subring of R be an integral domain? Justify.
- In a ring with unity and without zero divisors, prove that the only idempotents are zero and unity.
- In a ring with unity prove that
(i) a unit can not be a zero divisor.
(ii) a zero divisor can not be a unit.

8. Prove that the direct sum of two integral domains need not be an integral domain.
9. Give an example of a division ring R which is not a field. Find a subring of R which is not a division ring.
10. Let R be a ring with unity. Prove that
 - (i) if u has a right inverse then u is not a right zero divisor.
 - (ii) if u has more than one right inverse then u is a left zero divisor.
11. Let R be the ring of all real valued functions defined on $[a, b]$ under pointwise addition and pointwise multiplication.
 - (i) Find all units of R .
 - (ii) Find all the zero divisor of R .
12. Let r be a nilpotent element of a commutative ring R . Prove that
 - (i) r is either zero or a zero divisor.
 - (ii) if R has unity 1, then $1 + r$ and $1 - r$ are units.
13. Let R be the set of all sequences of integers (a_1, a_2, a_3, \dots) where all except finitely many a_i are zero. Prove that
 - (i) R is a ring.
 - (ii) R is a commutative.
 - (iii) R does not have unity.

9.12 Solved Problems

Problem 9.16. R is a ring with unity e . If every non-zero element of R has a unique right inverse, prove that R is a division ring.

Solution: Let $0 \neq x \in R$. Suppose there exists a unique right inverse y of x , i.e. $xy = e$.

$$\begin{aligned} \text{Now } x(y + yx - e) &= xy + xyx - xe \\ &= e + ex - x \\ &= e + x - x = e \end{aligned}$$

$\therefore y + yx - e$ is also a right inverse of x . By uniqueness of the right inverse, we have $y + yx - e = y \implies yx - e = 0 \implies yx = e$

Hence y is also a left inverse of x , so that inverse of x is unique, namely y .

Thus each non zero element has an inverse.

$\therefore R$ is a division ring.

Problem 9.17. Let R be a ring such that $x^3 = x$, for all $x \in R$. Then R is a commutative ring.

Solution: We have $x^3 = x \quad \forall x \in R \dots (1)$

Step 1 Let $x, y \in R$, then $x^2y - x^2yx^2 \in R$.

$$\begin{aligned} \text{Also } (x^2y - x^2yx^2)^2 &= (x^2y - x^2yx^2)(x^2y - x^2yx^2) \\ &= 0 \quad \text{on using distributive law and (1)} \end{aligned}$$

$$\begin{aligned} \therefore (x^2y - x^2yx^2)^3 &= 0 \\ \Rightarrow x^2y - x^2yx^2 &= 0 \quad \text{using (1)} \\ \Rightarrow x^2y &= x^2yx^2 \end{aligned}$$

Similarly $yx^2 = x^2yx^2$

Thus $x^2y = yx^2 \dots (2)$.

Step 2 For $x \in R$, $x^2 - x \in R$, so that

$$(x^2 - x)^3 = x^2 - x \quad \text{by (1)}$$

On simplifying, we get

$$3x^2 = 3x \dots (3)$$

$$\text{Also } (x^2 - x)^2 = 2x^2 - 2x \quad \text{using (1)}$$

$$= x - x^2 \quad \text{using (3)}$$

$$\therefore (x^2 - x)^2 = x - x^2 \dots (4)$$

Step 3 Let $x, y \in R$

$$\text{Then } (x^2 - x)^2 y = y(x^2 - x)^2 \quad \text{by (2)}$$

$$\text{so that } (x^2 - x)y = y(x^2 - x) \quad \text{by (4)}$$

Hence $xy = yx$.

Thus R is commutative.

Problem 9.18. Let S be any set and $\mathcal{P}(S)$ the power of S . Describe the units and zero divisors of $\mathcal{P}(S), \Delta, \cap$. Under what conditions is $\mathcal{P}(S)$ a field?

Solution: The unity of the ring $\mathcal{P}(S)$ is S and zero element is ϕ .

To find units:

If $A \in \mathcal{P}(S)$ is a unit, then for some $B \in \mathcal{P}(S)$

$$A \cap B = S,$$

so that $A=B=S$.

Hence S is the only unit in $\mathcal{P}(S)$.

To find zero divisors:

If $A \in \mathcal{P}(S)$ is a zero divisor, then for some $B \in \mathcal{P}(S)$, we must have

$$A \cap B = \phi.$$

since $A \cap (S \setminus A) = \phi \quad \forall A \in \mathcal{P}(S)$, therefore every element of $\mathcal{P}(S)$ other than ϕ and S is a zero divisor.

$\mathcal{P}(S)$ is a field if every element other than ϕ is a unit. Hence S should not have any proper subsets. This is possible when $S = \{a\}$.

$\therefore \mathcal{P}(S)$ is a field when S is a singleton.

9.13 Supplementary Exercises

1. Indicate whether the following statements are true or false, with proper justification. Also correct the false statements.

- (i) Every field is an integral domain.
- (ii) Every unit in a ring is not a zero divisor.
- (iii) The product of two units is a unit.
- (iv) The sum of two units is a unit.
- (v) Every non-zero element in \mathbb{R}^2 is a unit.
- (vi) \mathbb{R}^2 is an integral domain.
- (vii) In $e[0, 1]$, the ring of all real valued continuous functions under point-wise addition and multiplication, $f(x)=x$, the identity function is the unit element of the ring.
- (viii) $\left\{ \begin{pmatrix} x & x \\ x & x \end{pmatrix}, |x \in \mathbb{Q} \right\}$ under usual addition and multiplication is a field.

- (ix) Every ring with unity has at least two units.
- (x) A subring of a field is a subfield.
- (xi) A subring of an integral domain is an integral domain.
- (xii) Every element of a ring has a multiplicative inverse.
- (xiii) In a ring R , for $a \in R$, $2x = a$ always has a solution in R .
- (xiv) In a ring R , if $x^2 = a^2$ then $x = \pm a$.
- (xv) $(a + b)^2 = a^2 + b^2 + 2ab$, holds in a ring R .
- (xvi) The sum of two nilpotent elements is nilpotent.
- (xvii) A non-zero nilpotent can be an idempotent.
- (xviii) Subring of a ring with unity is a ring with unity.
- (xix) Subring of a non-commutative ring is always non-commutative.
- (xx) nZ has zero divisors if n is not prime
- (xxi) $x^2 + 2x + 4 = 0$ has solution in \mathbb{Z}_6 .
- (xxii) No subring of \mathbb{Q} is a field.
- (xxiii) The sum of two subrings is a subring.
- (xxiv) The union of two subring is a subring.
- (xxv) The intersection of two subring is a subring.
- (xxvi) $24\mathbb{Z} \subseteq 48\mathbb{Z}$.
- (xxvii) Subring of a non-commutative ring is always non-commutative.
- (xxviii) Every subring of a ring with zero divisors is a ring with zero divisors.
- (xxix) A subring of a ring without unity, cannot have unity.
- (xxx) If R is a ring with unity and S is a subring of R with unity, then the unities of R and S are the same.
- (xxx1) If R is a ring with unity 1_R and S is a subring of R with unity 1_S then $1_R * 1_S = 1_R$
- (xxx2) If A is a non-zero subring of R , then A^2 is always non-zero.
- (xxx3) The number of solutions of $x^2 + 6x + 9 = 0$ in \mathbb{Z}_{12} is 2.
- (xxx4) The number of solution of $x^2 + 12x + 9 = 0$ in \mathbb{Z}_8 is 2.
- (xxx5) Every Boolean ring is commutative.
- (xxx6) If 0 and 1 are the additive and multiplicative identities of a non-zero ring, then $0 \neq 1$.
- (xxx7) Every ring with unity has at least two units.
- (xxx8) Every ring with unity has at most two units.
- (xxx9) Every non-zero divisor in a ring with unity is a units.
- (xxx0) In a ring with unity, the set of units and the set of zero divisor are disjoint.
- (xxx1) An ideal of a subring need not to be an ideal of the ring.
- (xxx2) In a ring with unity every idempotent is a unit.
- (xxx3) There exists ring in which every non-zero element is a unit.

2. State which of the following sets under the indicated operations of addition and multiplication are defined and give a ring structure.
- $(\mathbb{N}, +, \cdot)$.
 - $(\mathbb{R}, +, \cdot)$.
 - $\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Q} \right\}$ under the usual addition and multiplication of matrices.
 - $\left\{ \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$ under the usual multiplication.
 - $\wp(S)$ the power set of S , a non empty set, w.r.t $A + B = A \cap B$.
 - $\{a+ib : a, b \in \mathbb{Z}\}$ under the usual addition and multiplication of complex number, where real part and imaginary part are added reduced and reduced modulo n .
3. Let n be a natural number which is not a perfect square prove that
- $\mathbb{Z}[\sqrt{n}] = \{a + b\sqrt{n} \mid a, b \in \mathbb{Z}\}$ is a commutative ring with unity which is not a field.
 - $\mathbb{Q}[\sqrt{n}] = \{a + b\sqrt{n} \mid a, b \in \mathbb{Q}\}$ is a field.
4. Give an example of an infinite Boolean ring.
5. Let $(G, *)$ be an Abelian group with identity element e .
On G define \cdot by
 $a \cdot b = e \forall a, b \in G$.
Prove that $(G, *, \cdot)$ is a ring. Is it commutative? Does it have unity?
6. Prove that a ring having a unique right unity has a unity.
7. Let R be a ring and n is an even positive integer such that $a^n = a$, for all $a \in R$. Prove that $a = -a$ for all $a \in R$.
8. If R is a ring with unity e . Prove that $\{S = ne \mid n \in \mathbb{Z}\}$ is a subring of R .
9. If $m, n \in \mathbb{Z}^+$ and l is the least common multiple of m and n , Prove that $m\mathbb{Z} \cap n\mathbb{Z} = l\mathbb{Z}$.
10. In $(\mathbb{Z}_n, \oplus_n, \odot_n)$ prove that every non-zero element is either a unit or a zero divisor.
11. If R is a ring then can we say that any non-zero element is either a unit or a zero divisor? Justify your answer.
12. List the units and zero divisors of $\mathbb{Z} \oplus \mathbb{Q}$.
13. Find two element a and b of a ring R , such that a and b are zero divisors and $a + b \neq 0$ is a not-zero divisor.
14. If m and n are relatively prime integers greater than 1, show that \mathbb{Z}_{mn} contains at least two idempotent elements other than zero and unity. Hence find the idempotents of \mathbb{Z}_{30} and \mathbb{Z}_{12} .

15. For what positive integer n does the ring \mathbb{Z}_n have no idempotent element other than zero and unity?
16. Give an example of a ring having element x and y such that $xy=0$ but $yx \neq 0$.
17. Find the units of the $M_2(\mathbb{Z})$.
18. Prove that
 (i) $U(\mathbb{Z}_{30}) = U(30)$
 (ii) $U(\mathbb{Z}_n) = U(n)$.
19. In a ring R , if $a \in R$ is a unit and $b \in R$ is such that $b^n = 0$, then prove that $a + b$ is a unit.
20. Find the units of $\mathbb{Z}[x]$.
21. Consider the algebraic structure $(R, +, \cdot)$ such that
 (i) $(R, +)$ is a group.
 (ii) (R^*, \cdot) is a group.
 (iii) \cdot is right as well as left distributive over $+$.
 Prove that R is a division ring.
22. Prove that $\{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} | a, b, c, d \in \mathbb{Q}\}$ is a field.
23. Prove that $\{a + b\alpha + c\alpha^2 | a, b, c \in \mathbb{Q}\}$, where $\alpha = \sqrt[3]{2}$ is a field.
24. Give an example of a Boolean ring with (i) 8 element (ii) 16 element.
25. Prove that the only Boolean ring which is an integral domain is \mathbb{Z}_2 .
26. Let R be the ring of integral quaternions.
 If $r = a + bi + cj + dk \in R$, then $\bar{r} = a - bi - cj - dk$, is called the conjugate of r . Define
 $N : R \rightarrow \mathbb{Z}$
 by $N(a + bi + cj + dk) = a^2 + b^2 + c^2 + d^2$
 Prove that
 (i) $N(r) = r\bar{r}, \forall r \in R$.
 (ii) $N(rs) = N(r)N(s)$.
 (iii) r is a unit if and only if $N(r) = 1$.
 (iv) the only units of R are $\pm 1, \pm i, \pm j, \pm k$.
27. Find the centre of
 (i) ring of integral quaternions.
 (ii) ring of real quaternions.
28. For a fixed $a \in R$, define
 $C(a) = \{r \in R | ra = ar\}$. Prove that
 (i) $C(a)$ is a subring of R containing a .
 (ii) Centre of $R = \bigcap C(a)$.
 (iii) If R is a division ring, then $C(a)$ is a division ring.
29. Find the units of $\mathbb{Q} \oplus \mathbb{Q}$.
30. Let R_1, R_2 be rings containing non-zero elements. Prove that $R_1 \oplus R_2$ has unity if and only if R_1 unity and R_2 both have unity.

9.14 Answers to Exercise

Exercise - 9.6

10. *Hint* let e be a left unity.
then $(x+e-xe)$ is a left unity, $\forall x \in R$.
13. $I, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$
22. *Hint:* Make multiplication table for \odot_n
(i) No (ii) Yes, 4 (iii) Yes, 6 (iv) No.
23.
(i) $M_2(\mathbb{Z}_2)$
(ii) $M_2(\mathbb{Z}_3)$
(iii) $M_m(\mathbb{Z}_n)$
25. *Hint:* $(m, n) = 1 \Rightarrow \exists \lambda, \mu \in \mathbb{Z}$ such that
 $\lambda m + \mu n = 1,$
then $\lambda m \pmod{mn}$ and $\mu n \pmod{mn}$ are idempotents in \mathbb{Z}_{mn} .
(i) 3,4 (ii) 4,9 (iii) 5,16.
28. *Hint:* In $M_2(\mathbb{Z}_2), A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$
 $A^2 = B^2 = 0$ but $A+B$ is not nilpotent.
Take $C = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{Z}_2),$ then $AC = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ which is not nilpotent.
29. (i) 0, 2 (ii) 0, 2, 4, 6 (iii) 0 (iv) 0, 6 (v) 0, 2, 4, 6, 8, 10, 12, 14 (vi) 0, 10 (vii) 6, 12, 18, 24 30.
30. *Hint* if e is an idempotent than $(ex - exe)^2 = 0 \forall x \in R$
 $\therefore ex - exe = 0.$ similarly $xe - exe = 0.$
31. (i) $\pm 1.$
(ii) 1,2,3,4
(iii) 1,5
(iv) 1,2,4,7,8,11,13,14
(v) (1,1), (-1,1), (1,-1), (-1,-1)
(vi) $\{(a, b) \mid a, b \in \mathbb{Q}^*\}$
(vii) all non-singular matrices
(viii) $([\sqrt{2}])^*.$
33. *Hint:* $a + ib \in \mathbb{Z}[i]$ is a unit if
 $|a + bi| = 1$
units are $\pm 1, \pm i.$
34. *Hint:* If $c \in R$ is such that $ac = ca = 1$
then $(a + b)(c - c^2b) = 1.$ Take $R = M_2(\mathbb{Z}), A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$

35. *Hint:* Proceed as done for $U(\mathbb{Z}[x])$.

36. *Hint:* $(ba)^n = b(ab)^{n-1}a$.

Exercise - 9.8

3. S_4 is not a subring. $S_2 \subset S_1, S_5 \subset S_4$.

5. *Hint:* If $S_3 \not\subset S_1, S_3 \not\subset S_2$ let

$x \in S_3$ such that $x \notin S_1, y \in S_3, y \notin S_2, x \in S, x \notin S_1 \implies x \in S_2$.
Similarly $y \in S_1, x, y \in S_3 \implies x - y \in S_3$.

10. (i) $\{f(\frac{1}{3}) | f(x) \in \mathbb{Z}[x], f(0) = 0\}$
(ii) $\{f(\frac{2}{3}) | f(x) \in \mathbb{Z}[x], f(0) = 0\}$

Exercise - 9.11

11. (i) $\{f : f(x) \neq 0 \text{ for any } x \in [a, b]\}$
(ii) $\{f : f(x) = 0 \text{ for some } x \in [a, b]\}$

Supplementary Exercises

1. (i) True
(ii) True
(iii) True
(iv) False
(v) False
(vi) False
(vii) False
(viii) True, unity is $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, inverse of $\begin{pmatrix} x & x \\ x & x \end{pmatrix}$ is $\begin{pmatrix} \frac{1}{4x} & \frac{1}{4x} \\ \frac{1}{4x} & \frac{1}{4x} \end{pmatrix}$.
(ix) False, \mathbb{Z}_2 .
(x) False, \mathbb{Z} is a subring of the field \mathbb{Q} , but not a subfield.
(xi) True
(xii) False
(xiii) False, $2x = 5$ in \mathbb{Z} has no solution.
(xiv) False
(xv) False
(xvi) False
(xvii) False
(xviii) False
(xix) False
(xx) False
(xxi) $x=2$
(xxii) True
(xxiii) False, $A = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}, |a, b \in \mathbb{Z} \right\}$, $B = \left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}, |a, b \in \mathbb{Z} \right\}$
are subring of $\mathbb{M}_2(\mathbb{Z})$, but $A+B$ is not a subring.

- (xxiv) False
 (xxv) True
 (xxvi) False
 (xxvii) False, $\mathbb{M}_2(\mathbb{Z})$ is non-comm. But $S = \left\{ \begin{pmatrix} n & 0 \\ 2n & 0 \end{pmatrix}, |n \in \mathbb{Z} \right\}$ is commutative subring.
 (xxviii) False
 (xxix) False, $R = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, |a, b \in \mathbb{Z} \right\}$ is a subring of R without unity $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
 (xxx) False, $\mathbb{M}_2(\mathbb{Q})$ is a ring without unity. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} S = \left\{ \begin{pmatrix} x & x \\ x & x \end{pmatrix}, |x \in \mathbb{Z} \right\}$ is a subring of $\mathbb{M}_2(\mathbb{Q})$ with unity $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.
 (xxxii) False, $1_R * 1_S = 1_S$
 (xxxiii) True
 (xxxiv) False it has 3 solutions. Given equation can be written as $x^2 - 6x + 9 = 0 \therefore (x - 3)^2 = 0$.
 (xxxv) True
 (xxxvi) True
 (xxxvii) False, \mathbb{Z}_2
 (xxxviii) False, $\mathbb{Z}_8, \mathbb{Z}_5$
 (xxxix) False
 (xxxx) True
 (xxxxi) True
 (xxxxii) False, True only for $e = 1$.
 (xxxxiii) True

4. $(\wp(S), \Delta, \cap)$ where S is an infinite set.

7. *Hint:* use $(-a)^n = a^n$ as n is even.

12. Units: $(\pm 1, q)$ where $0 \neq q \in \mathbb{Q}$
 Zero divisor : $(0, q)(n, 0) q \neq 0, n \neq 0$

13. *Hint:* In $\mathbb{Z}_6, a = 2, b = 3$.

14. 6, 25 in \mathbb{Z}_3 and 4, 9 in \mathbb{Z}_{12} .

15. n is prime.

19. *Hint:* If $c \in R$ is such that $ac = ca = 1$, $(a + b)(c - c^2b + c^3b^2 + \dots + (-1)^{n-1}c^n b^{n-1}) = 1$

20. *Hint:* $\mathbb{Z}, \mathbb{Z}[x]$ have the same units.

21. *Hint:* Expand $(a+b)(1+1)$ in two way and equate.

24. (i) subring of $M_3(\mathbb{Z}_2)$ consisting of diagonal matrices.
(ii) subring of $M_4(\mathbb{Z}_2)$ consisting of diagonal matrices.
25. *Hint:* Let $a \in R$ then $a^2 = a \implies a(a - 1) = 0 \implies a = 1, 0 \implies R = \{0, 1\}$.
27. (i) \mathbb{Z} , (ii) \mathbb{R} .
29. $\mathbb{Q}^* \times \mathbb{Q}^*$

This page is intentionally left blank.

UNIT - 4

Chapter 10

System of Linear Equations

An equation of the type $ax = b$, where a and b are real constants and x is an unknown, is called a linear equation in x . If x_1, x_2, \dots, x_n are unknowns (called variables) then a relation of the type

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b \quad (10.1)$$

where a_1, \dots, a_n and b are given real constants is called a linear equation in x_1, \dots, x_n . The constants a_1, \dots, a_n are called the coefficients of the variables x_1, \dots, x_n respectively. By a solution of equation (10.1) is meant a set of values of x_1, \dots, x_n which satisfy it.

Example 10.1.

1. $3x - 0.2y = 5$ is a linear equation in two variables x and y . $x = 1, y = -10$ is a solution. More solutions also exist.
2. $\sqrt{3}x - 4y + 3z = -12$ is a linear equation in three variables x, y, z . $x = 0, y = 3, z = 0$ is one solution. Another solution is $x = -4\sqrt{3}, y = z = 0$.
3. $2x^2 + y = 5$ is not a linear equation, because the term x^2 is present.
4. $xy + 4x = 2$ is also not a linear equation because the term xy is present.

A system of m linear equations, in n unknowns x_1, x_2, \dots, x_n (or simply a linear system), is a set of m linear equations each in the same n unknowns. It can be conveniently denoted by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots &\quad \quad \quad \vdots \quad \quad \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \quad (10.2)$$

the two subscripts i and j in a_{ij} are used as follows. The first subscript i indicates that we are dealing with the i th equation, while the second subscript j is

associated with the j th variable x_j . Thus the i th equation is

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$$

in (10.2), for $i = 1, \dots, m$; $j = 1, \dots, n$; a_{ij}, b_i are known as constants. Generally they are real numbers.

A solution to a linear system (10.2) is a sequence of n numbers s_1, s_2, \dots, s_n which has the property that each equation in (10.2) is satisfied when $x_1 = s_1, x_2 = s_2, \dots, x_n = s_n$ are substituted in (10.2). The solution is written as an ordered n -tuple (s_1, s_2, \dots, s_n) . The set of all solutions is called the solution set of the linear system. A linear system is said to be consistent if it has a solution, otherwise it is said to be inconsistent. Two linear systems, having the same solution set are called equivalent systems, i.e., each solution of a first system is a solution of the second system and vice versa.

If $b_1 = b_2 = \dots = b_m = 0$, then the system is said to be homogenous and non-homogenous otherwise.

Example 10.2.

1. $3x - 2y = 1, 5x + 6y = 11$ is a linear system of equations in the variables x and y . $x = 1, y = 1$ is a solution of both the equations and is therefore a solution of the system of equation. The given system of equation is consistent. Hence the solution set is $\{(1, 1)\}$.
2. $3x + 4y = 5, 6x + 8y = 81$ has no solution. Hence it is an inconsistent system. The solution set is the empty set.
3. $x - y + 3z = 6, x + 3y - 3z = -4, 5x + 3y + 3z = 10$ is a system of equations in the three variables x, y, z . $(2, -1, 1)$ is a solution of this linear system, so that system of equation is consistent. $(5, -4, -1)$ is also a solution. Thus the given system has more than one solutions. In fact, if k is any real number $\{(\frac{7}{2} - \frac{3}{2}k, \frac{-5}{2} + \frac{3}{2}k, k) : k \text{ is a real number}\}$ is always a solution. Thus the solution set is $\{(\frac{7}{2} - \frac{3}{2}k, \frac{-5}{2} + \frac{3}{2}k, k) : k \text{ is a real number}\}$.

The set of all solutions of

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

with at least one a_i non zero, is called a hyperplane in \mathbb{R}^n .

When $n = 2$, it is a line in \mathbb{R}^2 and when $n = 3$, it is a plane in \mathbb{R}^3 .

Geometrical Interpretation

The graph of a linear equation $ax + by = c$ in two variables x and y is a line. For every solution $x = x_1, y = y_1$ of this equation, the point (x_1, y_1) lies on the line and vice versa. Thus a single linear equation in two variables has infinitely many solutions.

Let us now consider two linear equations,

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

A solution of this system is the set of values x and y which satisfy both the

equations, i.e., it lies on both the lines represented by these equations. Thus, it is the point of intersection of these two lines, if it exists.

There are three possibilities

- (i) The two lines meet at a point, so that there is exactly one solution.
- (ii) The two lines are parallel, i.e., they never meet. Hence there is no solution.
- (iii) The two lines become identical. This amounts to saying that they meet at every point. Hence there are infinitely many solutions.

Thus a solution set of a system is intersection of the graphs of each equation of this system.

10.1 Matrix Notation

Consider the linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

of m equations in the n unknowns x_1, x_2, \dots, x_n . The above system can be written as $AX = b$, where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

A is called the coefficient matrix of the system.

If $b = 0$, then the system is said to be homogeneous. If the right hand side of the system is attached to A as the $(n+1)$ th column, then the matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & \vdots & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & \vdots & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & \vdots & b_m \end{pmatrix}$$

is called the augmented matrix of the system and is denoted by $\left(A \quad \vdots \quad b \right)$.

Within an augmented matrix, the horizontal and vertical subarrays

$$\left(a_{i1} \quad a_{i2} \quad \dots \quad a_{in} \quad \vdots \quad b_i \right) \text{ and } \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

are the i th-row (which represents the i th equation) and the j th column (which are the coefficients of the j th variable x_j) of the augmented matrix respectively. Clearly, there is a one-to-one

correspondence between the columns of the coefficient matrix and the variables of the system. The last column $(b_1 \ b_2 \ \dots \ b_m)^t$ of the augmented matrix represents homogeneity of the system and so no variables corresponds to it.

Example 10.3. Consider a system of equations

$$\begin{aligned}x_1 + 2x_2 + 3x_3 - x_4 &= 5 \\2x_1 - x_2 + x_3 + x_4 &= -2 \\x_1 + 3x_2 - 4x_3 + 5x_4 &= 6\end{aligned}$$

These are 3 equations in 4 unknowns x_1, x_2, x_3, x_4 .

$$\text{If } A = \begin{pmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & 1 & 1 \\ 1 & 3 & -4 & 5 \end{pmatrix}$$

then A is the coefficient matrix (or matrix of coefficients) of the given system. The matrix A with an extra column, which is the right hand side, is the augmented matrix of the given system. Thus

$$\left(A \ : \ b \right) = \begin{pmatrix} 1 & 2 & 3 & -1 & \vdots & 5 \\ 2 & -1 & 1 & 1 & \vdots & -2 \\ 1 & 3 & -4 & 5 & \vdots & 6 \end{pmatrix}$$

10.2 Solving a Linear System

In this chapter, we make a systematic study of the theoretical aspects of the solution of the linear equations and give some computational procedures. The basic approach adopted in solving a given system is to find an equivalent system which is easier to solve. This is illustrated in the following example (you have already done this in middle school). The corresponding augmented matrix of the system will be given along side.

Example 10.4. Consider the system of equations

$$x_1 - x_2 + x_3 = 6, \quad 2x_1 + 3x_2 + 4x_3 = 8, \quad 5x_1 - 2x_2 + 6x_3 = 27$$

This system will be solved by the process of elimination. The corresponding augmented matrix will be written alongside.

The system is :

$$\begin{aligned}x_1 - x_2 + x_3 &= 6 \quad \dots (1) \\2x_1 + 3x_2 + 4x_3 &= 8 \quad \dots (2) \\5x_1 - 2x_2 + 6x_3 &= 27 \quad \dots (3)\end{aligned} \quad \left(\begin{array}{cccc} 1 & -1 & 1 & 6 \\ 2 & 3 & 4 & 8 \\ 5 & -2 & 6 & 27 \end{array} \right) = A$$

Eliminate x_1 from equations (2) and (3), using equation (1). For this we apply

$$(2) - 2 \times (1) : \quad 5x_2 + 2x_3 = -4$$

$$(3) - 5 \times (1) : \quad 3x_2 + x_3 = -3$$

Thus we get the equivalent system of equations

$$\begin{aligned}x_1 - x_2 + x_3 &= 6 \quad \dots (4) \\5x_2 + 2x_3 &= -4 \quad \dots (5) \\3x_2 + x_3 &= -3 \quad \dots (6)\end{aligned} \quad \left(\begin{array}{cccc} 1 & -1 & 1 & 6 \\ 0 & 5 & 2 & -4 \\ 0 & 3 & 1 & -3 \end{array} \right) = B$$

Eliminate x_2 from equation (6) using equation (5). For this we apply

$$(6) - \frac{3}{5} \times (5): \quad \frac{-1}{5}x_3 = \frac{-3}{5}$$

Thus we get the equivalent system of equations

$$\begin{array}{rcl} x_1 - x_2 + x_3 & = & 6 \quad \dots (7) \\ 5x_2 + 2x_3 & = & -4 \quad \dots (8) \\ \frac{-1}{5}x_3 & = & \frac{-3}{5} \quad \dots (9) \end{array} \quad \left(\begin{array}{cccc} 1 & -1 & 1 & 6 \\ 0 & 5 & 2 & -4 \\ 0 & 0 & \frac{-1}{5} & \frac{-3}{5} \end{array} \right) = C$$

Solving equation (9) for x_3 , equation (8) for x_2 and equation (7) for x_1 , we get, $x_3 = 3, x_2 = -2, x_1 = 1$

Note that the augmented matrix of this system is a triangular matrix.

Observe that on A , if we apply the operations

Row 2 \rightarrow Row 2 + (-2) Row 1

Row 3 \rightarrow Row 3 + (-5) Row 1,

we get the matrix B . On B if we apply,

Row 3 \rightarrow Row 3 + $\frac{-3}{5}$ Row 2, we get the matrix C .

The system of equations corresponding to matrix C are the equations (7), (8) and (9).

This example illustrates that operations on equations in a linear system correspond to operations on the corresponding rows of the augmented matrix. Such operations are called elementary row operations. There are 3 types of elementary row operations.

10.3 Elementary Row Operations (ERO)

E1 (*Interchange*) Interchange of two rows. It is denoted by $R_i \leftrightarrow R_j$ when the i th and j th rows are interchanged.

E2 (*Scaling*) Multiply each element of a row by a non-zero constant. It is denoted by $R_i \rightarrow cR_i$ when the elements of the i th row are multiplied by $c \neq 0$.

E3 (*Replacement*) Replace the elements of the row by the sum of itself and a multiple of another row. It is denoted by $R_i \rightarrow R_i + k R_j$ when the i th row is replaced by the sum of the i th row and k times the j th row.

Note that in operation E_1 two rows of a matrix are affected whereas in operations E_2 and E_3 only one row is affected. Row operations can be applied to any matrix, not merely to those which arise as the augmented matrix of the linear system.

Two matrices A and B are said to be *row equivalent* if one can be obtained from the other by a sequence of elementary row operations. Symbolically we write $A \sim B$.

Elementary row operations are reversible. If B is the matrix obtained from A by $R_i \leftrightarrow R_j$ then A can be obtained from B by $R_i \rightarrow R_j$.

If D is the matrix obtained from C by $R_i \rightarrow cR_i, c \neq 0$, then C can be obtained from D by $R_i \leftrightarrow c^{-1}R_i$.

If F is the matrix obtained from E by $R_i \rightarrow R_i + k R_j$, then E can be obtained from F by $R_i \rightarrow R_i + (-k) R_j$.

At the moment we are interested in row operations on the augmented matrix

of a system of linear equations. If a linear system is changed into a new one by applying elementary row operations. Then by considering each type of elementary row operations it can be seen that any solution of the original system remains a solution of the new system. Conversely, since the original system can be produced via row operations on the new system, each solution of the new system is also a solution of the original system. Thus, "If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set."

Example 10.5. Let $A = \begin{pmatrix} 2 & 5 & 6 \\ 1 & 3 & -4 \\ 7 & 2 & 8 \end{pmatrix}$

Apply $R_1 \leftrightarrow R_2$ on A .

Then $A \sim \begin{pmatrix} 1 & 3 & -4 \\ 2 & 5 & 6 \\ 7 & 2 & 8 \end{pmatrix} = B(\text{say})$

Apply $R_3 \rightarrow (-1)R_3$ on B

Then $B \sim \begin{pmatrix} 1 & 3 & -4 \\ 2 & 5 & 6 \\ -7 & -2 & -8 \end{pmatrix} = C(\text{say}).$

Then $B \sim C$. Thus $A \sim C$

Apply $R_3 \rightarrow R_3 + 2R_2$ on C

Then $C \sim \begin{pmatrix} 1 & 3 & -4 \\ 2 & 5 & 6 \\ -3 & 8 & 4 \end{pmatrix} = D(\text{say})$

We shall now explain the steps involved in solving the given system of equations using augmented matrix. Consider a system of four linear equations in four unknowns in order to explain the procedure explicitly.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 &= b_4 \end{aligned}$$

Form the augmented matrix

$$\left(A \ : \ b \right) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \vdots & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & \vdots & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & \vdots & b_3 \\ a_{41} & a_{42} & a_{43} & a_{44} & \vdots & b_4 \end{pmatrix}$$

Step 1 Make the elements in the first column below a_{11} , zero.

- (a) If $a_{11} \neq 0$ proceed to (b). If $a_{11} = 0$ and $a_{i1} \neq 0$ for some $i = 2, 3, 4$, then apply $R_1 \leftrightarrow R_i$. If $a_{i1} = 0$, $i = 1, 2, 3, 4$, then go to step 2.
- (b) By applying $R_i \rightarrow R_i + \frac{-a_{i1}}{a_{11}} R_1$, $i = 2, 3, 4$, we get the matrix

$$\left(\begin{array}{c} A \\ \vdots \\ b \end{array} \right) \sim \left(\begin{array}{cccccc} b_{11} & b_{12} & b_{13} & b_{14} & \vdots & c_1 \\ 0 & b_{22} & b_{23} & b_{24} & \vdots & c_2 \\ 0 & b_{32} & b_{33} & b_{34} & \vdots & c_3 \\ 0 & b_{42} & b_{43} & b_{44} & \vdots & c_4 \end{array} \right) = \left(\begin{array}{c} B \\ \vdots \\ c \end{array} \right)$$

Step 2 Make the elements in second columns below b_{22} , zero.

Consider the matrix $\left(\begin{array}{c} B \\ \vdots \\ c \end{array} \right)$. Using the element b_{22} (in a similar way as a_{11}), make the elements below b_{22} zero, as in step 1. Then

$$\left(\begin{array}{c} B \\ \vdots \\ c \end{array} \right) \sim \left(\begin{array}{cccccc} b_{11} & b_{12} & b_{13} & b_{14} & \vdots & c_1 \\ 0 & c_{22} & c_{23} & c_{24} & \vdots & d_2 \\ 0 & c_{32} & c_{33} & c_{34} & \vdots & d_3 \\ 0 & c_{42} & c_{43} & c_{44} & \vdots & d_4 \end{array} \right) = \left(\begin{array}{c} C \\ \vdots \\ d \end{array} \right)$$

Step 3 Continue this process till $\left(\begin{array}{c} A \\ \vdots \\ b \end{array} \right)$ is equivalent to a triangular matrix. Thus

$$\left(\begin{array}{c} A \\ \vdots \\ b \end{array} \right) \sim \left(\begin{array}{cccccc} d_{11} & d_{12} & d_{13} & d_{14} & \vdots & p_1 \\ 0 & d_{22} & d_{23} & d_{24} & \vdots & p_2 \\ 0 & 0 & d_{33} & d_{34} & \vdots & p_3 \\ 0 & 0 & 0 & d_{44} & \vdots & p_4 \end{array} \right)$$

Hence the given system of equations is equivalent to

$$\begin{aligned} d_{11}x_1 + d_{12}x_2 + d_{13}x_3 + d_{14}x_4 &= p_1 \\ d_{22}x_2 + d_{23}x_3 + d_{24}x_4 &= p_2 \\ d_{33}x_3 + d_{34}x_4 &= p_3 \\ d_{44}x_4 &= p_4 \end{aligned}$$

Solve these equations for x_4, x_3, x_2, x_1 respectively.

10.4 Solved Problems

Problem 10.1. Solve the linear system with the augmented matrix

$$(i) \left(\begin{array}{cccc} 1 & -2 & 3 & \vdots & 4 \\ 2 & -1 & -3 & \vdots & 5 \\ 3 & 0 & 1 & \vdots & 2 \\ 3 & -3 & 0 & \vdots & 7 \end{array} \right)$$

$$(ii) \begin{pmatrix} 1 & 2 & 3 & 1 & \vdots & 8 \\ 1 & 3 & 0 & 1 & \vdots & 7 \\ 1 & 0 & 2 & 1 & \vdots & 3 \end{pmatrix}$$

Solution:

$$(i) \left(A \ : \ b \right) = \begin{pmatrix} 1 & -2 & 3 & \vdots & 4 \\ 2 & -1 & -3 & \vdots & 5 \\ 3 & 0 & 1 & \vdots & 2 \\ 3 & -3 & 0 & \vdots & 7 \end{pmatrix}$$

Applying $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - 3R_1, R_4 \rightarrow R_4 - 3R_1$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & -2 & 3 & \vdots & 4 \\ 0 & 3 & -9 & \vdots & -3 \\ 0 & 6 & -8 & \vdots & -10 \\ 0 & 3 & -9 & \vdots & -5 \end{pmatrix}$$

Applying $R_3 \rightarrow R_3 - 2R_2, R_4 \rightarrow R_4 - R_2$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & -2 & 3 & \vdots & 4 \\ 0 & 3 & -9 & \vdots & -3 \\ 0 & 0 & 10 & \vdots & -4 \\ 0 & 0 & 0 & \vdots & -2 \end{pmatrix}$$

Thus the system of equations is

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 4 \\ 3x_2 - 9x_3 &= -3 \\ 10x_3 &= -4 \\ 0x_3 &= -2 \end{aligned}$$

the last equation gives $0 = -2$ which is not possible. Hence the system is inconsistent.

$$(ii) \left(A \ : \ b \right) = \begin{pmatrix} 1 & 2 & 3 & 1 & \vdots & 8 \\ 1 & 3 & 0 & 1 & \vdots & 7 \\ 1 & 0 & 2 & 1 & \vdots & 3 \end{pmatrix}$$

Applying $R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - R_1$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & 2 & 3 & 1 & \vdots & 8 \\ 0 & 1 & -3 & 0 & \vdots & -1 \\ 0 & -2 & -1 & 0 & \vdots & -5 \end{pmatrix}$$

Applying $R_3 \rightarrow R_3 + 2R_2$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & 2 & 3 & 1 & \vdots & 8 \\ 0 & 1 & -3 & 0 & \vdots & -1 \\ 0 & 0 & -7 & 0 & \vdots & -7 \end{pmatrix}$$

Thus the system of equations is,

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + x_4 &= 8 \\ x_2 - 3x_3 &= -1 \\ -7x_3 &= -7 \end{aligned}$$

$$\therefore x_3 = 1, x_2 = 2$$

$$x_1 + 2x_2 + 3x_3 + x_4 = 8$$

$$\text{or } x_1 + x_4 = 1$$

$$\text{If } x_4 = k \text{ then } x_1 = 1 - k$$

Hence the solution is

$x_1 = 1 - k, x_2 = 2, x_3 = 1, x_4 = k$, where k is any real number. Thus the given system is consistent and has infinitely many solutions.

Problem 10.2. Let the following matrix be the augmented matrix of a system of equations.

$$\begin{pmatrix} 1 & -4 & 7 & \vdots & g \\ 0 & 3 & -5 & \vdots & h \\ -2 & 5 & -9 & \vdots & k \end{pmatrix}$$

Determine a relation between g, h and k so that the system is

(i) consistent,

(ii) inconsistent.

Solution: Let $A = \begin{pmatrix} 1 & -4 & 7 & \vdots & g \\ 0 & 3 & -5 & \vdots & h \\ -2 & 5 & -9 & \vdots & k \end{pmatrix}$

Step 1 Apply $R_3 \rightarrow R_3 + 2R_1$ to A . Then

$$A \sim \begin{pmatrix} 1 & -4 & 7 & \vdots & g \\ 0 & 3 & -5 & \vdots & h \\ 0 & -3 & 5 & \vdots & k + 2g \end{pmatrix}$$

Step 2 $R_3 \rightarrow R_3 + R_2$

$$A \sim \begin{pmatrix} 1 & -4 & 7 & \vdots & g \\ 0 & 3 & -5 & \vdots & h \\ 0 & 0 & 0 & \vdots & h+k+2g \end{pmatrix}$$

The last equation will be
 $0x_3 = h + k + 2g$

- (i) If the system is consistent then we must have $h + k + 2g = 0$
 (ii) If the system is inconsistent then $h + k + 2g \neq 0$

Problem 10.3. Determine if the following system of equations is consistent

$$\begin{aligned} x_1 - 3x_2 + 4x_3 &= -4 \\ 3x_1 - 7x_2 + 7x_3 &= -8 \\ -4x_1 + 6x_2 - 2x_3 &= 7 \end{aligned}$$

If it is, find the solution.

Solution: The augmented matrix of the system is:

$$\left(A \ : \ b \right) = \begin{pmatrix} 1 & -3 & 4 & \vdots & -4 \\ 3 & -7 & 7 & \vdots & -8 \\ -4 & 6 & -2 & \vdots & 7 \end{pmatrix}$$

Step 1 Applying $R_2 \rightarrow R_2 + (-3)R_1, R_3 \rightarrow R_3 + 4R_1,$

$$\left(A \ : \ b \right) \begin{pmatrix} 1 & -3 & 4 & \vdots & -4 \\ 0 & 2 & -5 & \vdots & 4 \\ 0 & -6 & 14 & \vdots & -9 \end{pmatrix}$$

Step 2 Applying $R_3 \rightarrow R_3 + 3R_2,$

$$\left(A \ : \ b \right) \begin{pmatrix} 1 & -3 & 4 & \vdots & -4 \\ 0 & 2 & -5 & \vdots & 4 \\ 0 & 0 & -1 & \vdots & 3 \end{pmatrix}$$

The augmented matrix is equivalent to a triangular matrix. The given system of equations is equivalent to

$$\begin{aligned} x_1 - 3x_2 + 4x_3 &= -4 \\ 2x_2 - 5x_3 &= 4 \\ -x_3 &= 3 \end{aligned}$$

Solving, we get $x_3 = -3, x_2 = \frac{-11}{2}, x_1 = \frac{-17}{2}.$

Hence the given system has a solution, so that it is consistent. Solution is unique and is $(\frac{-17}{2}, \frac{-11}{2}, -3).$

Problem 10.4. Solve the following system of equations

$$\begin{aligned}x_1 + 3x_2 - 2x_3 &= 3 \\2x_1 + 6x_2 - 2x_3 + 4x_4 &= 18 \\x_2 + x_3 + 3x_4 &= 10\end{aligned}$$

Solution: The augmented matrix of the system is:

$$\left(A \ : \ b \right) = \begin{pmatrix} 1 & 3 & -2 & 0 & \vdots & 3 \\ 2 & 6 & -2 & 4 & \vdots & 18 \\ 0 & 1 & 1 & 3 & \vdots & 10 \end{pmatrix}$$

Applying $R_2 \rightarrow R_2 - 2R_1$

$$\left(A \ : \ b \right) \begin{pmatrix} 1 & 3 & -2 & 0 & \vdots & 3 \\ 0 & 0 & 2 & 4 & \vdots & 12 \\ 0 & 1 & 1 & 3 & \vdots & 10 \end{pmatrix}$$

Applying $R_2 \leftrightarrow R_3$ (In order to make the element in the (2,2)th position non-zero)

$$\left(A \ : \ b \right) \begin{pmatrix} 1 & 3 & -2 & 0 & \vdots & 3 \\ 0 & 1 & 1 & 3 & \vdots & 10 \\ 0 & 0 & 2 & 4 & \vdots & 12 \end{pmatrix}$$

Thus the system of equations is

$$\begin{aligned}x_1 + 3x_2 - 2x_3 &= 3 \\x_2 + x_3 + 3x_4 &= 10 \\2x_3 + 4x_4 &= 12\end{aligned}$$

By back substitution, we get

$$\begin{aligned}x_3 &= 6 - 2x_4, \\x_2 &= 4 - x_4, \\x_1 &= 3 - x_4\end{aligned}$$

Hence the given system is consistent and has infinitely many solutions, given by

$$\begin{aligned}x_1 &= 3 - k \\x_2 &= 4 - k \\x_3 &= 6 - 2k \\x_4 &= k\end{aligned}$$

where k is any real number.

Problem 10.5. Determine whether the following system of equations is consistent

$$\begin{aligned}3x_1 + x_2 - 4x_3 &= 7 \\x_1 - 2x_2 + 3x_3 &= 6 \\5x_1 - 3x_2 + 2x_3 &= 5\end{aligned}$$

If it is, find the solution.

Solution: The augmented matrix of the system is:

$$\left(A \ : \ b \right) = \begin{pmatrix} 3 & 1 & -4 & \vdots & 7 \\ 1 & -2 & 3 & \vdots & 6 \\ 5 & -3 & 2 & \vdots & 5 \end{pmatrix}$$

Step 1 For ease of calculations we interchange the first and second rows.
Applying $R_1 \leftrightarrow R_2$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & -2 & 3 & \vdots & 6 \\ 3 & 1 & -4 & \vdots & 7 \\ 5 & -3 & 2 & \vdots & 5 \end{pmatrix}$$

Applying $R_2 \rightarrow R_2 + (-3)R_1$

$R_3 \rightarrow R_3 + (-5)R_1$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & -2 & 3 & \vdots & 6 \\ 0 & 7 & -13 & \vdots & -11 \\ 0 & 7 & -13 & \vdots & -25 \end{pmatrix}$$

Step 2 Applying $R_3 \rightarrow R_3 - R_2$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & -2 & 3 & \vdots & 6 \\ 0 & 7 & -13 & \vdots & -11 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

The augmented matrix is equivalent to a triangular matrix. The given system of equations is equivalent to

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 6 \\ 7x_2 - 13x_3 &= -25 \\ 0x_3 &= 0 \end{aligned}$$

The last equation is satisfied for any value of x_3 . Solving for x_2 and x_1 in terms of x_3 we get,

$$\begin{aligned} x_2 &= \frac{13}{7}x_3 - \frac{25}{7} \\ x_1 &= \frac{5}{7}x_3 - \frac{8}{7} \end{aligned}$$

Thus the given system is consistent and the solution is $(\frac{5}{7}k - \frac{8}{7}, \frac{13}{7}k - \frac{25}{7}, k)$ where k is any real number. Hence there are infinitely many solutions.

Problem 10.6. Determine if the following system of equation is consistent

$$\begin{aligned} x_2 + 4x_3 &= -5 \\ x_1 + 3x_2 + 5x_3 &= -2 \\ 3x_1 + 7x_2 + 7x_3 &= 6 \end{aligned}$$

If it is, find the solution.

Solution: The augmented matrix of the system is:

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 0 & 1 & 4 & \vdots & -5 \\ 1 & 3 & 5 & \vdots & -2 \\ 3 & 7 & 7 & \vdots & 6 \end{pmatrix}$$

Step 1 Since the element in the (1, 1)th position is zero, we shall make it non-zero by applying $R_2 \leftrightarrow R_1$

$$\begin{aligned} \text{Thus } \left(A \ : \ b \right) &\sim \begin{pmatrix} 1 & 3 & 5 & \vdots & -2 \\ 0 & 1 & 4 & \vdots & -5 \\ 3 & 7 & 7 & \vdots & 6 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 3 & 5 & \vdots & -2 \\ 0 & 1 & 4 & \vdots & -5 \\ 0 & -2 & -8 & \vdots & 12 \end{pmatrix} \text{ Applying } R_3 \longrightarrow R_3 + (-3)R_1 \\ &\sim \begin{pmatrix} 1 & 3 & 5 & \vdots & -2 \\ 0 & 1 & 4 & \vdots & -5 \\ 0 & 0 & 0 & \vdots & 2 \end{pmatrix} \text{ Applying } R_2 \longrightarrow R_2 + 2R_1 \end{aligned}$$

The augmented matrix is equivalent to a triangular matrix. The given system of equations is equivalent to

$$\begin{aligned} x_1 + 3x_2 + 5x_3 &= -2 \\ x_2 + 4x_3 &= -5 \\ 0x_3 &= 2 \end{aligned}$$

The third equation is not satisfied for any value of x_3 . Hence the given system does not have a solution and is inconsistent.

Problem 10.7. Determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system

$$(i) \begin{pmatrix} 1 & h & -3 \\ -2 & 4 & 6 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 2 & -3 & h \\ -6 & 9 & 5 \end{pmatrix}$$

Solution:

$$(i) \text{ Let } A = \begin{pmatrix} 1 & h & -3 \\ -2 & 4 & 6 \end{pmatrix}$$

$$A \sim \begin{pmatrix} 1 & h & -3 \\ 0 & 4+2h & 0 \end{pmatrix} \text{ Applying } R_2 \longrightarrow R_2 + 2R_1$$

The system of equations is

$$\begin{aligned} x_1 + hx_2 &= -3 \\ (2h+4)x_2 &= 0 \end{aligned}$$

The system is consistent for every value of h .

$$(ii) A = \begin{pmatrix} 2 & -3 & h \\ -6 & 9 & 5 \end{pmatrix}$$

$$A \sim \begin{pmatrix} 2 & -3 & h \\ 0 & 0 & 5+3h \end{pmatrix} \text{ Applying } R_2 \longrightarrow R_2 + 3R_1$$

The system of equations is

$$\begin{aligned} 2x_1 - 3x_2 &= h \\ 0x_2 &= 5 + 3h \end{aligned}$$

Thus $5 + 3h = 0$ so that $h = -\frac{5}{3}$.
Hence the given system is consistent when $h = -\frac{5}{3}$.

Problem 10.8. Find an equation relating a , b and c so that the linear system

$$\begin{aligned} 2x + 2y + 3z &= a \\ 3x - y + 5z &= b \\ x - 3y + 2z &= c \end{aligned}$$

is consistent for any values of a , b and c that satisfy that equation.

Solution: The augmented matrix of the system is:

$$\left(A \ : \ b \right) = \begin{pmatrix} 2 & 2 & 3 & \vdots & a \\ 3 & -1 & 5 & \vdots & b \\ 1 & -3 & 2 & \vdots & c \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & -3 & 2 & \vdots & c \\ 3 & -1 & 5 & \vdots & b \\ 2 & 2 & 3 & \vdots & a \end{pmatrix} \quad \text{Applying } R_1 \longleftrightarrow R_3$$

$$\sim \begin{pmatrix} 1 & -3 & 2 & \vdots & c \\ 0 & 8 & -1 & \vdots & b-3c \\ 0 & 8 & -1 & \vdots & a-2c \end{pmatrix} \quad \text{Applying } R_2 \longrightarrow R_2 - 3R_1, R_3 \longrightarrow R_3 - 2R_1$$

$$\sim \begin{pmatrix} 1 & -3 & 2 & \vdots & c \\ 0 & 8 & -1 & \vdots & b-3c \\ 0 & 0 & 0 & \vdots & a-b+c \end{pmatrix} \quad \text{Applying } R_3 \longrightarrow R_3 - R_2$$

Thus the corresponding system of equations is

$$\begin{aligned} x - 3y + 2z &= c \\ 8y - z &= b - 3c \\ 0z &= a - b + c \end{aligned}$$

The 3rd equation is consistent if $a - b + c = 0$.
Hence the given system is consistent if $a - b + c = 0$.

Problem 10.9. Find the values of k for which the resulting linear system has

- (i) No solution
- (ii) unique solution, and
- (iii) infinitely many solutions.

$$\begin{aligned}x + y + z &= 2 \\x + 2y + z &= 3 \\x + y + (k^2 - 5)z &= k\end{aligned}$$

Solution: The augmented matrix of the system is:

$$\begin{aligned}\left(A \ : \ b \right) &= \begin{pmatrix} 1 & 1 & 1 & \vdots & 2 \\ 1 & 2 & 1 & \vdots & 3 \\ 1 & 1 & k^2 - 5 & \vdots & k \end{pmatrix} \\ \sim \begin{pmatrix} 1 & 1 & 1 & \vdots & 2 \\ 0 & 1 & 0 & \vdots & 1 \\ 0 & 0 & k^2 - 6 & \vdots & k - 2 \end{pmatrix} & \text{Applying } R_2 \rightarrow R_2 - R_1, \ R_3 \rightarrow R_3 - R_1\end{aligned}$$

Thus the system of equations is

$$x + y + z = 2 \quad (1)$$

$$y = 1 \quad (2)$$

$$(k^2 - 6)z = k - 2 \quad (3)$$

Two cases arise:

Case 1 $k^2 - 6 \neq 0$. Then,

$$z = \frac{k - 2}{k^2 - 6}.$$

Case 2 $k^2 = 6$.

The system has no solution. Since for these values of k , equation(3) is absurd as its left hand side is zero, while the right hand side is non-zero. Thus the system has

- (i) no solution when $k = \pm\sqrt{6}$
- (ii) unique solution when $k \neq \pm\sqrt{6}$
- (iii) infinitely many solutions, is not possible.

Problem 10.10. Construct three different augmented matrices for linear system whose solution set is $x_1 = -2$, $x_2 = 1$, $x_3 = 0$.

Solution: The system of equations

$$x_1 = -2$$

$$x_2 = 1$$

$$x_3 = 0$$

is one system, having the given solution set. Its augmented matrix is

$$\left(A \ : \ b \right) = \begin{pmatrix} 1 & 0 & 0 & \vdots & -2 \\ 0 & 1 & 0 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 0 \end{pmatrix}$$

Any system of equations whose augmented matrix is equivalent to $\left(A \ : \ b \right)$ will have the same solution set. Thus the required augmented matrices will be obtained by applying ERO to $\left(A \ : \ b \right)$. Applying $R_1 \rightarrow R_1 + 2R_2$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 0 \end{pmatrix} = \left(B \ : \ d \right) \text{ (say)}$$

Applying $R_3 \rightarrow R_3 + R_1 + 3R_2$

$$\left(A \ : \ b \right) \sim \begin{pmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 1 \\ 1 & 5 & 1 & \vdots & 3 \end{pmatrix} = \left(C \ : \ e \right) \text{ (say)}$$

Thus $\left(A \ : \ b \right)$, $\left(B \ : \ d \right)$, $\left(C \ : \ e \right)$ are the required three different augmented matrices.

10.5 Exercise

- Find the matrices obtained by performing operations on A, where

$$A = \begin{pmatrix} 2 & 0 & 4 & 2 \\ 3 & -2 & 5 & 6 \\ -1 & 3 & 1 & 1 \end{pmatrix}$$

- Interchange the 1st and 3rd row.
 - Multiply the 2nd row by -3.
 - Adding -3 times third row to the first row.
 - Adding -1 times first row and the third row to the second row.
- Find ERO that transforms the first matrix into the second and then find the reverse row operation that transforms second into first matrix.

$$(i) \begin{pmatrix} 1 & 2 & -1 & 3 \\ 4 & 1 & 6 & 8 \\ -2 & 0 & 1 & -4 \end{pmatrix}, \begin{pmatrix} -2 & 0 & 1 & -4 \\ 4 & 1 & 6 & 8 \\ 1 & 2 & -1 & 3 \end{pmatrix}$$

$$(ii) \begin{pmatrix} -1 & 1 & 0 & -1 \\ 2 & -2 & 3 & -4 \\ 5 & -7 & 6 & -8 \\ 0 & 1 & -4 & 3 \end{pmatrix}, \begin{pmatrix} -1 & 1 & 0 & -1 \\ 2 & -2 & 3 & -4 \\ -10 & 14 & -12 & 16 \\ 0 & 1 & -4 & 3 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 1 & 3 & 9 \\ 3 & -6 & 7 \\ 0 & -2 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 9 \\ 3 & -6 & 7 \\ 0 & 1 & -3 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 2 & 4 & 6 & -2 \\ 3 & 1 & 4 & 6 \\ -2 & 0 & 1 & -1 \end{pmatrix}, \begin{pmatrix} 2 & 4 & 6 & -2 \\ 3 & 1 & 4 & 6 \\ -1 & 2 & 4 & -2 \end{pmatrix}$$

$$(v) \begin{pmatrix} 1 & 3 & 4 \\ 2 & -1 & 0 \\ 3 & 4 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 4 \\ 0 & -7 & -8 \\ 3 & 4 & -1 \end{pmatrix}$$

3. Find three matrices which are row equivalent to the matrix

$$\begin{pmatrix} 4 & 3 & -1 & 5 \\ -4 & 2 & -11 & 0 \\ 2 & -3 & 0 & -5 \end{pmatrix}$$

4. The augmented matrix of a linear system has been reduced by row operations to the form shown. In each case, continue the appropriate row operation and describe the solutions of the original system.

$$(i) \begin{pmatrix} 1 & 0 & -3 & -2 \\ 0 & -3 & 10 & 7 \end{pmatrix}$$

$$(ii) \begin{pmatrix} -1 & 3 & 1 \\ 0 & 17 & 17 \\ 3 & -4 & 2 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 0 & 0 & 1 & 2 \\ 2 & 3 & 0 & -2 \\ -1 & -3 & 6 & -5 \end{pmatrix}$$

5. Solve the linear system associated with the given augmented matrix

$$(i) \begin{pmatrix} 1 & 1 & 1 & \vdots & 0 \\ 1 & 1 & 0 & \vdots & 3 \\ 0 & 1 & 1 & \vdots & 1 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 1 & 2 & 3 & \vdots & 0 \\ 1 & 1 & 1 & \vdots & 0 \\ 5 & 7 & 9 & \vdots & 0 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 1 & 2 & 3 & \vdots & 0 \\ 1 & 1 & 1 & \vdots & 0 \\ 1 & 1 & 2 & \vdots & 0 \\ 1 & 3 & 3 & \vdots & 0 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 1 & 2 & 1 & \vdots & 7 \\ 2 & 0 & 1 & \vdots & 4 \\ 1 & 0 & 2 & \vdots & 5 \\ 1 & 2 & 3 & \vdots & 11 \\ 2 & 1 & 4 & \vdots & 12 \end{pmatrix}$$

$$(v) \begin{pmatrix} 1 & 1 & 3 & -3 & \vdots & 0 \\ 0 & 2 & 1 & -3 & \vdots & 3 \\ 1 & 0 & 2 & -1 & \vdots & -1 \end{pmatrix}$$

$$(vi) \begin{pmatrix} 4 & 2 & -1 & \vdots & 5 \\ 3 & 3 & 6 & \vdots & 1 \\ 5 & 1 & -8 & \vdots & 8 \end{pmatrix}$$

$$(vii) \begin{pmatrix} 1 & 2 & -3 & \vdots & 4 \\ 2 & -3 & 5 & \vdots & 6 \\ 4 & -13 & 21 & \vdots & 26 \end{pmatrix}$$

6. Solve the following system of equations.

$$(i) \quad \begin{aligned} 2x_2 + 3x_3 - 4x_4 &= 1 \\ 2x_3 + 3x_4 &= 4 \\ 2x_1 + 2x_2 - 5x_3 + 2x_4 &= 4 \\ 2x_1 - 6x_3 + 9x_4 &= 7 \end{aligned}$$

$$(ii) \quad \begin{aligned} x + 2y + 3z &= 9 \\ 2x - y + z &= 8 \\ 3x - z &= 3 \end{aligned}$$

$$(iii) \quad \begin{aligned} x + y + 2z - 5w &= 3 \\ 2x + 5y - z - 9w &= -3 \\ 2x + y - z + 3w &= -11 \\ x - 3y + 2z + 7w &= -5 \end{aligned}$$

$$(iv) \quad \begin{aligned} x + y + z + w &= 4 \\ x + y + z - w &= 2 \\ x - y + z - w &= 0 \end{aligned}$$

$$\begin{aligned}
 \text{(v)} \quad & x_1 + 6x_2 + 3x_3 + 8x_4 = 0 \\
 & 2x_1 + 4x_2 + 6x_3 - x_4 = 0 \\
 & 3x_1 + 10x_2 + 9x_3 + 7x_4 = 0 \\
 & 4x_1 + 16x_2 + 12x_3 + 15x_4 = 0
 \end{aligned}$$

$$\begin{aligned}
 \text{(vi)} \quad & x_1 - 3x_2 + 2x_3 = 0 \\
 & 7x_1 - 21x_2 + 14x_3 = 0 \\
 & -3x_1 + 9x_2 - 6x_3 = 0
 \end{aligned}$$

$$\begin{aligned}
 \text{(vii)} \quad & 2x + 4y - 5z = 0 \\
 & x - 5y + 8z = 4 \\
 & 3x + 13y - 18z = 4
 \end{aligned}$$

7. Find the values of k for which the resulting linear system has

- (a) No solution
- (b) Unique solution
- (c) Infinitely many solutions

$$\begin{aligned}
 \text{(i)} \quad & x + y = 3 \\
 & x + (k^2 - 8)y = k
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad & x + y - z = 2 \\
 & x + 2y + z = 3 \\
 & x + y + (k^2 - 5)z = k
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad & x + y + z = 2 \\
 & 2x + 3y + 2z = 5 \\
 & 2x + 3y + (k^2 - 1)z = k + 1
 \end{aligned}$$

8. Construct 4 different augmented matrices for linear systems whose solution set is

- (i) $x_1 = 4, x_2 = -5$
- (ii) $x_1 = -3, x_2 = 2, x_3 = -1$

9. If a system $Ax = b$ of linear equations over \mathbb{R} has two distinct solutions u and v then show that there exists infinitely many solutions.

10. Show that the equations $2x + 4y = 3, x + 2y = 4$ are consistent over the field \mathbb{Z}_5 but inconsistent over \mathbb{R} .

11. Show that the equations $2x + 4y = 1, 4x + 3y = 2$ have a unique solution in \mathbb{R} . Do they have a unique solution over \mathbb{Z}_5 ? If yes, prove it. If no, find all solutions.

10.6 Row Reduction and Echelon Forms

In this section we shall refine the method studied above into a row reduction algorithm which will enable us to answer the fundamental existence and uniqueness question of the solution of a system of linear equation. The general systematic procedures for finding the solutions will be explained by an example.

Example 10.6. *Solve the system of linear equations*

$$\begin{aligned} 2y + 4z &= 2 \\ x + 2y + 2z &= 3 \\ 3x + 4y + 6z &= -1 \end{aligned}$$

We can work with the augmented matrix only. However, to compare the operations on the system of linear equations with those on the augmented matrix, we work on the system and the augmented matrix in parallel. The augmented matrix of the system is

$$\left(\begin{array}{cccc|c} 0 & 2 & 4 & \vdots & 2 \\ 1 & 2 & 2 & \vdots & 3 \\ 3 & 4 & 6 & \vdots & -1 \end{array} \right)$$

Step 1 Since the coefficient of x in the first equation is zero, while that in the second equation is non-zero, we interchange these two equations. Thus, we get

$$\begin{aligned} x + 2y + 2z &= 3 \\ 2y + 4z &= 2 \\ 3x + 4y + 6z &= -1 \end{aligned}$$

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & \vdots & 3 \\ 0 & 2 & 4 & \vdots & 2 \\ 3 & 4 & 6 & \vdots & -1 \end{array} \right)$$

Step 2 Using the 1st equation, we eliminate x from the 3rd equation. To do this we add (-3) times the first equation to the equation.

$$\begin{aligned} x + 2y + 2z &= 3 \\ 2y + 4z &= 2 \\ -2y &= -10 \end{aligned}$$

$$\left(\begin{array}{cccc|c} 1 & 2 & 2 & \vdots & 3 \\ 0 & 2 & 4 & \vdots & 2 \\ 0 & -2 & 0 & \vdots & -10 \end{array} \right)$$

Thus x is eliminated from the 2nd and 3rd equations. The coefficient of x in the 1st equation(row) is called the first pivot. In this case it is 1. The second and third equations have two unknowns y and z . Leave the first equation(row) alone,

and the same elimination procedure is applied to the second and third equations (rows). The pivot to eliminate y from the third equation is the coefficient (in this case 2) of y in the second equation (row).

Step 3 Add the second equation (row) to the third equation (row):

$$\textcircled{1}x + 2y + 2z = 3 \quad (10.3)$$

$$\textcircled{2}y + 4z = 2 \quad (10.4)$$

$$\textcircled{4}z = -8 \quad (10.5)$$

$$\begin{pmatrix} \boxed{1} & 2 & 2 & \vdots & 3 \\ 0 & \boxed{2} & 4 & \vdots & 2 \\ 0 & 0 & \boxed{4} & \vdots & -8 \end{pmatrix}$$

The elimination process (steps 1 to 3) done above is called forward elimination. The process is called **Gaussian elimination**.

Step 4 Normalize the non-zero rows by dividing them with their pivots. Thus the pivots become 1, and we get

$$x + 2y + 2z = 3$$

$$y + 2z = 1$$

$$z = -2$$

$$\begin{pmatrix} 1 & 2 & 2 & \vdots & 3 \\ 0 & 1 & 2 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & -2 \end{pmatrix}$$

Step 5 The last equation gives $z = -2$. Substituting $z = -2$ into the second equation gives $y = 5$. Putting the values of y and z in the first equation we get $x = -3$. This process is called **back substitution**.

This computation is shown below, that is, eliminating numbers above the leading 1s. Adding -2 times the third equation (row) to the second and the first equations (rows),

$$x + 2y = 7$$

$$y = 5$$

$$z = -2$$

$$\begin{pmatrix} 1 & 2 & 0 & \vdots & 7 \\ 0 & 1 & 0 & \vdots & 5 \\ 0 & 0 & 1 & \vdots & -2 \end{pmatrix}$$

Adding -2 times the second equation (row) to the first equation (row)

$$x = -3$$

$$y = 5$$

$$z = -2$$

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & -3 \\ 0 & 1 & 0 & \vdots & 5 \\ 0 & 0 & 1 & \vdots & -2 \end{pmatrix}$$

The whole process to obtain the solution is called **Gauss Jordan elimination method**.

Thus by applying a finite sequence of elementary row operations, the augmented matrix of the system of linear equations can be transformed into triangular form, which is row equivalent to the original augmented matrix. By applying back substitution, the solution of this system and therefore of the original system is obtained. The triangular form of the matrix can be further simplified by applying row operations, into a simpler matrix from which it is easy to decide whether the system is consistent or not, and if it is consistent, the solution(or solutions) is obtained. No back substitution is needed.

We shall now illustrate the procedure to reduce a given matrix to the simpler forms described above. Before doing this we shall introduce some terminology.

Definition 10.1. Let A be a $m \times n$ matrix with r non-zero rows, $0 \leq r \leq m$. Then A is in echelon form(or row echelon form) if it has the following properties

1. The first r rows of A are non-zero(and the last $m - r$ rows are rows of zeros)
2. Suppose the first non-zero element in the i th row occurs in the p_i th position, for $i = 1, 2, \dots, r$; then
 $p_1 < p_2 < \dots < p_r$

a_{ip_i} is called the leading entry of the i th row.

Example 10.7. Consider the following matrices

$$A = \begin{pmatrix} \boxed{1} & 0 & 0 & 2 \\ 0 & \boxed{4} & 0 & 3 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} \boxed{11} & 0 & 0 & 2 \\ 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{14} \end{pmatrix}$$

$$C = \begin{pmatrix} \boxed{1} & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{11} & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} \boxed{1} & 0 & 0 & 2 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & \boxed{31} & 0 \end{pmatrix}$$

$$E = \begin{pmatrix} 0 & \boxed{1} & 0 & 0 & -2 & 4 \\ 0 & 0 & \boxed{1} & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & \boxed{1} & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$F = \begin{pmatrix} 0 & 0 & \boxed{4} \downarrow & 3 & 5 & 7 & 2 \\ 0 & 0 & 0 \rightarrow & 0 \rightarrow & \boxed{21} \downarrow & -2 & 8 \\ 0 & 0 & 0 & 0 & 0 \rightarrow & \boxed{13} \downarrow & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 \rightarrow & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The leading entries of the non-zero rows are indicated in boxes. Consider the matrix A , which is a 4×4 matrix. Here $r = 3$. The zero row is at the bottom. $p_1 = 1$, $p_2 = 2$, $p_3 = 3$ so that $p_1 < p_2 < p_3$ is satisfied. Thus A is in row echelon form.

Consider the matrix B , which is a 4×4 matrix. There are no rows of zeros. $p_1 = 1$, $p_2 = 2$, $p_3 = 3$, $p_4 = 4$ so that $p_1 < p_2 < p_3 < p_4$ is satisfied. Thus B is in row echelon form.

Consider the matrix C , which is a 5×4 matrix. Here $r = 3$. There are two rows of zero rows. Thus condition 1 is not satisfied. Hence C is not in row echelon form.

D is a 3×4 matrix in which condition 1 is satisfied. $p_1 = 1$, $p_2 = 4$, $p_3 = 3$. Thus $p_2 \not< p_3$ so that condition 2 is not satisfied. Thus D is not in row echelon form.

E is a 5×6 matrix. Here $r = 3$. Condition 1 is satisfied. $p_1 = 2$, $p_2 = 3$, $p_3 = 5$. Thus $p_1 < p_2 < p_3$ so condition 2 is satisfied so that E is in row echelon form.

F is a 5×7 matrix. Here $r = 4$. Condition 1 is satisfied. $p_1 = 3$, $p_2 = 5$, $p_3 = 6$, $p_4 = 7$. Thus $p_1 < p_2 < p_3 < p_4$ so that condition 2 is satisfied. Hence F is in echelon form.

Remark 10.1.

1. The matrices C and D , which are not in row echelon form, can be transformed to this form by applying suitable row operations. Try yourself.
2. Note that we start with the leading entry in the upper left corner and move down by one step (so that we come to the next non-zero row) and move forward along that row to catch the leading entry of that row. Continue this process till the leading entry of the last non-zero row is reached. If this movement gives a staircase pattern then it is in row echelon form.

Example 10.8. $\begin{pmatrix} \boxed{4} \downarrow & 0 & 10 & 3 \\ 0 \rightarrow & \boxed{2} \downarrow & 3 & 4 \\ 0 & 0 \rightarrow & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, staircase pattern

A is a 4×4 matrix with 2 non-zero rows so that $r = 2$. The zero rows are at the bottom. $p_1 = 1, p_2 = 2$ so that $p_1 < p_2$. Thus A is a row echelon form. $a_{1p_1} = a_{11} = 1; a_{2p_2} = a_{22} = 1$. Thus leading entries are 1, so condition 2 is satisfied. Also in the 1st column and 2nd column, the leading 1 is only non-zero entry. Hence condition 3 is satisfied so that A is in reduced row echelon form.

B is a 5×10 matrix, with $r = 5$. $p_1 = 2, p_2 = 4, p_3 = 5, p_4 = 6, p_5 = 9$ so that $p_1 < p_2 < p_3 < p_4 < p_5$. Thus C is in row echelon form.

$a_{1p_1} = a_{12} = 1; a_{2p_2} = a_{24} = 1; a_{3p_3} = a_{35} = 1; a_{4p_4} = a_{46} = 1; a_{5p_5} = a_{59} = 1$. Thus leading entries are 1, so condition 2 is satisfied. Also each leading 1 is the only non-zero entry in that column. Hence condition 3 is satisfied so that B is in reduced row echelon form.

$$C = \begin{pmatrix} 0 & \boxed{1} & \frac{-1}{3} & 0 & \frac{2}{5} \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

C is a 4×5 matrix, with two non-zero rows, so that $r = 2$. $p_1 = 2, p_2 = 5$ so that $p_1 < p_2$. Thus B is a row echelon form.

$a_{1p_1} = a_{12} = 1; a_{2p_2} = a_{25} = 1$. Thus leading entries are 1. In the 2nd column the leading entry is the only non-zero entry in that column but in 5th column the leading entry is not the only non-zero entry. Hence C is not in reduced echelon form.

Remark 10.2. If a $m \times n$ matrix A has r non-zero rows in the reduced echelon form, then the number of leading 1's is r . The columns containing the leading 1's form the 1st r columns of the $m \times m$ unit matrix in order. That is, these are,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ } i\text{th position}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \text{ } m\text{th position of } I_m$$

$$\text{If } A = \begin{pmatrix} 0 & \boxed{1} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

A is in reduced echelon form. The leading 1's of the 1st and 2nd row are in boxes. The corresponding columns are $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, which are the 1st two columns of I_4 .

Example 10.10. The following matrices are in echelon form. The leading entries (denoted by ■) may have any non-zero value; the starred entries (*)

may have any value including zero.

$$A = \begin{pmatrix} \boxed{\blacksquare} & \boxed{*} & * & \boxed{*} \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

$$B = \begin{pmatrix} 0 & \boxed{\blacksquare} & * & \boxed{*} & * & \boxed{*} & \boxed{*} & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{pmatrix}$$

The columns containing the leading entries are enclosed within a box.

The following matrices are in reduced echelon form. The leading entries are 1 and there are zeros below and above each leading 1.

$$C = \begin{pmatrix} \boxed{1} & \boxed{0} & * & \boxed{0} \\ 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

$$D = \begin{pmatrix} 0 & \boxed{1} & * & \boxed{0} & * & \boxed{0} & \boxed{0} & * \\ 0 & 0 & 0 & 1 & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{pmatrix}$$

The columns containing the leading 1's are shaded. In C they are the 1st three columns of I_5 . In D they are the four columns of I_4 .

Any non-zero matrix may be row reduced (i.e. transformed by E-row operations) to echelon form by using different sequences of row operations.

Theorem 10.1. Every $m \times n$ matrix is row equivalent to a matrix in row echelon form.

In fact, the row echelon form is not unique as will be shown later. The row echelon form may be further reduced by a sequence of E-row operations to a form which is unique. This is called the reduced row echelon form.

Theorem 10.2. Every matrix is row equivalent to one and only one reduced echelon matrix.

If a matrix A is row equivalent to an echelon matrix U , we call U an **echelon form** (or row echelon form) of A , if U is in reduced echelon form, we call U the **reduced echelon form** of A .

When row operations on a matrix produce an echelon form; further row operations to obtain the reduced echelon form do not change the position of the leading entries. Since the reduced echelon form is unique, the leading entries are always in the same position in any echelon form obtained from a given matrix. These leading entries correspond to leading 1's in the reduced echelon form.

Definition 10.3. A **pivot position** in a matrix A is a location in A , that corresponds to a leading 1 in the reduced echelon form of A . A **pivot column** is a column of A that contains a pivot position.

Note that the pivot position in a matrix are the same as the position of the leading entries in its echelon form. The columns containing these leading entries are the pivot columns. Thus to locate the pivot columns it is sufficient to reduce the matrix to its echelon form.

Note For a $m \times n$ matrix we make the following observations:

1. The number of pivots is always less than or equal to $\min(m, n)$.
2. Each row and each column can have at most one pivot element.
3. For the matrix to be in echelon form, the elements above the pivot element can be any number (denoted by *), zero or non-zero, whereas those below it must be 0, the elements to the right of a pivot can be *, whereas those to the left of it must be 0's. If \blacksquare indicates a pivot element, then diagrammatically this can be represented as:

$$\begin{array}{c} *'s \\ \uparrow \\ 0's \leftarrow \blacksquare \rightarrow *'s \\ \downarrow \\ 0's \end{array}$$

4. When a matrix is in echelon form, the movement from one point to another is described as follows:

We can move only downwards from a pivot by one step. Horizontally, we may move only to the right and it can be any number of steps.

We shall now give a systematic procedure to transform a non-zero matrix to row echelon form.

Example 10.11. Transform the following matrix to row echelon form

$$A = \begin{pmatrix} 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 2 & 2 & -5 & 2 & 4 \\ 0 & 2 & 0 & -6 & 9 & 7 \end{pmatrix},$$

We shall explain the procedure adopted to transform A to row echelon form in a stepwise manner.

Step 1 Counting from left to right, find the first non-zero column (i.e. a column having atleast one non-zero entry). This column is called the pivotal column. For A the second column is the pivotal column

$$A = \begin{pmatrix} 0 & \boxed{0} & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 2 & 2 & -5 & 2 & 4 \\ 0 & 2 & 0 & -6 & 9 & 7 \end{pmatrix}$$

Step 2 Counting from top to bottom in the pivotal column, identify the first non-zero entry. The element is called the pivot. We encircle it in A

$$A = \begin{pmatrix} 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & \boxed{2} & 2 & -5 & 2 & 4 \\ 0 & 2 & 0 & -6 & 9 & 7 \end{pmatrix},$$

Step 3 Bring the pivot element to the first row. This may require an interchange of rows. In our case we need to apply $R_1 \leftrightarrow R_3$. Thus

$$A \sim \begin{pmatrix} 0 & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 2 & 0 & -6 & 9 & 7 \end{pmatrix} = A_1$$

Step 4 In the pivotal column, below the pivot, zero. This is done by adding a suitable multiples of the first row to the subsequent rows. In our case, we apply. Thus $R_4 \rightarrow R_4 + (-1)R_1$ to A_1

$$A \sim \begin{pmatrix} 0 & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & -2 & -1 & 7 & 3 \end{pmatrix} = A_2(\text{say})$$

Step 5 Identify B as the $(m-1) \times n$ submatrix of A_2 obtained by hiding the first row of A_2 . Repeat steps 1 to 4 on B . Be careful not to erase the first row of A_2 .

In this case

$$B = \begin{pmatrix} \boxed{0} & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & -2 & -1 & 7 & 3 \end{pmatrix}$$

Apply $R_1 \leftrightarrow R_2$ to B

$$B_1 = \begin{pmatrix} \boxed{0} & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & -2 & -1 & 7 & 3 \end{pmatrix}$$

Apply $R_3 \rightarrow R_3 + R_1$ to B_1

$$B_2 = \begin{pmatrix} \boxed{0} & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 2 & 3 & 4 \end{pmatrix}$$

Step 6 In the light of continuing the process we now hide the top two rows of the above matrix (take the full matrix and not just B_2). This amounts to hiding an additional top row of B_2 .

$$\begin{pmatrix} \boxed{0} & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 2 & 3 & 4 \end{pmatrix}$$

On C , apply $R_4 \rightarrow R_4 + (-1)R_3$

$$C_1 = \begin{pmatrix} \boxed{0} & \boxed{2} & \boxed{2} & \boxed{-5} & \boxed{2} & \boxed{4} \\ \boxed{0} & \boxed{0} & \boxed{2} & \boxed{3} & \boxed{-4} & \boxed{1} \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Step 7 Now hide the top 3 rows of the above matrix, taking the full matrix. This amounts to hiding an additional top row of C_1 . Thus

$$D = \begin{pmatrix} \boxed{0} & \boxed{2} & \boxed{2} & \boxed{-5} & \boxed{2} & \boxed{4} \\ \boxed{0} & \boxed{0} & \boxed{2} & \boxed{3} & \boxed{-4} & \boxed{1} \\ \boxed{0} & \boxed{0} & \boxed{0} & \boxed{2} & \boxed{3} & \boxed{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Since D has no pivotal column so the process is completed. Thus

$$A \sim \begin{pmatrix} 0 & 2 & 2 & -5 & 2 & 4 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = E$$

and the matrix E is in echelon form.

The pivot columns are columns 2, 3 and 4 of E . The general form of E is

$$\begin{pmatrix} 0 & \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The columns 2, 3 and 4 of A are the pivot columns and the pivot positions of A are the positions corresponding to the leading entries of E .

$$A = \begin{pmatrix} 0 & \boxed{0} & 2 & 3 & -4 & 1 \\ 0 & 0 & \boxed{0} & 2 & 3 & 4 \\ 0 & 2 & -2 & \boxed{-5} & 2 & 4 \\ 0 & 2 & 0 & -6 & 9 & 7 \end{pmatrix}$$

A pivot (as illustrated above) is a non-zero number in a pivot position which is used to create zeros using row operations. In the above example, the pivots are 2, 2 and 2. Note that these numbers are not the same as the actual elements of A in the highlighted pivot positions as shown. In fact if a different sequence of row operations is used, we will get a different set of pivots.

Remark 10.3. Row echelon form of a matrix is not unique. For instance, if we apply the operations $R_1 \rightarrow R_1 + R_2$ to E , (in the above example) then

$$A \sim \begin{pmatrix} 0 & 2 & 4 & -2 & -2 & 5 \\ 0 & 0 & 2 & 3 & -4 & 1 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = F(\text{say})$$

F is also in row echelon form. Thus E and F are both in echelon form and are equivalent to the matrix A .

Example 10.12. Reduce the matrix in the above illustration to reduced echelon form.

Going through steps 1-7 as in the above illustration, we first reduce the given matrix to row echelon form. Thus

$$A \sim \begin{pmatrix} 0 & \boxed{2} & 2 & -5 & 2 & 4 \\ 0 & 0 & \boxed{2} & 3 & -4 & 1 \\ 0 & 0 & 0 & \boxed{2} & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = E$$

The pivots have been encircled.

Step 8 Starting with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

The rightmost pivot in row 3 is 2. Make it 1 by applying $R_3 \rightarrow \frac{1}{2}R_3$ to E . Thus

$$A \sim \begin{pmatrix} 0 & \boxed{2} & 2 & -5 & 2 & 4 \\ 0 & 0 & \boxed{2} & 3 & -4 & 1 \\ 0 & 0 & 0 & \boxed{1} & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Create zeros in the column 4 by applying $R_1 \rightarrow R_1 + 5R_3$, $R_2 \rightarrow R_2 - 3R_3$.

$$A \sim \begin{pmatrix} 0 & 2 & 2 & 0 & \frac{19}{2} & 14 \\ 0 & 0 & \boxed{2} & 0 & -\frac{17}{2} & -5 \\ 0 & 0 & 0 & \boxed{1} & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The next pivot is in row 2 and is 2. Scale this row to make pivot as 1. Applying $R_2 \rightarrow \frac{1}{2}R_2$, we get

$$A \sim \begin{pmatrix} 0 & \boxed{2} & 2 & 0 & \frac{19}{2} & 14 \\ 0 & 0 & \boxed{1} & 0 & -\frac{17}{4} & -\frac{5}{2} \\ 0 & 0 & 0 & \boxed{1} & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Create zeros in the 3rd column above the pivot element $R_1 \rightarrow R_1 - 2R_2$. Thus

$$A \sim \begin{pmatrix} 0 & \boxed{2} & 0 & 0 & 18 & \frac{33}{2} \\ 0 & 0 & \boxed{1} & 0 & -\frac{17}{4} & -\frac{5}{2} \\ 0 & 0 & 0 & \boxed{1} & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The next pivot (which is the last pivot) is in row 1 and it is 2. Scale this row to make the pivot as 1. Apply $R_1 \rightarrow \frac{1}{2}R_1$. Thus

$$A \sim \begin{pmatrix} 0 & \boxed{1} & 0 & 0 & 9 & \frac{33}{4} \\ 0 & 0 & \boxed{1} & 0 & -\frac{17}{4} & -\frac{5}{2} \\ 0 & 0 & 0 & \boxed{1} & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = G$$

Now G is in reduced echelon form. Hence A is transformed to a matrix in reduced echelon form.

Problem 10.11. Find a matrix in row echelon form that is row equivalent to the given matrix. Give 2 possible answers.

$$A = \begin{pmatrix} 2 & -1 & 0 & 1 & 4 \\ 1 & -2 & 1 & 4 & -3 \\ 5 & -4 & 1 & 6 & 5 \\ -7 & 8 & -3 & \frac{-1}{4} & 1 \end{pmatrix}$$

Solution:

Step 1 Counting from left, the 1st non-zero column is the pivot column. The pivot position is at the top. For ease of calculation apply $R_1 \leftrightarrow R_2$ (so that the pivot becomes 1)

$$A \sim \begin{pmatrix} \boxed{1} & -2 & 1 & 4 & -3 \\ 2 & -1 & 0 & 1 & 4 \\ 5 & -4 & 1 & 6 & 5 \\ -7 & 8 & -3 & -14 & 1 \end{pmatrix} = A_1(\text{say})$$

Apply $R_2 \rightarrow R_2 - 2R_1$, $R_3 \rightarrow R_3 - 5R_1$, $R_4 \rightarrow R_4 + 7R_1$

to make all the entries below the pivot in the pivotal column zero. Then

$$A \sim \begin{pmatrix} 1 & -2 & 1 & 4 & -3 \\ 0 & \boxed{3} & -2 & -7 & 10 \\ 0 & 6 & -4 & -14 & 20 \\ 0 & -6 & 4 & 14 & -20 \end{pmatrix} = A_2(\text{say})$$

Step 2 Ignoring the 1st row of A_2 , the 2nd column is the pivotal column and the element in the (2,2)th position is the pivot. It is circled in A_2 . To make all the entries, below the pivotal column zero, apply, $R_3 \rightarrow R_3 - 2R_2$, $R_4 \rightarrow R_4 + 2R_2$. Thus

$$A \sim \begin{pmatrix} 1 & -2 & 1 & 4 & -3 \\ 0 & 3 & -2 & -7 & 10 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A_3(\text{say})$$

Step 3 In A_3 the last two rows are rows of zeros. Thus A has been reduced to row echelon form. To get another row echelon form, apply a row operation on A_3 so that the work done above (i.e., that of making zeros) is not destroyed.

$$A \sim \begin{pmatrix} 1 & -2 & 1 & 4 & -3 \\ 0 & 6 & -4 & -14 & 20 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A_4(\text{say})$$

A_3, A_4 are in row echelon form and they are row equivalent to A . These are two possible answers.

Problem 10.12. Find the matrix in reduced row echelon form that is row equivalent to the given matrix (encircle the pivot positions in the final matrix) and in the original matrix and in the original matrix. Also list the pivot columns.

$$\begin{pmatrix} 1 & -2 & 0 & 2 \\ 2 & -3 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 1 & 1 & 0 & 2 \\ 2 & -6 & -2 & 1 \end{pmatrix}$$

Solution:

Problem 10.13. Determine whether the given matrices are in reduced echelon form, row echelon form or neither.

$$(i) A = \begin{pmatrix} 1 & 1 & 0 & 4 & 0 & 2 \\ 0 & 3 & 0 & 6 & 0 & 3 \\ 0 & 0 & 0 & -5 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

$$(ii) B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$(iii) C = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(iv) D = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(v) E = \begin{pmatrix} 0 & 1 & 12 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(vi) F = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Solution:

(i) A is a 4×6 matrix and it has no rows of zeros. The leading entries of non-zero rows are circled.

$$A = \begin{pmatrix} \boxed{1} \downarrow & 1 & 0 & 4 & 0 & 2 \\ 0 \rightarrow & \boxed{3} \downarrow & 0 & 6 & 0 & 3 \\ 0 & 0 \rightarrow & 0 \rightarrow & \boxed{-5} \downarrow & 0 & 4 \\ 0 & 0 & 0 & 0 \rightarrow & 0 \rightarrow & \boxed{7} \end{pmatrix}$$

The route from one leading entry to the other is shown by arrows. Since it forms a staircase, therefore it is in echelon form.

Since all the leading entries are not one (the leading entry of 2nd row is 3) therefore it is not in reduced echelon form.

Thus A is in echelon form only.

(ii) B is a 4×4 matrix. It has a row of zeros, which is not the bottom most row. Hence it is not in echelon form.

(iii) C is a 3×3 matrix. It has a row of zeros, which is the bottommost row.

The leading entries are circled.
$$\begin{pmatrix} \boxed{1} \downarrow & 2 & 0 \\ 0 \rightarrow & 0 \rightarrow & \boxed{1} \\ 0 & 0 & 0 \end{pmatrix}$$

The route from one leading entry to the other is shown by arrows. Since it forms a staircase, therefore it is in echelon form. Moreover the leading

entries are 1 and the pivot columns are $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, i.e. all the elements above and below the leading entry are zero. Thus it is in reduced echelon form also. Hence C is both in echelon as well as reduced echelon form.

(iv) D is a 5×6 matrix with 2 rows of zeros. These rows of zeros are in the bottommost position. Thus there are 3 pivot columns. The leading entries are circled.

$$\begin{pmatrix} 0 & \boxed{1} \downarrow & 0 & 0 & 0 & 0 \\ 0 & 0 \rightarrow & 0 \rightarrow & \boxed{4} \downarrow & 0 & 0 \\ 0 & 0 & 0 & 0 \rightarrow & 0 \rightarrow & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The route from one leading entry to the other, shown by arrows, forms a staircase. Hence, it is in echelon form. Since all the leading entries are not 1 (the 2nd leading entry is 4), therefore it is not in reduced echelon form.

Thus D is in echelon form

(v) E is a 4×5 matrix with 2 rows of zeros, which are the bottommost rows.

The leading entries are circled.

$$\begin{pmatrix} 0 & \boxed{1} \downarrow & 12 & 2 & 0 \\ 0 & 0 \rightarrow & 0 \rightarrow & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The route from one leading entry to the other is shown by arrows. Since it forms a staircase, therefore it is in echelon form. The leading entries are 1 and the pivot columns are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

The second pivot column is not $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

Hence it is not in the reduced echelon form.

Hence E is in the echelon form.

(vi) F is a 3×4 matrix with one row of which is in the bottommost positions.

The leading entries are circled.

$$\begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The elements below the first leading entry are not all zero, so that it is not in echelon form. Thus F is neither in echelon form nor in reduced echelon form.

Problem 10.14. Describe the possible echelon forms of a non-zero 2×3 matrix.

Solution: We shall use \blacksquare to indicate a leading non-zero entry and $*$ to indicate an entry which can take any value, including zero. For a 2×3 matrix there can be at most 2 pivot columns.

Two cases arise:

Case 1 There is only one pivot column. This pivot column can be the 1st column or the 2nd column or the 3rd column. This can be classified as follows:

Pivot column	Echelon form
1st	$\begin{pmatrix} \blacksquare & * & * \\ 0 & 0 & 0 \end{pmatrix}$
2nd	$\begin{pmatrix} 0 & \blacksquare & * \\ 0 & 0 & 0 \end{pmatrix}$
3rd	$\begin{pmatrix} 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{pmatrix}$

Case 2 There are two pivot columns. This can be classified as follows:

Pivot column	Echelon form
1st and 2nd	$\begin{pmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \end{pmatrix}$
1st and 3rd	$\begin{pmatrix} \blacksquare & * & * \\ 0 & 0 & \blacksquare \end{pmatrix}$
2nd and 3rd	$\begin{pmatrix} 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \end{pmatrix}$

10.7 Exercise

- Determine whether the given matrix is in reduced echelon form, row echelon form, or neither.

$$(i) A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

$$(ii) A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 15 \\ 0 & 0 & 1 & 0 & -14 \\ 0 & 0 & 0 & -1 & 13 \end{pmatrix}$$

$$(iii) A_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(iv) A_4 = \begin{pmatrix} 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & -10 \\ 0 & 0 & 0 & 1 & 41 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(v) A_5 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(vi) A_6 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(vii) A_7 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(viii) A_8 = \begin{pmatrix} 0 & 1 & 0 & 0 & 5 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 1 & 0 & -2 & 3 \end{pmatrix}$$

$$(ix) A_9 = \begin{pmatrix} 0 & 1 & 0 & 0 & 12 \\ 0 & 0 & 1 & 1 & -11 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(x) A_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 & -6 & 3 \\ 0 & 1 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(xi) A_{11} = \begin{pmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

2. Find a matrix in row echelon form that is equivalent to the given matrix. Give two possible answers in each.

$$(i) B_1 = \begin{pmatrix} 1 & -3 & 2 & 1 & 2 \\ 3 & -9 & 10 & 2 & 9 \\ 2 & -6 & 4 & 2 & 4 \\ 2 & -6 & 8 & 1 & 7 \end{pmatrix}$$

$$(ii) B_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}$$

$$(iii) B_3 = \begin{pmatrix} 0 & -1 & 2 & 3 \\ 2 & 3 & 4 & 5 \\ 1 & 3 & -1 & 2 \\ 3 & 2 & 4 & 1 \end{pmatrix}$$

$$(iv) B_4 = \begin{pmatrix} 1 & 2 & -3 & 1 \\ -1 & 0 & 3 & 4 \\ 0 & 1 & 2 & -1 \\ 2 & 3 & 0 & -3 \end{pmatrix}$$

$$(v) B_5 = \begin{pmatrix} 2 & 1 & 0 & 0 & 1 \\ 3 & 0 & 3 & 0 & 2 \\ 5 & 7 & -9 & 2 & 3 \end{pmatrix}$$

$$(vi) B_6 = \begin{pmatrix} 0 & 2 & 4 & 3 & 0 \\ 0 & 5 & 10 & 15/2 & 0 \\ 0 & 1 & 2 & 3/2 & 4 \\ 0 & 2 & 4 & 3 & 2 \end{pmatrix}$$

10.8 Vector Equations

Suppose a Hyundai showroom stocks five different models of cars: Santro, Ascent, Verna, i10 and i20. Each month the number of cars of each model in stock are recorded. If in the month of January 2009, there are 10 cars of Santro, 6 of Ascent, 4 of Verna, 7 of i10 and 8 of i20, then we can represent these stock

quantities by a column: $\begin{pmatrix} 10 \\ 6 \\ 4 \\ 7 \\ 8 \end{pmatrix}$ or a row: $(10 \ 6 \ 4 \ 7 \ 8)$

Such an ordered set of numbers, which is distinguished not only by the elements it contains, but also by the order in which they appear, is called a vector. If it is represented by a 1×5 matrix it is called a row vector, if it is represented as a 5×1 matrix it is called a column vector. Since it contains 5 entries it is also called a 5-vector. Each entry is called a component. Clearly

the vector $(10 \ 6 \ 4 \ 7 \ 8)$ and the column vector $\begin{pmatrix} 10 \\ 6 \\ 4 \\ 7 \\ 8 \end{pmatrix}$ contain exactly

the same information — the numbers and their order are the same, only the way they are written is different.

10.9 Vectors in \mathbb{R}^2

A matrix with only one column is called a column vector. They are denoted by boldface letters. If $u = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $v = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$, $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$, where w_1, w_2 are any real numbers, then u, v, w are vectors with two entries. The set of all vectors with two real entries is denoted by \mathbb{R}^2 . The \mathbb{R} stands for the set of real numbers that appear as entries in the vectors, and the exponent 2 indicates that there are two entries in each vector. These entries are also called the components of the vector.

Definition 10.4. Two vectors in \mathbb{R}^2 are equal if and only if their corresponding entries are equal. If $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ then $u = v$, iff $u_1 = v_1$ and $u_2 = v_2$.

Definition 10.5. If $u \in \mathbb{R}^2$, $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and c is any real number, then $cu = \begin{pmatrix} cu_1 \\ cu_2 \end{pmatrix}$. The number c is called a scalar and cu is called the scalar multiple of u by c .

Definition 10.6. If u and $v \in \mathbb{R}^2$, $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ then the sum of the vectors u and v denoted by $u+v$, is $\begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}$.

Thus to obtain the sum of two vectors we add their corresponding components.

Example 10.13. 1. If $u = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $v = \begin{pmatrix} 1+1 \\ 2+1 \end{pmatrix}$, is $u = v$?

Yes, because corresponding components are equal.

2. Find k such that $\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} k+2 \\ 2 \end{pmatrix}$.

The two vectors being equal we must have $-1 = k+2$ so that $k = -3$.

3. If $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$ find $2u$, $(-3)v$, $(-1)v$, $2u + (-3)v$.

$$2u = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \times 1 \\ 2 \times 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad (-3)v = (-3) \begin{pmatrix} -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ -9 \end{pmatrix}.$$

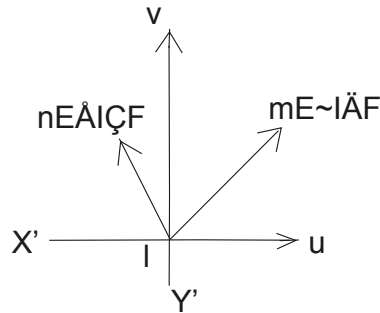
$$(-1)v = (-1) \begin{pmatrix} -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

$$2u + (-3)v = \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 6 \\ -9 \end{pmatrix} = \begin{pmatrix} 8 \\ -5 \end{pmatrix}.$$

10.10 Geometric Descriptions of \mathbb{R}^2

We consider a rectangular coordinate system $X'OX$, $Y'OY$ in the plane. With each element $\begin{pmatrix} a \\ b \end{pmatrix}$ of \mathbb{R}^2 , we associate the point $P(a, b)$ in the plane.

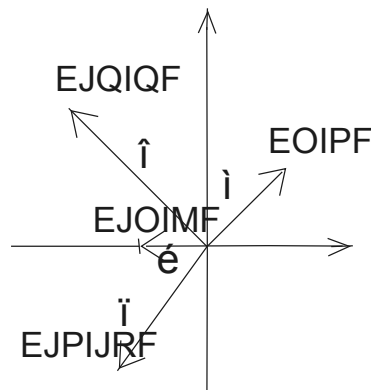
The point (a, b) gives rise to the directed line segment \overrightarrow{OP} , O being the initial point and P being the terminal point.



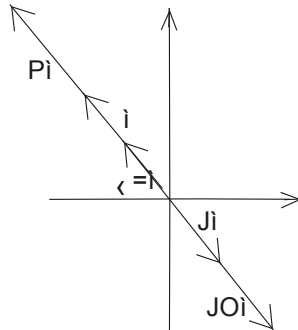
Conversely, if $Q(c, d)$ is any point in the plane then $\begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2$. Also with the directed line segment \overrightarrow{OQ} (with O as the initial point) we associate the vector $\begin{pmatrix} c \\ d \end{pmatrix}$ in \mathbb{R}^2 . Thus every vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ in \mathbb{R}^2 can be represented by a point $P(x_1, x_2)$ in the plane and also by a directed line segment with O as the initial point, $P(x_1, x_2)$ as termination point.

Example 10.14.

1. Represent $u = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $v = \begin{pmatrix} -4 \\ 4 \end{pmatrix}$, $w = \begin{pmatrix} -3 \\ -5 \end{pmatrix}$, $p = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$ by directed line segments.



2. If $u = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$, represent $3u$, $\frac{1}{2}u$, $(-1)u$, $(-2)u$ geometrically. Verify that $(-1)u = -u$ and $(-2)u = 2(-u)$.

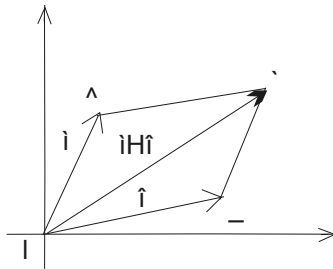


Remark 10.4. In general, the length of the arrow for cu is $|c|$ times the length of the arrow for u . If c is positive, the direction of cu is same as that of u , whereas if c is negative the direction of cu is opposite to the direction of u .

The sum of two vectors has a useful geometric representation. The following rule can be verified by analytic geometry.

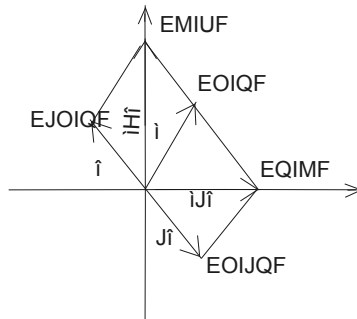
Parallelogram Rule for Addition

If u and $v \in \mathbb{R}^2$ are represented by \vec{OA} and \vec{OB} then $u + v$ corresponds to the diagonal \vec{OC} of the parallelogram with adjacent sides OA and OB .



Parallelogram Rule

Example 10.15. If $u = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$, $v = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$, represent $u + v$, $u - v$ geometrically.



Vectors in \mathbb{R}^3

Vectors in \mathbb{R}^3 are 3×1 matrices, vectors with 3 entries. Geometrically, they are represented by points in a three-dimensional coordinate space and by arrows from the origin.

10.11 Vectors in \mathbb{R}^n

If n is a positive integer, \mathbb{R}^n denotes the collection of all ordered n -tuples of real numbers, usually written as $n \times 1$ matrices, such as $u = \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_n \end{pmatrix}$.

u_1, u_2, \dots, u_n are called components of the vector u . A vector all of whose components are zero is called the zero vector. It is denoted by 0 . (The number of components in 0 will be clear from the context.)

Equality of vectors, scalar multiplication of a vector and addition of vectors is defined as in \mathbb{R}^2 but we will give a formal definition again.

Definition 10.7. Let u and $v \in \mathbb{R}^n$ where $u = \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_n \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{pmatrix}$.

Then $u = v$ iff $u_1 = v_1, u_2 = v_2, \dots, u_n = v_n$. Thus the two vectors are equal iff their corresponding components are equal.

Definition 10.8. If $u \in \mathbb{R}^n$ and c is any real number, then $cu = \begin{pmatrix} cu_1 \\ cu_2 \\ \cdot \\ \cdot \\ cu_n \end{pmatrix}$.

cu is called the scalar multiple of u by c .

To obtain cu , each component of u is multiplied by c .

Definition 10.9. Let u and $v \in \mathbb{R}^n$ where $u = \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_n \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{pmatrix}$.

Then $u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \cdot \\ \cdot \\ u_n + v_n \end{pmatrix}$, is the sum of the vectors u and v .

Thus the sum of two vectors is obtained by adding their corresponding components.

These operations on vectors have the following properties, which can be verified directly from the corresponding properties for real numbers.

Algebraic Properties of \mathbb{R}^n

For all u, v, w in \mathbb{R}^n and all scalars c and d .

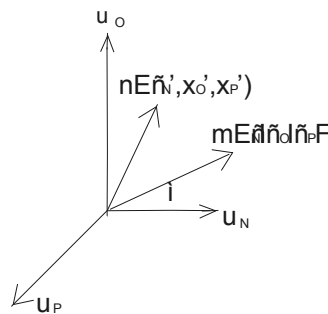
1. $u + v = v + u$
2. $(u + v) + w = u + (v + w)$
3. $u + 0 = 0 + u = u$
4. $u + (-u) = 0 = (-u) + u$
5. $c(u + v) = cu + cv$
6. $(c+d)u = cu + du$
7. $c(du) = (cd)u = d(cu)$
8. $1u = u$

where $-u$ denotes $(-1)u$. For simplicity of notation, we used ‘vector subtraction’ and write $u-v$ instead of $u+(-1)v$.

Points in \mathbb{R}^n

Let us first consider \mathbb{R}^3 . Let $u = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ be any vector in \mathbb{R}^3 . Then the point $P(x_1, x_2, x_3)$ is the head of the vector u . On the other hand, for any point $Q(x'_1, x'_2, x'_3)$ there corresponds the vector $\vec{OQ} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$.

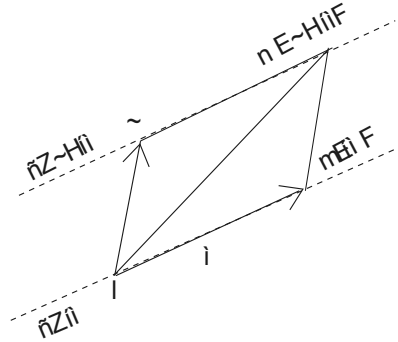
This correspondence can be extended to \mathbb{R}^n . For every point $P(x_1, x_2, \dots, x_n)$ in \mathbb{R}^n , there is an associated vector \vec{OP} in \mathbb{R}^n , and conversely. Because of this correspondence the point P associated with the vector u is denoted by $P(u)$.



Lines in \mathbb{R}^n

Let u be any vector. For any scalar t , tu is a vector along u . As t varies the points tu lie along the line containing u . Thus, $x = tu$, $t \in \mathbb{R}$ represents a line containing the vector u . When we add a fixed vector a to any vector tu on this line corresponding to the point $P(tu)$ we get a point $Q(a + tu)$ in such a way that $\vec{PQ} = a$. Thus all the points $a + tu$, $t \in \mathbb{R}$ lie on a line parallel to u and

passing through a . Thus $x = a + tu$, $t \in \mathbb{R}$ represents a line through the point $P(a)$ and parallel to u .



Planes in \mathbb{R}^n

Let u, v be any two non-collinear vectors. For any scalars α, β , the vector $\alpha u + \beta v$ is the diagonal of the parallelogram whose adjacent sides are $\alpha u, \beta v$. Therefore, $\alpha u + \beta v$ is a vector in the plane containing the vectors u and v for all scalars α and β . Thus

$$x = \alpha u + \beta v, \alpha, \beta \in \mathbb{R}$$

represents a plane containing vectors u and v . When we add a fixed vector a to $\alpha u + \beta v$, the point P associated with $\alpha u + \beta v$ is shifted to Q in such a way that $\overrightarrow{PQ} = a$. Thus all points $a + \alpha u + \beta v$ with $\alpha, \beta \in \mathbb{R}$ lie on the plane parallel to the plane $x = \alpha u + \beta v, \alpha, \beta \in \mathbb{R}$ and containing the point $A(a)$.

$x = a + \alpha u + \beta v, \alpha, \beta \in \mathbb{R}$ represents a plane through the point $A(a)$ and parallel to the plane $x = \alpha u + \beta v, \alpha, \beta \in \mathbb{R}$.

Linear Combination of Vectors

Definition 10.10. If $u_1, u_2, \dots, u_k \in \mathbb{R}^n$ and c_1, c_2, \dots, c_k are scalars then the vector $c_1 u_1 + c_2 u_2 + \dots + c_k u_k$ is called a linear combination of u_1, u_2, \dots, u_k . The scalars c_1, c_2, \dots, c_k are called weights in the linear combination.

Example 10.16.

1. If $v_1, v_2 \in \mathbb{R}^n$ then $\sqrt{5}v_1 - 3v_2, -2v_1 + \frac{3}{2}v_2, \frac{2}{3}v_2 = (0v_1 + \frac{2}{3}v_2), 0 = (0v_1 + 0v_2)$ are linear combinations of v_1, v_2 .

2. If $u, v \in \mathbb{R}^3, u = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}, v = \begin{pmatrix} 0 \\ 4 \\ -5 \end{pmatrix}$ then $-2u + 3v = -2 \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 4 \\ -5 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ -6 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \\ -15 \end{pmatrix} = \begin{pmatrix} -4 \\ 14 \\ -21 \end{pmatrix}$

Also, $0 = 0u + 0v$. Thus 0 is a linear combination of u and v . Note that 0 is always a linear combination of any set of vectors as all the weights can be taken to be zero.

Example 10.17. Express $\begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}$ as a linear combination of $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$.

$$\text{Let } u_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, b = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}.$$

In order to express b as a linear combination of u_1 and u_2 , we need to find x_1 and x_2 such that $x_1u_1 + x_2u_2 = b \dots(1)$.

$$\text{Thus } x_1 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}$$

$$\text{ie. } \begin{pmatrix} x_1 + x_2 \\ x_1 + 2x_2 \\ 2x_1 + 5x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}$$

The two vectors are equal if and only if the corresponding components are equal. Hence

$$\begin{aligned} x_1 + x_2 &= 5 \\ x_1 + 2x_2 &= 3 \\ 2x_1 + 5x_2 &= 4 \end{aligned}$$

We shall solve this system by reducing the augmented matrix to reduced echelon form:

$$\left(A : b \right) = \left(\begin{array}{cc|c} 1 & 1 & 5 \\ 1 & 2 & 3 \\ 2 & 5 & 4 \end{array} \right)$$

$$\text{Applying } R_2 \rightarrow R_2 - R_1 \text{ and } R_3 \rightarrow R_3 - 2R_1 \sim \begin{pmatrix} 1 & 1 & \vdots & 5 \\ 0 & 1 & \vdots & -2 \\ 0 & 3 & \vdots & -6 \end{pmatrix}$$

$$\text{Applying } R_3 \rightarrow R_3 - 3R_2 \sim \begin{pmatrix} 1 & 1 & \vdots & 5 \\ 0 & 1 & \vdots & -2 \\ 0 & 0 & \vdots & 0 \end{pmatrix} R_1 \rightarrow R_1 - R_2 \sim \begin{pmatrix} 1 & 0 & \vdots & 7 \\ 0 & 1 & \vdots & -2 \\ 0 & 0 & \vdots & 0 \end{pmatrix}$$

Thus the solution is

$$x_1 = 7, x_2 = -2$$

Hence b is a linear combination of u_1 and u_2 .

In fact

$$7 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}$$

Note that in the above illustration, the original vectors u_1, u_2, b form the columns of the augmented matrix which is to be reduced to the reduced echelon form.

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

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ u_1 & u_2 & b \end{array}$$

Example 10.18. Is $\begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix}$ a linear combination of $\begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ -4 \\ -1 \end{pmatrix}$, and $\begin{pmatrix} 1 \\ -5 \\ 7 \end{pmatrix}$?

Let $u_1 = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$, $u_2 = \begin{pmatrix} 2 \\ -4 \\ -1 \end{pmatrix}$, $u_3 = \begin{pmatrix} 1 \\ -5 \\ 7 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix}$

We need to find weights x_1, x_2, x_3 (if possible) such that $x_1 u_1 + x_2 u_2 + x_3 u_3 = b$

In view of the above discussion, the weights x_1, x_2 and x_3 are the solution of the linear system whose augmented matrix is $\begin{pmatrix} u_1 & u_2 & u_3 & : & b \end{pmatrix}$

$$= \begin{pmatrix} 1 & 2 & 1 & : & 2 \\ -3 & -4 & -5 & : & -5 \\ 2 & -1 & 7 & : & 3 \end{pmatrix} \text{ which is equivalent to } \begin{pmatrix} 1 & 2 & 1 & : & 2 \\ 0 & 2 & -2 & : & 1 \\ 0 & -5 & 5 & : & -1 \end{pmatrix} \text{ by } R_2 \rightarrow R_2 + 3R_1 \text{ and } R_3 \rightarrow R_3 - 2R_1$$

which is equivalent to $\begin{pmatrix} 1 & 2 & 1 & : & 2 \\ 0 & 2 & -2 & : & 1 \\ 0 & 0 & 0 & : & 3/2 \end{pmatrix}$ by $R_3 \rightarrow R_3 + 5/2R_2$

This leads to an inconsistent system as there is a pivot element in the augmented column. Thus b is not a linear combination of u_1, u_2 and u_3 .

One of the main areas of study in linear algebra is to study the set of all vectors which can be expressed as a linear combination of a given set of vectors, say $\{u_1, u_2, \dots, u_n\}$, where each $u_i \in \mathbb{R}^m$.

Definition 10.11. If $u_1, u_2, \dots, u_n \in \mathbb{R}^m$, then the set of all linear combinations of u_1, u_2, \dots, u_n is denoted by $\text{span}\{u_1, u_2, \dots, u_n\}$. Thus $\text{span}\{u_1, u_2, \dots, u_n\} = \{c_1u_1 + c_2u_2 + \dots + c_nu_n : c_i \in \mathbb{R}, i = 1, 2, \dots, n\}$.

The span of the null set is $\{0\}$.

Note that $\text{span}\{u_1, u_2, \dots, u_n\}$ contains every scalar multiple of u_i for $i=1,2,\dots,n$ (for example $cu_2 = 0u_1 + cu_2 + \dots + 0u_n$). In particular $0 = 0u_1 + 0u_2 + \dots + 0u_n$ so that $0 \in \text{span}\{u_1, u_2, \dots, u_n\}$.

Note: If $S = \{u_1, u_2, \dots, u_n\}$, then $\text{span}S = \text{span}\{u_1, u_2, \dots, u_n\}$.

Definition 10.12. Let $u_1, u_2, \dots, u_n, b \in \mathbb{R}^m$. An equation of the form $x_1u_1 + x_2u_2 + \dots + x_nu_n = b$... (1)

where $x_i \in \mathbb{R}, i = 1(1)n$ are unknowns is called a vector equation.

By a solution of this equation is meant a set of values of the unknowns x_i , which satisfy the equation. Thus $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$ is a solution of (1) if $c_1u_1 + c_2u_2 + \dots + c_nu_n = b$.

We shall now study the conditions for a vector to lie the span of a given set of vectors. In this regard, we have the following results.

Theorem 10.3. *If $b, u_i \in \mathbb{R}^m$, $i=1(1)n$, then the following statements are equivalent:*

- (1) $b \in \text{span}\{u_1, u_2, \dots, u_n\}$.
- (2) b is a linear combination of u_1, u_2, \dots, u_n .
- (3) The vector equation $x_1u_1 + x_2u_2 + \dots + x_nu_n = b$ has a solution.

Proof: (1) \Rightarrow (2) Let $b \in \text{span}\{u_1, u_2, \dots, u_n\}$. Then there exist weights c_1, c_2, \dots, c_n such that $c_1u_1 + c_2u_2 + \dots + c_nu_n = b$. Hence b is a linear combination of u_1, u_2, \dots, u_n .

(2) \Rightarrow (3) Let (2) hold. Then there exist c_1, c_2, \dots, c_n such that $c_1u_1 + c_2u_2 + \dots + c_nu_n = b$. Thus $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$ is a solution of the vector equation $x_1u_1 + x_2u_2 + \dots + x_nu_n = b$.

Hence (3) is true.

(3) \Rightarrow (1) Suppose (3) holds. Let $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$ be a solution of the vector equation. Then $c_1u_1 + c_2u_2 + \dots + c_nu_n = b$ so that $b \in \text{span}\{u_1, u_2, \dots, u_n\}$.

Hence (1) holds. Thus (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1), so that the statements are equivalent. \square

Theorem 10.4. *If $b, u_i \in \mathbb{R}^m$, $i=1,2,\dots,n$, then the following statements are equivalent:*

- (1) The vector equation $x_1u_1 + x_2u_2 + \dots + x_nu_n = b$ has a solution.
- (2) The linear system corresponding to the augmented matrix $[u_1 \ u_2 \ \dots \ u_n \ ; \ b]$ has a solution.

Proof: Let $u_1, u_2, \dots, u_n, b \in \mathbb{R}^m$, and $u_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix}$ for $i = 1(1)n$; $b_i =$

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Then $x_1u_1 + x_2u_2 + \dots + x_nu_n = b \quad \dots(1)$

$$\Leftrightarrow x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} x_1a_{11} + x_2a_{12} + \dots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \dots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \dots + x_na_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

\Leftrightarrow

$$\begin{aligned} x_1 a_{11} + x_2 a_{12} + \cdots + x_n a_{1n} &= b_1 \\ x_1 a_{21} + x_2 a_{22} + \cdots + x_n a_{2n} &= b_2 \\ &\vdots \\ x_1 a_{m1} + x_2 a_{m2} + \cdots + x_n a_{mn} &= b_m \end{aligned}$$

The augmented matrix of this system is

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right)$$

The columns of this matrix are u_1, u_2, \dots, u_n, b in order. Hence the vector equation (1) is equivalent to the linear equation (2).

Hence $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$ is a solution of (1) if and only if it is a solution of (2). \square

The above two theorems can be combined and restated as follows:

Theorem 10.5. *If $b, u_i \in \mathbb{R}^m, i=1,2,\dots,n$, then the following statements are equivalent:*

- (1) $b \in \text{span}\{u_1, u_2, \dots, u_n\}$.
- (2) The vector equation $x_1 u_1 + x_2 u_2 + \dots + x_n u_n = b$ has a solution.
- (3) The linear system corresponding to the augmented matrix $[u_1 \ u_2 \ \dots \ u_n \vdots \ b]$ has a solution.

We are now interested in finding the span of a given set of vectors. Thus, if $u_1, u_2, \dots, u_n \in \mathbb{R}^m$, the following questions arise:

- 1) Do u_1, u_2, \dots, u_n span \mathbb{R}^m ?
 - 2) If the answer to (1) is in the negative, then what is $\text{span}\{u_1, u_2, \dots, u_n\}$?
- To answer these questions we have the following theorems:

Theorem 10.6. *If $u_1, u_2, \dots, u_n \in \mathbb{R}^m$, the following statements are equivalent:*

1. u_1, u_2, \dots, u_n span \mathbb{R}^m
2. Every $b \in \mathbb{R}^m$ is a linear combination of u_1, u_2, \dots, u_n .
3. For every $b \in \mathbb{R}^m$, the vector equation $x_1 u_1 + x_2 u_2 + \dots + x_n u_n = b$ has a solution.
4. For every $b \in \mathbb{R}^m$, the linear system corresponding to the augmented matrix $[u_1 \ u_2 \ \dots \ u_n \vdots \ b]$ has a solution.
5. The matrix $[u_1 \ u_2 \ \dots \ u_n]$ has a pivot position in every row.

Proof: (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) are obvious from the definitions.

(4) \Rightarrow (5)

Let $A = [u_1 \ u_2 \ \dots \ u_n]$.

Suppose (4) holds. Thus the augmented column b does not have a pivot position. Hence every position lies in A. For any $b \in \mathbb{R}^m$, let $\left(\begin{array}{c} U \\ \vdots \\ c \end{array} \right)$ be the reduced echelon form of $\left(\begin{array}{c} A \\ \vdots \\ b \end{array} \right)$. Then U have a row of zeroes, for if it does then we can always choose a $b \in \mathbb{R}^m$ such that the matrix $\left(\begin{array}{c} U \\ \vdots \\ c \end{array} \right)$ has a row of the form $\left(\begin{array}{c} 0 \ 0 \ \dots \ 0 \ \vdots \ 1 \end{array} \right)$, i.e. augmented column has a pivot position, which is a contradiction. Thus every row of A has a pivot position.

(5) \Rightarrow (1)

Let $A = [u_1 \ u_2 \ \dots \ u_n]$.

Suppose (5) holds. Let, if possible, assume that (1) does not hold. Then, there exists $b \in \mathbb{R}^m$ such that $b \notin \text{span}\{u_1, u_2, \dots, u_n\}$

ie. b is not a linear combination of u_1, u_2, \dots, u_n .

Thus the vector equation $x_1u_1 + x_2u_2 + \dots + x_nu_n = b$ has no solution, so that the system with augmented matrix $[u_1 \ u_2 \ \dots \ u_n \ : \ b] (= (A \ : \ b))$ has no solution.

Hence b has a pivot position. Let $(U \ : \ c)$ be the reduced echelon form of $(A \ : \ b)$. Then U has row of the form $(0 \ 0 \ \dots \ 0 \ : \ 1)$, so that U has a row of zeroes. Hence A does not have a pivot position in the corresponding row. This contradicts (5). Hence our assumption is wrong so that (1) holds.

Thus

$$(5) \Rightarrow (1)$$

Thus all statements are equivalent. \square

Example 10.19. Does $b \in \text{span}\{u_1, u_2\}$, where $u_1 = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix}$,

$b = \begin{pmatrix} 0 \\ 2 \\ -4 \end{pmatrix}$. If yes, express b as in terms of u_1 and u_2 .

Using theorem (10.5).

$$b \in \text{span}\{u_1, u_2\}$$

\Leftrightarrow The vector equation $x_1u_1 + x_2u_2 = b$ has a solution.

\Leftrightarrow The linear system corresponding to the augmented matrix $[u_1 \ u_2 \ : \ b]$ has a solution.

$$\text{Now, } [u_1 \ u_2 \ : \ b] = \begin{pmatrix} 1 & 1 & \vdots & 0 \\ -2 & -4 & \vdots & 2 \\ -1 & 3 & \vdots & -4 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & \vdots & 0 \\ 0 & -2 & \vdots & 2 \\ 0 & 4 & \vdots & -4 \end{pmatrix} \text{ by } R_2 \rightarrow R_2 + 2R_1 \text{ and } R_3 \rightarrow R_3 + R_1$$

$$\sim \begin{pmatrix} 1 & 1 & \vdots & 0 \\ 0 & -2 & \vdots & 2 \\ 0 & 0 & \vdots & 0 \end{pmatrix} \text{ by } R_3 \rightarrow R_3 + 2R_2$$

$$\sim \begin{pmatrix} 1 & 1 & \vdots & 0 \\ 0 & 1 & \vdots & -1 \\ 0 & 0 & \vdots & 0 \end{pmatrix} \text{ by } R_2 \rightarrow -1/2R_2$$

The corresponding system of equations is $x_1 + x_2 = 0$, $x_2 = -1$. Hence the solution is $x_1 = 1$, $x_2 = -1$, so that $u_1 - u_2 = b$.

Hence $b = u_1 - u_2$.

Example 10.20. Is b a linear combination of u_1 and u_2 , where $u_1 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$,

$$u_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}. \text{ If yes, find the combination.}$$

Using Theorem 10.5.

b a linear combination of u_1 and u_2

\Leftrightarrow The vector equation $x_1u_1 + x_2u_2 = b$ has a solution.

\Leftrightarrow The linear system corresponding to the augmented matrix $[u_1 \ u_2 \ \vdots \ b]$ has a solution.

$$\text{Now, } [u_1 \ u_2 \ \vdots \ b] = \begin{pmatrix} 3 & 2 & \vdots & 4 \\ 2 & 1 & \vdots & 3 \\ 1 & 0 & \vdots & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & \vdots & 1 \\ 2 & 1 & \vdots & 3 \\ 3 & 2 & \vdots & 4 \end{pmatrix} \text{ by } R_1 \leftrightarrow R_3$$

$$\sim \begin{pmatrix} 1 & 0 & \vdots & 1 \\ 0 & 1 & \vdots & 1 \\ 0 & 2 & \vdots & 1 \end{pmatrix} \text{ by } R_2 \rightarrow R_2 - 2R_1 \text{ and } R_3 \rightarrow R_3 - 3R_1$$

$$\sim \begin{pmatrix} 1 & 0 & \vdots & 1 \\ 0 & 1 & \vdots & 1 \\ 0 & 0 & \vdots & -1 \end{pmatrix} \text{ by } R_3 \rightarrow R_3 - 2R_2$$

The corresponding linear system is

$$x_1 + 0x_2 = 1, 0x_1 + x_2 = 1, 0x_1 + 0x_2 = -1.$$

The last equation gives $0 = -1$, which is not true.

Hence the linear system does not have a solution, so that b is not a linear combination of u_1 and u_2 .

Example 10.21. Does the following set of vectors $\{u_1, u_2, u_3, u_4\}$ span \mathbb{R}^3 , where $u_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 6 \\ 3 \\ 0 \end{pmatrix}$, $u_3 = \begin{pmatrix} 4 \\ -1 \\ 2 \end{pmatrix}$ and $u_4 = \begin{pmatrix} 2 \\ -5 \\ 4 \end{pmatrix}$

Using Theorem 10.6, $\{u_1, u_2, u_3, u_4\}$ span \mathbb{R}^3 iff the matrix $(u_1 \ u_2 \ u_3 \ u_4)$ has a pivot position in every row. We shall reduce this matrix to echelon form.

$$\begin{aligned} A &= (u_1 \ u_2 \ u_3 \ u_4) \\ &= \begin{pmatrix} 1 & 6 & 4 & 2 \\ 2 & 3 & -1 & -5 \\ -1 & 0 & 2 & 4 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 6 & 4 & 2 \\ 0 & -9 & -9 & -9 \\ -1 & 6 & 6 & 6 \end{pmatrix} \text{ by applying } R_2 \rightarrow R_2 - 2R_1 \text{ and } R_3 \rightarrow R_3 + R_1 \end{aligned}$$

$$\sim \begin{pmatrix} \textcircled{1} & 6 & 4 & 2 \\ 0 & \textcircled{-9} & -9 & -9 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ by } R_3 \rightarrow R_3 + 2/3R_2 \\ = U \text{ (say)}$$

The pivots in U have been encircled. The pivot positions in A are the (1, 1)th and (2, 2)th positions. Thus the third row of A does not have any pivot position in every row. Thus the given vectors do not span \mathbb{R}^3 .

Example 10.22. Find the span of the given vectors in the above example.

Let $b \in \mathbb{R}^3$. Then $b \in \text{span}\{u_1, u_2, u_3, u_4\}$ iff $\begin{pmatrix} u_1 & u_2 & u_3 & u_4 & : & b \end{pmatrix}$ is the augmented matrix of a consistent linear system. We shall reduce this matrix to echelon form $\begin{pmatrix} A : b \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & : & b \end{pmatrix}$

Proceeding as in the above example

$$\sim \begin{pmatrix} 1 & 6 & 4 & 2 & & : & b_1 \\ 0 & -9 & -9 & -9 & & : & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & : & -1/3b_1 + 2/3b_2 + b_3 \end{pmatrix} \\ = \begin{pmatrix} U : b \end{pmatrix} \text{ (say)}$$

For the system to be consistent, we must have

$$-1/3b_1 + 2/3b_2 + b_3 = 0$$

$$\text{ie } b_1 = 2b_2 + 3b_3$$

$$\text{ie } \text{Span}\{u_1, u_2, u_3, u_4\} = \left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3 : b_1 = 2b_2 + 3b_3 \right\}.$$

Problem 10.15. Construct a 4×4 matrix A , with non-zero entries and a vector $b \in \mathbb{R}^4$ such that b is not in the set spanned by the columns of A .

Solution: The given problem is equivalent to finding a system of equations, with augmented matrix $\begin{pmatrix} A : b \end{pmatrix}$, which has no solution. If the echelon form of the matrix $\begin{pmatrix} A : b \end{pmatrix}$ is $\begin{pmatrix} U : c \end{pmatrix}$, then the column c must contain a pivot. There are many possible matrices $\begin{pmatrix} U : c \end{pmatrix}$. One such is

$$\begin{pmatrix} \textcircled{1} & 0 & 0 & 0 & : & 1 \\ 0 & \textcircled{1} & 1 & 0 & : & 1 \\ 0 & 0 & 0 & \textcircled{1} & : & 1 \\ 0 & 0 & 0 & 0 & : & 1 \end{pmatrix}$$

The pivots have been encircled. To get the matrix A with non zero entries we can apply row operations on $\begin{pmatrix} U : c \end{pmatrix}$.

$$\text{Apply } R_4 \rightarrow R_4 + R_2 + R_1 + R_3$$

$$\begin{aligned}
(U:c) &\sim \begin{pmatrix} 1 & 0 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 1 & 0 & \vdots & 1 \\ 0 & 0 & 0 & 1 & \vdots & 1 \\ 1 & 1 & 1 & 1 & \vdots & 4 \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 1 & 0 & \vdots & 1 \\ 1 & 1 & 1 & 1 & \vdots & 3 \\ 1 & 1 & 1 & 1 & \vdots & 4 \end{pmatrix} \text{ Applying } R_3 \rightarrow R_3 + R_2 + R_1 \\
&\sim \begin{pmatrix} 1 & 0 & 0 & 0 & \vdots & 1 \\ 2 & 3 & 3 & 2 & \vdots & 8 \\ 1 & 1 & 1 & 1 & \vdots & 3 \\ 1 & 1 & 1 & 1 & \vdots & 4 \end{pmatrix} \text{ Applying } R_2 \rightarrow R_2 + R_3 + R_4 \\
&\sim \begin{pmatrix} 3 & 3 & 3 & 2 & \vdots & 9 \\ 2 & 3 & 3 & 2 & \vdots & 8 \\ 1 & 1 & 1 & 1 & \vdots & 3 \\ 1 & 1 & 1 & 1 & \vdots & 4 \end{pmatrix} \text{ Applying } R_1 \rightarrow R_1 + R_2 \\
\text{Thus } A &= \begin{pmatrix} 3 & 3 & 3 & 2 \\ 2 & 3 & 3 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \\
\text{and } b &= \begin{pmatrix} 9 \\ 8 \\ 3 \\ 4 \end{pmatrix}.
\end{aligned}$$

Problem 10.16. Let $u_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 4 \end{pmatrix}$, $u_2 = \begin{pmatrix} 2 \\ -1 \\ -5 \\ 2 \end{pmatrix}$, $u_3 = \begin{pmatrix} 1 \\ -1 \\ -4 \\ 0 \end{pmatrix}$. Find the value of h so that $b = \begin{pmatrix} 2 \\ 1 \\ 1 \\ h \end{pmatrix}$ lies in $\text{span}\{u_1, u_2, u_3\}$

Solution: $b \in \text{span}\{u_1, u_2, u_3\}$

\Leftrightarrow The linear system corresponding to the augmented matrix $(u_1 \ u_2 \ u_3 \ :b)$ has a solution.

$$\text{Now } \begin{pmatrix} u_1 & u_2 & u_3 & : & b \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & : & 2 \\ 1 & -1 & -1 & : & 1 \\ 2 & -5 & -4 & : & 1 \\ 4 & 2 & 0 & : & h \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & : & 2 \\ 0 & -3 & -2 & : & -1 \\ 0 & -9 & -6 & : & -3 \\ 0 & -6 & -4 & : & -8+h \end{pmatrix}$$

by applying $R_2 \rightarrow R_2 - R_1$, $R_3 \rightarrow R_3 - 2R_1$ and $R_4 \rightarrow R_4 - 4R_1$

$$\sim \begin{pmatrix} 1 & 2 & 1 & : & 2 \\ 0 & -3 & -2 & : & -1 \\ 0 & 0 & 0 & : & 0 \\ 0 & 0 & 0 & : & -6+h \end{pmatrix} \text{ by } R_3 \rightarrow R_3 - 3R_2 \text{ and } R_4 \rightarrow R_4 - 2R_2$$

The corresponding system is:

$$x_1 + 2x_2 + x_3 = 2, -3x_2 - 2x_3 = -1, 0 = 0, 0 = -6 + h$$

The last equation gives $h=6$. Thus the system has solution when $h=6$ i.e. $b \in \text{span}\{u_1, u_2, u_3\}$ when $h = 6$.

10.12 Exercise

- If $u = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, $v = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ then
 - Compute $u + v$, $2u - v$, $-2u + 1/2v$
 - Display the vectors computed in (i) using arrows on a x-y plane
- A boat is traveling east across a river at the rate of 6kms/hour while the river's current is flowing south at a rate of 2km/hour.
 - Find the resultant velocity of the boat.
 - If the speed of the boat is halved, then what is the resultant velocity of the boat?
- Using the figure write the following vectors as a linear combination of u and v
 - a, b, c, d
 - z, y, z
 - Is every vector in \mathbb{R}^2 a linear combination of u and v ?
 - Display the vectors $2u - 3v$, $-3u - 4u$
- Write the system of equations that is equivalent to the given vector equation

$$\begin{aligned} \text{(i)} \quad & x_1 \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} -1.5 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ -2.5 \end{pmatrix} \\ \text{(ii)} \quad & x_1 \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \text{(iii)} \quad & x_1 \begin{pmatrix} -1 \\ 2 \\ 0 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ -1 \\ 4 \\ -6 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 1 \\ 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} \end{aligned}$$

5. Write the vector equation corresponding to the given system of linear equations

$$\text{(i)} \quad 2x + 3y - 4z = -1$$

$$x + 11z = 10$$

$$14y + 8z = -23$$

$$\text{(ii)} \quad -x + y - 3z = 12$$

$$2y + 5z = 6$$

$$-3x + 4y = -1$$

$$3y - 4z = -1$$

$$\text{(iii)} \quad 2y + 3z - 4w = 0$$

$$3x - 2z = 11$$

$$2y - 13z + 14w = 18$$

6. Is b a linear combination of the given vectors? If yes, express b as a linear combination of the vectors.

$$1. \quad u_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

$$2. \quad u_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, u_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, b = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

$$3. \quad u_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, u_3 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, u_4 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$4. \quad u_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

7. Is b a linear combination of the given vectors? If yes, express b as a linear combination of the vectors

$$\text{(i)} \quad u_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$$

$$\text{(ii)} \quad u_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ -7 \\ -8 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

$$\begin{aligned} \text{(iii)} \quad u_1 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, u_3 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} \\ \text{(iv)} \quad u_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, u_4 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, b = \\ &\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \end{aligned}$$

8. Is b a linear combination of the given vectors? If yes, express b as a linear combination of the vectors

$$\begin{aligned} \text{(i)} \quad u_1 &= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\ \text{(ii)} \quad u_1 &= \begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ 4 \\ 3 \\ -1 \end{pmatrix}, b = \begin{pmatrix} 5 \\ 0 \\ 1 \\ 3 \end{pmatrix} \\ \text{(iii)} \quad u_1 &= \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \\ u_4 &= \begin{pmatrix} 6 \\ 1 \\ 4 \\ 8 \end{pmatrix}, u_4 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 10 \\ 1 \\ 6 \\ 14 \end{pmatrix} \\ \text{(iv)} \quad u_1 &= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 2 \end{pmatrix}, u_2 = \begin{pmatrix} 3 \\ 5 \\ 4 \\ 3 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 6 \\ 4 \\ 5 \end{pmatrix} \end{aligned}$$

9. List 3 vectors in $\text{span}\{u_1, u_2\}$ where $u_1 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

10. List 4 vectors in $\text{span}\{u_1, u_2, u_3\}$ where $u_1 = \begin{pmatrix} 5 \\ 3 \\ -1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} -4 \\ 0 \\ 1 \\ 3 \end{pmatrix},$

$$u_3 = \begin{pmatrix} -4 \\ -2 \\ 1 \\ 1 \end{pmatrix}$$

11. Given u_1, u_2, u_3 where $u_1 = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}, u_2 = \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}, u_3 = \begin{pmatrix} 10 \\ -11 \\ 3 \end{pmatrix}.$

Answer the following:

- (i) Does u_3 lie in $\text{span}\{u_1, u_2\}$
- (ii) Does u_1 lie in $\text{span}\{u_2, u_3\}$
- (iii) Does u_2 lie in $\text{span}\{u_1, u_3\}$

12. Give the geometrical of the following vectors

(i) $\text{span}\{u_1\}$, where $u_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

(ii) $\text{Span}\{u_1, u_2\}$, where $u_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$

$\text{Span}\{u_1, u_2\}$, where $u_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

(iii) $\{u_1, u_2, u_3\}$, where $u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$, $u_3 = \begin{pmatrix} 3 \\ -4 \\ 5 \end{pmatrix}$

13. Let $A = (u_1 \ u_2 \ u_3)$ where $u_1 = \begin{pmatrix} 3 \\ 8 \\ -3 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}$, $u_3 =$

$\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$ and let $b = \begin{pmatrix} 1 \\ 5 \\ -3 \end{pmatrix}$, $W = \text{span}\{u_1, u_2, u_3\}$

- (i) Does b belong to $\{u_1, u_2, u_3\}$? How many vectors are there in $\{u_1, u_2, u_3\}$?
- (ii) Does b belong to W ? How many vectors are there in W ?
- (iii) Show that u_2 belongs to W
- (iv) Does u_2 belong to $\text{span}\{u_1, u_2\}$?

14. Determine if b lies in the span of the columns of A , where

(i) $A = \begin{pmatrix} 1 & -1 & 3 \\ -2 & 1 & -10 \\ 3 & -3 & 10 \end{pmatrix}$, $b \in \begin{pmatrix} 2 \\ -5 \\ 9 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 1 & 0 & 0 & -2 \\ 2 & 0 & 1 & 3 \\ 3 & -1 & 2 & 4 \end{pmatrix}$, $b \in \begin{pmatrix} 4 \\ -5 \\ 6 \end{pmatrix}$

(iii) $A = \begin{pmatrix} 3 & 2 \\ -2 & -2 \\ 4 & 9 \end{pmatrix}$, $b \in \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

15. Let $u_1 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. If $u_1 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, $b = \begin{pmatrix} h \\ k \end{pmatrix}$, for what values of h and k does $b \in \text{span}\{u_1, u_2\}$?

16. Construct a 3×3 matrix A , with non-zero entries and a vector $b \in \mathbb{R}^3$, such that b is not in the set spanned by the columns of A .

17. Construct a 4×3 matrix A , with non zero entries, and a vector $b \in \mathbb{R}^3$, such that b is not in the set spanned by the columns of A .

18. Which of the following sets of vectors span \mathbb{R}^n ?

(i) $\{u_1, u_2\}$ where $u_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

(ii) $\{u_1, u_2, u_3\}$ where $u_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

(iii) $\{u_1, u_2, u_3\}$ where $u_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

(iv) $\{u_1, u_2\}$ where $u_1 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

19. Which of the following sets of vectors span \mathbb{R}^3 ?

(i) $\{u_1, u_2\}$ where $u_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

(ii) $\{u_1, u_2, u_3\}$ where $u_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

(iii) $\{u_1, u_2, u_3, u_4\}$ where $u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$

$$u_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

20. Which of the following set of vectors span \mathbb{R}^4 ?

(i) $\{u_1, u_2, u_3, u_4\}$ where $u_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$

$$u_4 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}$$

(ii) $\{u_1, u_2, u_3, u_4, u_5\}$ where $u_1 = \begin{pmatrix} 6 \\ 4 \\ -2 \\ 4 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$

$$u_3 = \begin{pmatrix} 3 \\ 2 \\ -1 \\ 2 \end{pmatrix}, u_4 = \begin{pmatrix} 5 \\ 6 \\ -3 \\ 2 \end{pmatrix}, u_5 = \begin{pmatrix} 0 \\ 4 \\ -2 \\ -1 \end{pmatrix}$$

$$\begin{aligned} \text{(iii) } \{u_1, u_2, u_3\} \text{ where } u_1 &= \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 0 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \text{(iv) } \{u_1, u_2, u_3, u_4\} \text{ where } u_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \\ u_4 &= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

21. Do the points in the plane corresponding to $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} -3 \\ 1 \end{pmatrix}$ lie on a line through the origin?
22. Are the following statements true or false. If false, correct them.
- (i) $1/3u_1$ is a linear combination of u_1 and u_2 .

10.13 Solutions of Linear Systems

We shall now discuss the existence and uniqueness of the solutions of a system of linear equations. To do this we consider the following example.

Example 10.23. Solve the linear system

$$\begin{aligned} x_1 + x_2 + 2x_3 - 5x_4 &= 3 \\ 2x_1 + 5x_2 - x_3 - 9x_4 &= -3 \\ 2x_1 + x_2 - x_3 + 3x_4 &= -11 \\ x_1 - 3x_2 + 2x_3 + 7x_4 &= -5 \end{aligned}$$

Solution: To solve this system, we shall transform the augmented matrix of the system to reduced echelon form.

Step 1 The augmented matrix is

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 1 & 2 & -5 & 3 \\ 2 & 5 & -1 & -9 & -3 \\ 2 & 1 & -1 & 3 & -11 \\ 1 & -3 & 2 & 7 & -5 \end{array} \right)$$

Step 2 Apply $R_2 \rightarrow R_2 - 2R_1, R_3 \rightarrow R_3 - 2R_1, R_4 \rightarrow R_4 - R_1$ to $[A|b]$. Then

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 1 & 2 & -5 & 3 \\ 0 & 3 & -5 & 1 & -9 \\ 0 & -1 & -5 & 13 & -17 \\ 0 & -4 & 0 & 12 & -8 \end{array} \right) = B(\text{say})$$

Step 3 To B apply $R_2 \leftrightarrow R_3$. This is done to make the pivot element -1 as this will simplify the calculations

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 1 & 2 & -5 & 3 \\ 0 & -1 & -5 & 13 & -17 \\ 0 & 3 & -5 & 1 & -9 \\ 0 & -4 & 0 & 12 & -8 \end{array} \right)$$

Apply $R_3 \rightarrow R_3 + 3R_2, R_4 \rightarrow R_4 - 4R_2$

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 1 & 2 & -5 & 3 \\ 0 & -1 & -5 & 13 & -17 \\ 0 & 0 & -20 & 40 & -60 \\ 0 & 0 & 20 & -40 & 60 \end{array} \right) = C(\text{say})$$

Step 4 On C apply $R_4 \rightarrow R_4 + R_3$

$$(A|b) \sim \left(\begin{array}{cccc|c} \mathbf{1} & 1 & 2 & -5 & 3 \\ 0 & \mathbf{-1} & -5 & 13 & -17 \\ 0 & 0 & \mathbf{-20} & 40 & -60 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = D(\text{say})$$

The augmented matrix has been reduced to row-echelon form. We have bold-faced the pivot elements.

Step 5 Now we shall reduce the augmented matrix to reduced echelon form. We will start with the right-most pivot element. Scale row-3 of D by applying $R_3 \rightarrow \frac{-1}{20}R_3$

$$(A|b) \sim \left(\begin{array}{cccc|c} \mathbf{1} & 1 & 2 & -5 & 3 \\ 0 & \mathbf{-1} & -5 & 13 & -17 \\ 0 & 0 & \mathbf{1} & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Apply $R_2 \rightarrow R_2 + 5R_3, R_1 \rightarrow R_1 - 2R_3$. This is done to make the elements above the pivot 3rd, as zeros.

$$(A|b) \sim \left(\begin{array}{cccc|c} \mathbf{1} & 1 & 0 & -1 & -3 \\ 0 & \mathbf{-1} & 0 & 3 & -2 \\ 0 & 0 & \mathbf{1} & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = E(\text{say})$$

Scale row-2 of E (to make the 2nd pivot 1), by applying $R_2 \rightarrow (-1)R_2$

$$(A|b) \sim \left(\begin{array}{cccc|c} \mathbf{1} & 1 & 0 & -1 & -3 \\ 0 & \mathbf{1} & 0 & -3 & 2 \\ 0 & 0 & \mathbf{1} & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Apply $R_1 \rightarrow R_1 - R_2$ (This is done to make the elements above the pivot-2nd, as zeros)

$$(A|b) \sim \left(\begin{array}{cccc|c} \mathbf{1} & 0 & 0 & 2 & -5 \\ 0 & \mathbf{1} & 0 & -3 & 2 \\ 0 & 0 & \mathbf{1} & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = F(\text{say})$$

The augmented matrix has been transformed to reduced echelon form.

Step 6 The linear system represented by F is

$$\begin{aligned}x_1 + 2x_4 &= -5 \\x_2 - 3x_4 &= 2 \\x_3 - 2x_4 &= 3 \\0 &= 0\end{aligned}$$

Solving each equation for the unknown that corresponds to the leading entry in each row of F , we obtain

$$\begin{aligned}x_1 &= -5 - 2x_4 \\x_2 &= 2 + 3x_4 \\x_3 &= 3 + 2x_4\end{aligned}$$

The value of x_4 is not determined, and so it can take any arbitrary value. If we let $x_4 = k$, any real number, then a solution to the given linear system is

$$\begin{aligned}x_1 &= -5 - 2k \\x_2 &= 2 + 3k \\x_3 &= 3 + 2k \\x_4 &= k\end{aligned}$$

Since k can be assigned any real number, the given system has infinitely many solutions.

The variables which can take arbitrary values are called free variables and the others are called the basic variables. In the above example, x_4 is a free variable whereas x_1, x_2, x_3 are the basic variables.

Definition 10.13. *Among the variables in a system, the ones corresponding to the columns containing leading 1's (in the reduced echelon form of the augmented matrix) are called basic variables, and the ones corresponding to the other columns, if there are any, are called the free variables.*

Clearly the sum of the number of basic variables and the number of free variables is equal to the total number of unknowns: the number of columns. In general, a consistent system has infinitely many solutions if it has at least one free variable, and has a unique solution if it has no free variable. In fact, if a consistent system has a free variable (which always happens when the number of equations is less than that of unknowns), then by assigning arbitrary value to the free variable, one always obtains infinitely many solutions.

The basic variables can always be expressed in terms of the free variables, if any. When this is done, the solution, thus written, is called the general solution as it gives an explicit description of all solutions.

In the above example, the general solution is

$$\begin{aligned}x_1 &= -5 - 2x_4 \\x_2 &= 2 + 3x_4 \\x_3 &= 3 + 2x_4\end{aligned}$$

x_4 is free. Each different choice of x_4 determines a (different) solution of the system, and every solution of the system is determined by a choice of x_4 .

Given a system of linear equations, is it possible to know the number of free variables without actually solving the system?

The answer is “yes”.

In fact, if the row echelon form of the augmented matrix is U , then the variables corresponding to the columns of the leading entries are the basic variables and the remaining variables are of the free variables. For instance, if the row echelon form of the augmented matrix of a system is

$$\left(\begin{array}{cccccccc|c} \blacksquare & * & 0 & * & * & * & * & * \\ 0 & \blacksquare & 0 & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

then the basic variables are x_1, x_2, x_4 and x_6 and the free variables are x_3, x_5 and x_7 .

Example 10.24. Find the general solution of

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ 2x_1 - x_2 + x_3 &= 8 \\ 3x_1 - x_3 &= 3 \end{aligned}$$

Solution: *Step 1* The augmented matrix is

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 2 & 3 & 9 \\ 2 & -1 & 1 & 8 \\ 3 & 0 & -1 & 3 \end{array} \right)$$

Step 2 The augmented matrix is row equivalent to

$$(A|b) = \left(\begin{array}{cccc|c} \textcircled{1} & 0 & 0 & 2 \\ 0 & \textcircled{1} & 0 & -1 \\ 0 & 0 & \textcircled{1} & 3 \end{array} \right)$$

which is in the reduced echelon form (verify).

The leading 1's are encircled.

Step 3 The corresponding system of equations is

$$\begin{aligned} x_1 &= 2 \\ x_2 &= -1 \\ x_3 &= 3 \end{aligned}$$

Thus there are no free variables and the system has a unique solution given by $(x_1, x_2, x_3) = (2, -1, 3)$.

Example 10.25. Solve the linear system:

$$\begin{aligned} x_1 + 2x_2 - 3x_4 + x_5 &= 2 \\ x_1 + 2x_2 + x_3 - 3x_4 + x_5 + 2x_6 &= 3 \\ x_1 + 2x_2 - 3x_4 + 2x_5 + x_6 &= 4 \\ 3x_1 + 6x_2 + x_3 - 9x_4 + 4x_5 + 3x_6 &= 9 \end{aligned}$$

Solution: *Step 1* The augmented matrix is

$$(A|b) = \left(\begin{array}{cccccc|c} 1 & 2 & 0 & -3 & 1 & 0 & 2 \\ 1 & 2 & 1 & -3 & 1 & 2 & 3 \\ 1 & 2 & 0 & -3 & 2 & 1 & 4 \\ 3 & 6 & 1 & -9 & 4 & 3 & 9 \end{array} \right)$$

Step 2 The reduced row echelon form of the augmented matrix is

$$(A|b) \sim \left(\begin{array}{cccccc|c} \mathbf{1} & 2 & 0 & -3 & 0 & -1 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

The leading 1's are bold-faced. The variables corresponding to the pivot columns namely are x_1 , x_3 , x_5 are the basic variables and the remaining variables i.e. x_2 , x_4 and x_6 are the free variables.

Step 3 The corresponding system is

$$\begin{aligned} x_1 + 2x_2 - 3x_4 - x_6 &= 0 \\ x_3 + 2x_6 &= 1 \\ x_5 + x_6 &= 2 \end{aligned}$$

The solution is

$$x_1 = -2x_2 + 3x_4 + x_6$$

x_2 is free.

$$x_3 = 1 - 2x_6$$

x_4 is free.

$$x_5 = 2 - x_6$$

x_6 is free.

This system has three free variables.

10.14 Parametric Description of Solution Sets

Whenever a system is consistent and has free variables, the solution set can be described in terms of parameter(s). For instance, in above Example 10.23, the solution is

$$\begin{aligned} x_1 &= -5 - 2x_4 \\ x_2 &= 2 + 3x_4 \\ x_3 &= 3 + 2x_4 \end{aligned}$$

x_4 is free.

This is the parametric description of the solution, where x_4 is the parameter. The solution is in terms of one parameter.

In Example 10.25, the free variables x_2 , x_4 and x_6 act as the parameters. Here the solution set is in terms of three parameters.

The solution set of a linear system can have many parametric representations. For instance, in Example 10.23, the given linear system is reduced to the equivalent system

$$x_1 + 2x_4 = -5 \cdots (1)$$

$$x_2 - 3x_4 = 2 \cdots (2)$$

$$x_3 - 2x_4 = 3 \cdots (3)$$

$$(1) + (3) \Rightarrow x_1 + x_3 = -2$$

$$2 \times (2) - 3 \times (3) \Rightarrow 2x_2 - 3x_3 = -5$$

$$\Rightarrow x_2 - \frac{3}{2}x_3 = -\frac{5}{2}$$

Thus the equations (1), (2), (3) are equivalent to:

$$x_1 + x_3 = -2$$

$$x_2 - \frac{3}{2}x_3 = -\frac{5}{2}$$

$$x_3 - 2x_4 = 3$$

Solving in terms of x_3 , we get

$$x_1 = -2 - x_3$$

$$x_2 = -\frac{5}{2} + \frac{3}{2}x_3$$

$$x_4 = -\frac{3}{2} + \frac{1}{2}x_3$$

and x_3 is a parameter. If x_3 is assigned an arbitrary real number k , then the solution we get is $(x_1, x_2, x_3, x_4) = (-2 - k, -\frac{5}{2} + \frac{3}{2}k, k, -\frac{3}{2} + \frac{1}{2}k)$ where k is a parameter.

Similarly, we may choose x_1 (or x_2) as a parameter.

However, to be consistent, we make the convention of always using the free variables as the parameters for describing a solution set. Whenever, a system is inconsistent, the solution set is empty. In this case, the solution set has no parametric representation.

Example 10.26. Consider problem 10.4

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 3 & -2 & 0 & 3 \\ 2 & 6 & -2 & 4 & 18 \\ 0 & 1 & 1 & 3 & 10 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 3 & -2 & 0 & 3 \\ 0 & 1 & 1 & 3 & 10 \\ 0 & 0 & 2 & 4 & 12 \end{array} \right)$$

as done before.

This is the triangular form. We shall apply row operations to reduce $[A|b]$ to reduced echelon form.

Applying $R_3 \rightarrow \frac{1}{2}R_3$

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 3 & -2 & 0 & 3 \\ 0 & 1 & 1 & 3 & 10 \\ 0 & 0 & 1 & 2 & 6 \end{array} \right)$$

Applying $R_1 \rightarrow R_1 - 2R_3, R_2 \rightarrow R_2 - R_1$

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 3 & 0 & 4 & 15 \\ 0 & 1 & 0 & 1 & 4 \\ 0 & 0 & 1 & 2 & 6 \end{array} \right)$$

Applying $R_1 \rightarrow R_1 - 3R_2$

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 1 & 4 \\ 0 & 0 & 1 & 2 & 6 \end{array} \right)$$

This is the reduced row-echelon form. The corresponding system of equation is

$$\begin{aligned} x_1 + x_4 &= 3 \\ x_2 + x_4 &= 4 \\ x_3 + 2x_4 &= 6 \end{aligned}$$

Thus, by back substitution,

$$\begin{aligned} x_1 &= 3 - x_4 \\ x_2 &= 4 - x_4 \\ x_3 &= 6 - 2x_4 \end{aligned}$$

Since, there is no condition on x_4 , the variables x_1, x_2, x_3 can be uniquely obtained if x_4 is assigned any real value. Thus, the solution is $(x_1, x_2, x_3, x_4) = (3 - k, 4 - k, 6 - 2k, k)$ where k is any real number.

10.15 Existence and Uniqueness of Solutions

The basic question is: Given a system of linear equations, does it have a solution and if it does, is it unique? Before answering this we shall take up a few examples.

Example 10.27. Determine the existence and uniqueness of the solutions of the system

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 5 \\ x_1 + 3x_2 + 5x_3 + 7x_4 &= 11 \\ x_1 - x_3 - 2x_4 &= -6 \end{aligned}$$

Also, obtain the solution, if it exists.

Solution: *Step 1* The augmented matrix of the system is

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 7 & 11 \\ 1 & 0 & -1 & -2 & -6 \end{array} \right)$$

Step 2 The echelon form of the matrix is

$$C = \left(\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

Step 3 The equivalent system is

$$\begin{aligned} x_1 - x_3 - 2x_4 &= 0 \\ x_2 + 2x_3 + 3x_4 &= 0 \\ 0 &= 1 \end{aligned}$$

The last equation is inconsistent, so that the given system is inconsistent. So the solution set is empty.

Remark 10.5. *If an equivalent system gives rise to an inconsistent equation, then the given system is inconsistent.*

Example 10.28. *Determine the existence and uniqueness of the solutions of the system*

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ 2x_1 - x_2 + x_3 &= 8 \\ 3x_1 - x_3 &= 3 \end{aligned}$$

Also, obtain the solution, if it exists.

Solution:

Step 1 The augmented matrix of the system is

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 2 & -1 & 1 & 8 \\ 3 & 0 & -1 & 3 \end{array} \right)$$

Step 2 The row echelon form of the matrix is

$$(B) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right)$$

Verify it yourself.

Step 3: The equivalent system is

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ x_2 + x_3 &= 2 \\ x_3 &= 3 \end{aligned}$$

Using back-substitution, the solution is $x_1 = 2, x_2 = -1, x_3 = 3$ which is unique.

Remark 10.6. Observe that in the matrix B there is no row of the form $[0 \ 0 \ 0 \ | \ b]$, with $b \neq 0$, so that the system is consistent. Moreover, each column is a pivot column. This gives that the solution is unique.

Example 10.29. Determine the existence and uniqueness of the solutions of the system

$$\begin{aligned}x_2 + 2x_3 + 4x_4 &= -1 \\x_1 + 3x_2 - x_3 &= 5 \\2x_1 + 4x_3 + x_4 &= 3\end{aligned}$$

Also, obtain the solution, if it exists.

Solution:

Step 1 The augmented matrix of the system is

$$(A|b) = \left(\begin{array}{cccc|c} 0 & 1 & 2 & 4 & -1 \\ 1 & 3 & -1 & 0 & 5 \\ 2 & 0 & 4 & 1 & 3 \end{array} \right)$$

Step 2 The row echelon form of the matrix $(A|b)$ is

$$E = \left(\begin{array}{cccc|c} 1 & 3 & -1 & 0 & 5 \\ 0 & 1 & 2 & 4 & -1 \\ 0 & 0 & 18 & 25 & -13 \end{array} \right)$$

Step 3 The matrix E does not have any row of the form $[0 \ 0 \ 0 \ 0 \ | \ b]$ with $b \neq 0$ so that there will not be an inconsistent equation. Thus, the system is consistent. Moreover, there are 4 variables, since, there are 3 leading entries. Therefore, there are 3 basic variables and one free variable. Hence, the system has infinitely many solutions. The equivalent system is

$$\begin{aligned}x_1 + 3x_2 - x_3 &= 5 \\x_2 + 2x_3 + 4x_4 &= -1 \\18x_3 + 25x_4 &= -13\end{aligned}$$

Solving in terms of x_4 (free variable).

$$\begin{aligned}x_1 &= \frac{1}{18}(53 + 41x_4) \\x_2 &= \frac{1}{9}(4 - 11x_4) \\x_3 &= \frac{1}{18}(-13 - 25x_4)\end{aligned}$$

Thus, the solution is $(x_1, x_2, x_3, x_4) = (\frac{1}{18}(53+41k), \frac{1}{9}(4-11k), -\frac{1}{18}(13+25k), k)$ where k is any real number.

Summarizing the observations from the above examples we note that the echelon form $[C|d]$ of the augmented matrix $[A|b]$ enables us to draw conclusions about the existence and uniqueness of the solutions of the given system.

- If the echelon form contains a row of the form $[0\ 0\ \dots\ 0\ | e]$, with e non-zero, then, the system is inconsistent.
- If the system is consistent and the number of leading entries is equal to the number of variables (i.e. each column of the matrix C is a pivot column), then, each variable is a basic variable and there are no free variables. In this case the system has a unique solution.
- If the system is consistent and the number of leading entries is less than the number of variables, then, there is atleast one free variable. In this case, the system has infinitely many solutions.

These remarks justify the following theorem.

Theorem 10.7. *A linear system of equations is consistent if and only if the right-most column of the augmented matrix is not a pivot column, that is, if and only if an echelon form of the augmented matrix has no row of the form $[0\ 0\ 0\ \dots\ 0\ | b]$ with b non-zero.*

If a linear system is consistent, then, the solution set contains either

1. *A unique solution, when there are no free variables, or*
2. *Infinitely many solutions, when there is atleast one free variable.*

Steps to find the existence and uniqueness of the solution of a linear system:

- (i) Write the augmented matrix $[A|b]$.
- (ii) Reduce it to echelon form say $[C|d]$.
- (iii) If the augmented column d is a pivot column, then, the system is inconsistent, else consistent.
- (iv) If the system is consistent and every column is a pivot column then the solution is unique. If some column is not a pivot column, then, the system has infinitely many solutions.

Steps to find the solution of a linear system:

- (i) Write the augmented matrix $[A|b]$.
- (ii) Reduce it to echelon form say $[C|d]$.
- (iii) Decide whether the system is consistent, if it is, go to the next step else stop.
- (iv) Further reduce the echelon form to obtain the reduced echelon form.
- (v) Write the system of equations corresponding to the matrix obtained in the above step.
- (vi) Obtain the values of the basic variables in terms of the free variables (if any).

10.16 Homogenous System

In a linear system with augmented matrix $[A|b]$, if $b=0$, the system is said to be homogenous, and non-homogenous otherwise. A homogenous system always has $X = 0$ as a solution. Thus, such a system is always consistent. The solution $X = 0$ is called a trivial solution (or zero solution).

We would like to know whether a homogenous system has a solution other than the trivial solution, i.e., one in which atleast one of the variables is non-zero. Such a solution is called a non-trivial (or non-zero) solution. The following theorem tells us when a homogenous system has non-trivial solutions.

Theorem 10.8. *The homogenous system with augmented matrix $[A|O]$ has non-trivial solution if and only if the number of pivot columns of the coefficient matrix is less than the number of variables. Using theorem 10.7 we conclude that*

1. *If the number of equations is less than the number of variables, then, the system has non-trivial solution.*
2. *If every column of the coefficient matrix is a pivot column, then, the system has a unique solution namely the trivial solution.*

Example 10.30. *Solve the homogenous system*

$$2x + y - 2z = 0$$

$$-x + 3y - z = 0$$

$$x + 2y + 3z = 0$$

Solution: *Step 1* The augmented matrix of the system is

$$(A|b) = \left(\begin{array}{ccc|c} 2 & 1 & -2 & 0 \\ -1 & 3 & -1 & 0 \\ 1 & 2 & 3 & 0 \end{array} \right)$$

Step 2 The reduced row echelon form of the matrix is

$$(A|b) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right)$$

Step 3 The equivalent system is

$$x = 0$$

$$y = 0$$

$$z = 0$$

Thus, the given system has a unique solution $x = y = z = 0$.

Example 10.31. *Solve the homogenous system*

$$w + x + y + z = 0$$

$$w + x = 0$$

$$x + 2y + z = 0$$

Solution: *Step 1* The augmented matrix of the system is

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \end{array} \right)$$

Step 2 The reduced row echelon form of the matrix is

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right)$$

Step 3 The equivalent system is

$$\begin{aligned} w + z &= 0 \\ x - z &= 0 \\ y + z &= 0 \end{aligned}$$

Thus, the given system has a solution

$$\begin{aligned} w &= -z \\ x &= z \\ y &= -z \end{aligned}$$

z is free.

Problem 10.17. *Solve the system*

$$\begin{aligned} w + x + y + z &= 4 \\ -w + x + y + z &= 2 \\ -w + x - y + z &= 0 \\ -w + x - y - z &= -2 \end{aligned}$$

Solution: *Step 1* The augmented matrix of the system is

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ -1 & 1 & 1 & 1 & 2 \\ -1 & 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & -1 & -2 \end{array} \right)$$

Step 2 Transforming the above matrix to row echelon form by applying the elementary row transformations we get

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right)$$

Step 3 Transforming the above matrix to reduced row echelon form, we have

$$(A|b) \sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right)$$

The system corresponding to the above matrix is

$$\begin{aligned} w &= 1 \\ x &= 1 \\ y &= 1 \\ z &= 1 \end{aligned}$$

Thus, the system is consistent and has a unique solution.
It is $w = x = y = z = 1$.

The following example shows that the consistency of a system depends on the augmented column and not on the coefficient matrix.

Example 10.32. Solve the linear systems $AX=b$ and $AX=c$ where

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 2 & 4 & 8 & 6 \\ 1 & 2 & 4 & 5 \\ 3 & 5 & 11 & 8 \\ 4 & 6 & 14 & 0 \end{pmatrix}$$

$$b = \begin{pmatrix} -1 \\ 10 \\ 3 \\ 15 \\ 28 \end{pmatrix} \quad c = \begin{pmatrix} 2 \\ 6 \\ 7 \\ 8 \\ -10 \end{pmatrix}$$

Solution: In this case, we have to solve two linear systems but the coefficient matrix of both the systems is the same. Only the right hand sides are different. Thus, we will consider the matrix $[A | b | c]$. So that, we will get the reduction of augmented matrices of both the systems in one go.

Step 1

$$[A|b|c] = \left(\begin{array}{cccc|c|c} 0 & 0 & 0 & 1 & -1 & 2 \\ 2 & 4 & 8 & 6 & 10 & 6 \\ 1 & 2 & 4 & 5 & 3 & 7 \\ 3 & 5 & 11 & 8 & 15 & 8 \\ 4 & 6 & 14 & 0 & 28 & -10 \end{array} \right)$$

Step 2 Apply $R_1 \leftrightarrow R_3$ so that the (1,1)th element becomes non-zero.

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} \textcircled{1} & 2 & 4 & 5 & 3 & 7 \\ 2 & 4 & 8 & 6 & 10 & 6 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 3 & 5 & 11 & 8 & 15 & 8 \\ 4 & 6 & 14 & 0 & 28 & -10 \end{array} \right)$$

In order to make the elements of the 1st column below the pivot element(encircled) zero, apply $R_2 \rightarrow R_2 - 2R_1, R_4 \rightarrow R_4 - 3R_1, R_5 \rightarrow R_5 - 4R_1$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} 1 & 2 & 4 & 5 & 3 & 7 \\ 0 & 0 & 0 & -4 & 4 & -8 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & -1 & -1 & -7 & 6 & -13 \\ 0 & -2 & -2 & -20 & 16 & -38 \end{array} \right)$$

Step 3 In order to bring the pivot to the (2, 2)th position, apply $R_2 \leftrightarrow R_4$

$$\therefore [A|b|c] \sim \left(\begin{array}{cccc|c|c} 1 & 2 & 4 & 5 & 3 & 7 \\ 0 & \textcircled{1} & -1 & -7 & 0 & -13 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -4 & 4 & -8 \\ 0 & -2 & -2 & -20 & 16 & -38 \end{array} \right)$$

To make the elements below the 2nd column below the pivot element(encircled) zero apply $R_5 \rightarrow R_5 - 2R_2$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} 1 & 2 & 4 & 5 & 3 & 7 \\ 0 & \textcircled{1} & -1 & -7 & 0 & -13 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -4 & 4 & -8 \\ 0 & 0 & 0 & -6 & 16 & -12 \end{array} \right)$$

Step 4 Apply $R_4 \rightarrow -\frac{1}{4}R_4, R_5 \rightarrow \frac{1}{2}R_5$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} 1 & 2 & 4 & 5 & 3 & 7 \\ 0 & -1 & -1 & -7 & 0 & -13 \\ 0 & 0 & 0 & \textcircled{1} & -1 & 2 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -3 & 8 & -6 \end{array} \right)$$

Apply $R_4 \rightarrow R_4 - R_3, R_5 \rightarrow R_5 + 3R_3$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} 1 & 2 & 4 & 5 & 3 & 7 \\ 0 & -1 & -1 & -7 & 0 & -13 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \end{array} \right)$$

Step 5 To bring the row of zeros to the bottom-most position, apply $R_4 \leftrightarrow R_5$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} \textcircled{1} & 2 & 4 & 5 & 3 & 7 \\ 0 & \textcircled{-1} & -1 & -7 & 0 & -13 \\ 0 & 0 & 0 & \textcircled{1} & -1 & 2 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Step 6 We shall now transform the coefficient matrix to reduced echelon form.

Apply $R_2 \rightarrow R_2 + 7R_3$, $R_1 \rightarrow R_1 - 5R_3$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} \textcircled{1} & 2 & 4 & 0 & 8 & -3 \\ 0 & \textcircled{-1} & -1 & 0 & -7 & 1 \\ 0 & 0 & 0 & \textcircled{1} & -1 & 2 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Apply $R_1 \rightarrow R_1 + 2R_2$,

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} \textcircled{1} & 0 & 2 & 0 & -6 & -1 \\ 0 & \textcircled{-1} & -1 & 0 & -7 & 1 \\ 0 & 0 & 0 & \textcircled{1} & -1 & 2 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Apply $R_2 = (-1)R_2$

$$[A|b|c] \sim \left(\begin{array}{cccc|c|c} 1 & 0 & 2 & 0 & -6 & -1 \\ 0 & 1 & 1 & 0 & 7 & -1 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Step 7 Thus we get that,

$$[A|b] \sim \left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & -6 \\ 0 & 1 & 1 & 0 & 7 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = A_1(\text{say})$$

$$[A|c] \sim \left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) = A_2(\text{say})$$

A_1 is the reduced echelon form of $[A|b]$. Since, A_1 has a row of the form $[0 \ 0 \ 0 \ 0 \ | \ 5]$, hence, the system $AX = b$ is inconsistent. The reduced echelon form of $[A|c]$ is A_2 and the corresponding system is

$$\begin{aligned} x_1 + 2x_3 &= -1 \\ x_2 + x_3 &= -1 \\ x_4 &= 2 \end{aligned}$$

Hence, the solution is

$$\begin{aligned} x_1 &= -1 - 2x_3 \\ x_2 &= -1 - x_3 \\ x_3 &\text{ is free} \\ x_4 &= 2 \end{aligned}$$

Thus, $AX=c$ is consistent having infinitely many solutions.

Problem 10.18. Suppose that each matrix represents the augmented matrix for a system of linear equations. In each case, determine if the system is consistent. In case it is consistent, determine if the solution is unique. Here ■ represents a pivot position and * any entry, including zero.

(i)

$$A = \left(\begin{array}{ccc|c} \blacksquare & 0 & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 \end{array} \right)$$

(ii)

$$B = \left(\begin{array}{ccc|c} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \end{array} \right)$$

(iii)

$$C = \left(\begin{array}{ccc|c} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & 0 \end{array} \right)$$

(iv)

$$D = \left(\begin{array}{ccccc|c} 0 & \blacksquare & 0 & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & * \end{array} \right)$$

Solution:

(i) Since, the last column (augmented column) is not a pivot column, therefore, the system is consistent. Number of variables = 3, Number of pivots = 2. Thus, there is one free variable and the system has infinitely many solutions.

(ii) Since, the augmented column is a pivot column, therefore, the system is inconsistent.

(iii) Since, the last column is not a pivot column, therefore, the system is consistent.

Number of variables = 3

Number of pivots = 3

Hence, the system has a unique solution.

(iv) The last row is of the form $[0 \ 0 \ 0 \ 0 \ 0 \ | \ *]$. Two cases arise:

Case 1: '*' has the value 0. Then, the last row is $[0 \ 0 \ 0 \ 0 \ 0 \ | \ 0]$ and the system is consistent. In this case, there are three pivots so that there are $5-3 = 2$ free variables. Hence, there are infinitely many solutions.

Case 2: '*' has a non-zero value. Thus, the last column will be a pivot column so that the system is inconsistent.

Example 10.33. Consider a system of equations whose augmented matrix is 4×6 . Discuss the consistency and uniqueness of the solution when in the row echelon form of the augmented matrix.

- (i) *There are four pivot columns in the coefficient matrix.*
- (ii) *The sixth column is a pivot column.*
- (iii) *The coefficient matrix has a pivot position in every row.*

Solution: Since the augmented matrix is a 4×6 matrix, therefore,
 Number of equations = 4
 Number of variables = 5

- (i) With four pivots there should be a pivot in the last row (i.e. 4th row) of the coefficient matrix. Since, these pivots must occur in the coefficient matrix, therefore, it cannot be in the augmented column. Thus, the echelon form cannot have a row of the form $[0 \ 0 \ 0 \ 0 \ 0 \ | \ b]$, with b non-zero. Hence, the system is consistent. But, number of basic variables = number of pivot columns = 4.
 Hence, there is one free variable, so that there are infinitely many solutions.
- (ii) In this case, the last non-zero row will be $[0 \ 0 \ 0 \ 0 \ 0 \ | \ b]$ with b non-zero which gives inconsistency. Hence, the system will be inconsistent in this case.
- (iii) When there is a pivot in each of the 4 rows of the coefficient matrix, there are four pivots. So that there will be four pivot columns in the coefficient matrix. So the consistency follows by (i).

Example 10.34. *Consider a system of four equations in four variables. Discuss the consistency and uniqueness of the solution when in the row echelon form of the augmented matrix:*

1. *The coefficient matrix has a pivot in every column.*
2. *The coefficient matrix has pivot in every row.*
3. *The augmented matrix has pivot in augmented column.*
4. *The augmented matrix has four pivot columns.*
5. *The augmented matrix has three pivot columns.*
6. *The coefficient matrix has three pivot columns.*

Solution: In this case the coefficient matrix is a 4×4 matrix and the augmented matrix is a 4×5 matrix.

1. In this case, Number of pivot columns = 4. Thus, every column of the coefficient matrix is a pivot column, and there will be a pivot in the last row of the coefficient matrix. Therefore, the echelon form cannot have a row of the form $[0 \ 0 \ 0 \ 0 \ | \ b]$, with $b \neq 0$ and every variable is a basic variable. Hence the system is consistent and has a unique solution.
2. In this case also, the number of pivots = 4. By same argument as in (i), the system is consistent and has a unique solution.
3. If there is a pivot in the augmented column then there is a row of the $[0 \ 0 \ 0 \ 0 \ | \ b]$, with $b \neq 0$ which gives inconsistency. Hence, the system will be inconsistent in this case.

4. Since, there are four pivot columns in the augmented matrix, two cases arise:
 - (a) *Case 1:* Augmented column is not a pivot column. Thus, every pivot column occurs in coefficient matrix which is same as part (1).
 - (b) *Case 2:* Augmented column is a pivot column, which is same as (3).
5. Here also, two cases arise:
 - (a) *Case 1:* Augmented column is not a pivot column. Thus, there are three pivot columns in the coefficient matrix. The last row will be of the form $[0\ 0\ 0\ 0\ | \ 0]$, giving three basic variables and one free variable. Hence, the system is consistent having infinitely many solutions.
 - (b) *Case 2:* Augmented column is a pivot column. This is same as part (iii).
6. This is the same as case-a of (5).

Example 10.35. Consider a system of m equations in n variables. Discuss the consistency and uniqueness of the solution when

1. The system is under determined, i.e., $m < n$.
2. The system is over determined, i.e., $m > n$.

Solution: The augmented matrix is a $m \times (n+1)$ matrix and it will be transformed to row echelon form to discuss the consistency.

1. $m < n$: Thus, there can be at most m pivots, so that the number of pivot columns is less than the number of columns in the coefficient matrix. Two cases arise:
 - (a) *Case 1:* Augmented column is a pivot column. As discussed, earlier, such a system is inconsistent.
 - (b) *Case 2:* Augmented column is not a pivot column. Hence, the system is consistent. Thus, the number of pivot columns = $m < n$ = number of variables. Since, number of basic variables = number of pivot columns, therefore, the number of basic variables is less than the number of variables. Consequently, there is at least one free variable, so that there are infinitely many solutions.
2. $m > n$: Three cases arise:
 - (a) *Case 1:* The augmented column is a pivot column. Such a system is inconsistent.
 - (b) *Case 2:* Augmented column is not a pivot column, and number of pivot columns is n . Thus, the number of basic variables is equal to the number of variables so that there are no free variables. Hence, there is a unique solution.
3. *Case 3:* Augmented column is not a pivot column, and number of pivot columns is less than n . In this case, the number of basic variables is less than the number of variables so that there are free variables. Hence, the system has infinitely many solutions.

Example 10.36. Discuss the nature of the solution of the following system for different values of k . Obtain the solution, if it exists

$$\begin{aligned} kx + y + z &= 1 \\ x + ky + z &= k \\ x + y + kz &= k^2 \end{aligned}$$

Solution: The augmented matrix is

Step 1

$$A|b = \left(\begin{array}{ccc|c} k & 1 & 1 & 1 \\ 1 & k & 1 & k \\ 1 & 1 & k & k^2 \end{array} \right)$$

Step 2 Apply $R_1 \leftrightarrow R_3$

$$A|b \sim \left(\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 1 & k & 1 & k \\ k & 1 & 1 & 1 \end{array} \right)$$

Apply $R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - kR_1$

$$A|b \sim \left(\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & k-1 & 1-k & k-k^2 \\ 0 & 1-k & 1-k^2 & 1-k^3 \end{array} \right)$$

Step 3 $R_3 \rightarrow R_3 + R_2$

$$A|b \sim \left(\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & k-1 & 1-k & k-k^2 \\ 0 & 0 & 2-k-k^2 & 1+k-k^2-k^3 \end{array} \right) = A_1(\text{say})$$

The augmented matrix is reduced to echelon form.

Step 4 Three cases arise:

Case 1: Third row has a pivot and it is in the fourth column. Then $(k-1)(k+2) = 0$ and $(k-1)(k+1)^2 \neq 0$. This is possible for $k = -2$. In this case the system is inconsistent.

Case 2: Third row has a pivot and it is in the third column. Then, $(k-1)(k+2) \neq 0$. This is possible for $k \neq -2, 1$. In this case every column has a pivot, so the system is consistent and has a unique solution. Thus,

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & k-1 & 1-k & k-k^2 \\ 0 & 0 & 2-k-k^2 & 1+k-k^2-k^3 \end{array} \right)$$

Apply $R_2 \rightarrow \frac{1}{k-1}R_2, R_3 \rightarrow \frac{-1}{k-1}R_3$ (in order to simplify the calculations)

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & k & k^2 \\ 0 & 1 & -1 & -k \\ 0 & 0 & k+2 & (k+1)^2 \end{array} \right)$$

The corresponding system is

$$\begin{aligned}x + y + kz &= k^2 \\ y - z &= -k \\ (k + 2)z &= (k + 1)^2\end{aligned}$$

By back-substitution, we get $x = -\frac{k+1}{k+2}, y = \frac{1}{k+2}, z = \frac{(k+1)^2}{k+2}$.

Case 3: Third row does not have a pivot.

Thus, $(k - 1)(k + 2) = 0$ and $(k - 1)(k + 1)(k + 1) = 0$. This is possible when $k = 1$. In this case

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Thus, the system is consistent and has infinitely many solutions. The corresponding equations are

$$\begin{aligned}x + y + z &= 1 \\ \Rightarrow x &= 1 - y - z\end{aligned}$$

y, z are free.

Summarizing, we get, the system is inconsistent if $k = -2$. The system has infinitely many solutions if $k = 1$, and the solution is

$$\begin{aligned}x + y + z &= 1 \\ \Rightarrow x &= 1 - y - z\end{aligned}$$

y, z are free.

The system has a unique solution when $k \neq 1, -2$. In this case, the solution is $x = -\frac{k+1}{k+2}, y = \frac{1}{k+2}, z = \frac{(k+1)^2}{k+2}$.

Example 10.37. Find the real values of λ for which the following system has non-zero solution and also find the solution.

$$\begin{aligned}x + 2y + 3z &= \lambda x \\ 3x + y + 2z &= \lambda y \\ 2x + 3y + z &= \lambda z\end{aligned}$$

Solution: The given system is

$$\begin{aligned}(1 - \lambda)x + 2y + 3z &= 0 \\ 3x + (1 - \lambda)y + 2z &= 0 \\ 2x + 3y + (1 - \lambda)z &= 0\end{aligned}$$

Step 1 The augmented system is

$$[A|b] = \left(\begin{array}{ccc|c} 1-\lambda & 2 & 3 & 0 \\ 3 & 1-\lambda & 2 & 0 \\ 2 & 3 & 1-\lambda & 0 \end{array} \right)$$

Step 2 Apply $R_1 \rightarrow R_1 + R_2 + R_3$

$$[A|b] \sim \left(\begin{array}{ccc|c} 6-\lambda & 6-\lambda & 6-\lambda & 0 \\ 3 & 1-\lambda & 2 & 0 \\ 2 & 3 & 1-\lambda & 0 \end{array} \right)$$

If $6 - \lambda \neq 0$ then apply $R_1 \rightarrow \frac{1}{6-\lambda}R_1$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 3 & 1-\lambda & 2 & 0 \\ 2 & 3 & 1-\lambda & 0 \end{array} \right)$$

Apply $R_2 \rightarrow R_2 - 3R_1, R_3 \rightarrow R_3 - 2R_1$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -2-\lambda & -1 & 0 \\ 0 & 1 & -1-\lambda & 0 \end{array} \right)$$

Step 3 Apply $R_2 \leftrightarrow R_3$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1-\lambda & 0 \\ 0 & -2-\lambda & -1 & 0 \end{array} \right)$$

Apply $R_3 \rightarrow R_3 + (2 + \lambda)R_2$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1-\lambda & 0 \\ 0 & 0 & -(\lambda^2 + 3\lambda + 3) & 0 \end{array} \right)$$

Step 4 The matrix $[A|b]$ is reduced to echelon form. The corresponding system is

$$\begin{aligned} x + y + z &= 0 \\ y - (1 + \lambda)z &= 0 \\ -(\lambda^2 + 3\lambda + 3)z &= 0 \end{aligned}$$

For any real value of λ , $\lambda^2 + 3\lambda + 3 \neq 0$.

By back-substitution, we get $x = y = z = 0$

Step 5 If $\lambda = 6$, then from Step2, we get

$$[A|b] \sim \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 3 & -5 & 2 & 0 \\ 2 & 3 & -5 & 0 \end{array} \right)$$

Apply $R_3 \rightarrow R_3 - R_2$ (this is done to simplify the calculations). Apply $R_1 \leftrightarrow R_3$.

Apply $R_2 \rightarrow R_2 + 3R_1$. Apply $R_2 \rightarrow \frac{1}{19}R_2$.

$$[A|b] \sim \left(\begin{array}{ccc|c} -1 & 8 & -7 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

The corresponding system is

$$\begin{aligned} -x + 8y - 7z &= 0 \\ y - z &= 0 \\ 0 &= 0 \end{aligned}$$

By back substitution we get,

$$x=z, y=z$$

z is free.

Summarizing, we get,

If $\lambda \neq 6$ the system has only the trivial solution. If $\lambda = 6$ the solution is $x = z$, $y = z$, z is free.

Example 10.38. Discuss the existence and uniqueness of the solution of the system

$$\begin{aligned} x + y + z &= 1 \\ \alpha x + \beta y + \gamma z &= \varepsilon \\ \alpha^2 x + \beta^2 y + \gamma^2 z &= \varepsilon^2 \end{aligned}$$

Solution:

Step 1 The augmented matrix of the system is

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \varepsilon \\ \alpha^2 & \beta^2 & \gamma^2 & \varepsilon^2 \end{array} \right)$$

Step 2 Apply $R_2 \rightarrow R_2 - \alpha R_1, R_3 \rightarrow R_3 - \alpha^2 R_1$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & \beta - \alpha & \gamma - \alpha & \varepsilon - \alpha \\ 0 & \beta^2 - \alpha^2 & \gamma^2 - \alpha^2 & \varepsilon^2 - \alpha^2 \end{array} \right)$$

Step 3

Case 1: Taking $\beta - \alpha \neq 0$, i.e., $\alpha \neq \beta$, apply $R_2 \rightarrow \frac{1}{\beta - \alpha} R_2, R_3 \rightarrow \frac{1}{\beta - \alpha} R_3$

$$[A|b] = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{\gamma - \alpha}{\beta - \alpha} & \frac{\varepsilon - \alpha}{\beta - \alpha} \\ 0 & \beta + \alpha & \frac{\gamma^2 - \alpha^2}{\beta - \alpha} & \frac{\varepsilon^2 - \alpha^2}{\beta - \alpha} \end{array} \right)$$

Apply $R_3 \rightarrow R_3 - (\alpha + \beta)R_2$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{\gamma - \alpha}{\beta - \alpha} & \frac{\varepsilon - \alpha}{\beta - \alpha} \\ 0 & 0 & \frac{(\gamma - \alpha)(\gamma - \beta)}{\beta - \alpha} & \frac{(\varepsilon - \alpha)(\varepsilon - \beta)}{\beta - \alpha} \end{array} \right)$$

Apply $R_3 \rightarrow (\beta - \alpha)R_3$

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{\gamma - \alpha}{\beta - \alpha} & \frac{\varepsilon - \alpha}{\beta - \alpha} \\ 0 & 0 & (\gamma - \alpha)(\gamma - \beta) & (\varepsilon - \alpha)(\varepsilon - \beta) \end{array} \right) = A_1(\text{say})$$

The matrix has been reduced to row echelon form.

Step 4 We now have the following possibilities regarding α, β, γ .

1. If $\gamma \neq \alpha$ and $\gamma \neq \beta$. Thus, α, β, γ are all distinct. Thus, $(\gamma - \alpha)(\gamma - \beta) \neq 0$ so that the given system has a unique solution, as A_1 has last row of the form $[0 \ 0 \ p \ | \ q]$, with $p \neq 0$.
2. If $(\gamma - \alpha)(\gamma - \beta) = 0$. Then, $\gamma = \alpha$ or $\gamma = \beta$. In this case, the system is consistent if $(\varepsilon - \beta)(\varepsilon - \alpha) = 0$, i.e., $\varepsilon = \alpha$ or $\varepsilon = \beta$. In this case, A_1 has a row of the form $[0 \ 0 \ 0 \ | \ 0]$ so that the system has infinitely many solutions. Thus, if $\gamma = \alpha$ or β , then, the system is consistent and has infinitely many solutions if $\varepsilon = \alpha$ or β .
3. In case, $(\gamma - \alpha)(\gamma - \beta) = 0$, i.e., $\gamma = \alpha$ or $\gamma = \beta$ and $(\varepsilon - \alpha)(\varepsilon - \beta) \neq 0$ then the last row is of the form $[0 \ 0 \ 0 \ | \ r]$ with $r \neq 0$, so that the system is inconsistent.

Step 5 We have discussed the following cases:

1. α, β, γ are all distinct.
2. $\alpha \neq \beta$ and γ is one of α or β . By symmetry, this covers the cases when only two of α, β, γ are distinct.

We have yet to discuss the case when α, β, γ are all equal.

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & \varepsilon - \alpha \\ 0 & 0 & 0 & \varepsilon^2 - \alpha^2 \end{array} \right)$$

Thus, the system is inconsistent if $\varepsilon \neq \alpha$ and consistent if $\varepsilon = \alpha$. When $\varepsilon = \alpha$,

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

so that the system has infinitely many solutions.

Summarizing, we get that the given system is

- Inconsistent when:
 1. $\alpha = \beta = \gamma \neq \varepsilon$, i.e., α, β, γ are all equal but ε is different from them, or
 2. when $\alpha \neq \beta, \gamma$ is one of α or β and $\varepsilon \neq \alpha, \beta$, i.e., two of α, β, γ are distinct and ε is different from the two distinct values.
- Unique solution when
 $\alpha \neq \beta \neq \gamma$, i.e., α, β, γ are all distinct and ε can take any value.
- Infinitely many solutions
 1. when $\alpha \neq \beta, \gamma$ is either α or β and ε take the value α or β , i.e., two of α, β, γ are distinct and ε is equal to one of the two distinct values.
 2. or when $\alpha = \beta = \gamma = \varepsilon$, i.e., $\alpha, \beta, \gamma, \varepsilon$ are all equal.

10.17 Exercise

1. Let

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{pmatrix}$$

In each of the following parts, determine whether X is a solution to the linear system $AX = b$.

(i) $X = [-1 \ 2 \ -3]^t, b = [0 \ 0 \ 0]^t$

(ii) $X = [0 \ 0 \ 0]^t, b = [0 \ 0 \ 0]^t$

(iii) $X = [-1 \ 1 \ 2]^t, b = [3 \ 6 \ -5]^t$

(iv) $X = [1 \ 2 \ -3]^t, b = [2 \ -5 \ -2]^t$

2. Find the general solution of the systems whose augmented matrices are given:

(i)

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & -1 \\ 1 & -2 & 7 & -6 \end{array} \right)$$

(ii)

$$\left(\begin{array}{ccc|c} -6 & -9 & 12 & -15 \\ -1 & -1.5 & 2 & -2.5 \\ 2 & 3 & -4 & 5 \end{array} \right)$$

(iii)

$$\left(\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 2 & -5 & -6 & 0 & -5 \\ 0 & 1 & -6 & -3 & 1 & 2 \\ 1 & 3 & -11 & -9 & 1 & -3 \end{array} \right)$$

(iv)

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -2 & 3 \end{array} \right)$$

(v)

$$\left(\begin{array}{cccc|c} 2 & 1 & 1 & -2 & 1 \\ 3 & -1 & 1 & -6 & -2 \\ 1 & 1 & -1 & -1 & -1 \\ 6 & 0 & 1 & -9 & -2 \\ 5 & -1 & 2 & -8 & 3 \end{array} \right)$$

(vi)

$$\left(\begin{array}{cccc|c} 3 & 1 & 1 & -1 & 1 \\ 1 & 1 & 2 & 3 & 13 \\ 1 & -2 & 1 & 1 & 8 \end{array} \right)$$

3. Determine the existence and uniqueness of the solution of the following linear systems:

$$\begin{aligned}
 \text{(i)} \quad & 2x_1 + x_2 + 3x_3 = 1 \\
 & 4x_1 - x_3 = 5 \\
 & 3x_1 + x_2 + x_3 = 4
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad & 2x_1 + x_2 + x_3 = 3 \\
 & x_1 + 2x_2 - 4x_3 = 3 \\
 & x_2 - 3x_3 = 1 \\
 & -x_1 - 2x_3 = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad & x_1 + 2x_2 - x_3 + x_4 = 1 \\
 & 2x_1 + x_2 + 4x_3 + x_4 = 1 \\
 & x_1 + x_2 - x_3 - x_4 = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{(iv)} \quad & 2x_1 - 2x_2 + 4x_4 = 2 \\
 & -x_1 + 3x_3 + x_4 = 6 \\
 & 6x_1 - 6x_2 + x_3 + 8x_4 = 3
 \end{aligned}$$

4. Solve the following system of linear equations by reducing the augmented matrix is row echelon form:

$$\begin{aligned}
 \text{(i)} \quad & 6w - 6x + y + 8z = 12 \\
 & -w - x - 2z = -1 \\
 & w - 3y - z = -6 \\
 & w + 2x - 7y - 16z = -7
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad & 2x - y - 3z = 5 \\
 & 3y + z = -5 \\
 & x - 2y + 3z = 4 \\
 & 3x - 3y = 7
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad & w + x + y + z = 3 \\
 & -3w - 17x + y + 2z = 1 \\
 & 4w - 17x + 8y - 5z = 1 \\
 & -5x - 2y + z = 1
 \end{aligned}$$

$$\begin{aligned}
 \text{(iv)} \quad & 2y - z = 1 \\
 & 4x - 10y + 3z = 5 \\
 & 3x - 3y = 6
 \end{aligned}$$

$$\begin{aligned}
 \text{(v)} \quad & 2x + y + z = 1 \\
 & x + 2y - 4z = -1 \\
 & y - 3z = -1 \\
 & -x - 2z = -1
 \end{aligned}$$

5. Solve the following systems of linear equations by reducing the augmented matrix to reduced row echelon form.

$$\begin{aligned} \text{(i)} \quad & 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 = 9 \\ & 3x_2 - 6x_3 + 6x_4 + 4x_5 = -5 \\ & 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad & 3x_1 + 4x_2 + x_3 = 0 \\ & 2x_1 + 5x_2 + 3x_3 = 0 \\ & -x_1 + x_2 + 2x_3 = 0 \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad & -x_1 + 3x_3 + x_4 = 6 \\ & 2x_1 - 2x_2 + 4x_4 = 2 \\ & 6x_1 - 6x_2 + x_3 + 8x_4 = 12 \\ & x_1 + 2x_2 - 7x_3 - 16x_4 = 7 \end{aligned}$$

$$\begin{aligned} \text{(iv)} \quad & 2x_1 + x_2 + x_3 + x_4 = 1 \\ & x_1 + 2x_2 + x_3 + x_4 = 2 \\ & x_1 + x_2 + 2x_3 + x_4 = 3 \\ & x_1 + x_2 + x_3 + 2x_4 = 4 \end{aligned}$$

$$\begin{aligned} \text{(v)} \quad & x_1 + x_2 + x_3 = 12 \\ & x_1 + 2x_2 + 4x_3 = 15 \\ & x_1 + 3x_2 + 9x_3 = 16 \end{aligned}$$

6. Find the general solution of the systems whose augmented matrices are given. Give the solution in parametric form.

$$\text{(i)} \quad \left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 3 \\ 1 & 2 & 3 & -1 & 0 \\ -1 & 1 & 0 & -1 & 2 \end{array} \right)$$

$$\text{(ii)} \quad \left(\begin{array}{cccc|c} 0 & -1 & 3 & 0 & 1 \\ 2 & 3 & 2 & 2 & 10 \\ 3 & 5 & 5 & 3 & 18 \end{array} \right)$$

$$\text{(iii)} \quad \left(\begin{array}{ccccc|c} 0 & 1 & 0 & 0 & -4 & 1 \\ 1 & -2 & 0 & -1 & -4 & -1 \\ 1 & -3 & 0 & 0 & 9 & 2 \\ 1 & -3 & 0 & -1 & 0 & -2 \end{array} \right)$$

7. Show that the system

$$\begin{aligned}5x + 2y + 7z &= 4 \\3x + 26y + 2z &= 9 \\7x + 2y + 10z &= 5\end{aligned}$$

is consistent and also find the solution.

8. Solve the following system if consistent:

$$\begin{aligned}5x_1 + 3x_2 + 14x_3 &= 4 \\x_2 + 2x_3 &= 0 \\2x_1 + x_2 + 6x_3 &= 2\end{aligned}$$

9. Solve the following equations:

$$\begin{aligned}\text{(i)} \quad & x_1 + 3x_2 + 2x_3 = 0 \\& 2x_1 - x_2 + 3x_3 = 0 \\& 3x_1 - 5x_2 + 4x_3 = 0 \\& x_1 + 17x_2 + 4x_3 = 0\end{aligned}$$

$$\begin{aligned}\text{(ii)} \quad & x_1 - x_2 + x_3 = 0 \\& 3x_1 - x_2 + 4x_3 = 0 \\& 7x_1 - 3x_2 - 9x_3 = 0 \\& 4x_1 - 2x_2 + 5x_3 = 0\end{aligned}$$

10. If

$$A = \begin{pmatrix} 1 & 0 & 5 \\ 1 & 1 & 1 \\ 0 & 1 & -4 \end{pmatrix}$$

Find the general solution of the system:

$$\begin{aligned}\text{(i)} \quad & (2I_3 - A)X = 0 \\ \text{(ii)} \quad & (-4I_3 - A)X = 0\end{aligned}$$

11. Which of the following systems of linear equations possess trivial or non-trivial solutions?

$$\begin{aligned}\text{(i)} \quad & x_1 - 2x_2 + x_3 - x_4 = 0 \\& x_1 + x_2 - 2x_3 + 3x_4 = 0 \\& 4x_1 + x_2 - 5x_3 + 8x_4 = 0 \\& 5x_1 - 7x_2 + 2x_3 - x_4 = 0\end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad & x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\
 & 8x_1 + 5x_2 + x_3 + 4x_4 = 0 \\
 & 5x_1 + 6x_2 + 8x_3 + x_4 = 0 \\
 & 8x_1 + 3x_2 + 7x_3 + 2x_4 = 0
 \end{aligned}$$

12. Find the condition on a , b , c so that the following system is consistent:

$$\begin{aligned}
 \text{(i)} \quad & x + 2y - 3z = a \\
 & 2x + 3y + 3z = b \\
 & 5x + 9y - 6z = c
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad & x + 2y + 6z = a \\
 & 2x - 3y - 2z = b \\
 & 3x - y + 4z = c
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad & 4x + 2y + z = a \\
 & 2x - y + 3z = b \\
 & x + 3y - 2z = c
 \end{aligned}$$

13. For the system of linear equations:

$$\begin{aligned}
 8x_1 + x_2 &= b_1 \\
 x_3 + 3x_4 &= b_2 \\
 8x_1 + x_2 + 4x_3 + 8x_4 &= b_3 \\
 16x_1 + 2x_2 + x_3 + 2x_4 &= b_4
 \end{aligned}$$

Find the condition on b_1, b_2, b_3 and b_4 so that the system is consistent. When it is, find a general solution.

14. Given the echelon form of the augmented matrix of a linear system, discuss the nature of the solution:

$$\text{(i)} \quad \left(\begin{array}{cccc|c} \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & \blacksquare \end{array} \right)$$

$$\text{(ii)} \quad \left(\begin{array}{cccc|c} 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$$\text{(iii)} \quad \left(\begin{array}{ccc|c} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \end{array} \right)$$

(iv)

$$\left(\begin{array}{cccc|c} \blacksquare & 0 & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & * \end{array} \right)$$

(v)

$$\left(\begin{array}{ccccc|c} \blacksquare & * & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * \end{array} \right)$$

(vi)

$$\left(\begin{array}{ccc|cc} 0 & \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

15. Suppose the augmented matrix of a linear system is a 3×6 matrix. Is the system consistent in the following cases? Justify your answer.

- (i) The coefficient matrix has 3 pivot columns.
- (ii) The 6th column is a pivot column.
- (iii) The augmented matrix has 3 pivot columns.
- (iv) The coefficient matrix has a pivot in every row.
- (v) The augmented matrix has a pivot in every row.

16. Suppose the coefficient matrix of a linear system of four equations in four variables has a pivot in each column. Explain why the system has a unique solution.

17. Suppose the coefficient matrix of a linear is a 4×4 matrix. What would you have to know about the pivot columns in the augmented matrix in order to know that the linear system is consistent and has a unique solution.

18. Give an example of a matrix in echelon form whose corresponding linear system is:

- (i) inconsistent under determined system with 4 variables.
- (ii) inconsistent over determined system with 3 variables.
- (iii) consistent under determined system with 3 variables.
- (iv) consistent over determined system with 2 variables.

19. For what value of k does the following system of equations have a solution.

$$\begin{aligned} x + y + z &= 1 \\ x + 2y + 4z &= k \\ 1x + 4y + 10z &= k^2 \end{aligned}$$

Find the solution in each case.

20. Discuss the existence and uniqueness of the solutions of the following system of equations, for all values of k .

$$\begin{aligned}2x + 3ky + (3k + 4)z &= 0 \\x + (k + 4)y + (4k + 2)z &= 0 \\x + 2(k + 1)y + (3k + 4)z &= 0\end{aligned}$$

21. Discuss the existence and uniqueness of the system of equations:

$$\begin{aligned}x + y + z &= 1 \\ax + by + cz &= k \\a^2x + b^2y + c^2z &= k^2\end{aligned}$$

Also obtain the solution, if it exists:

- (a) a, b, c are all distinct.
- (b) $a \neq b, a = c$.
- (c) $a = b = c$.

10.18 Solution Sets of Linear Systems

So far we have learnt to obtain the solution of a given system of linear equations. Obtaining the solution set of a linear system is an important object of study in linear algebra and it will appear later in several different contexts. We shall now find the general solution of a given homogeneous and non-homogeneous system and express it using the vector notation. A geometric description of the general solution will also be given.

10.18.1 Homogeneous System

Let us recall that the homogeneous linear system is written as $\underline{A}\underline{X} = \underline{0}$ in matrix equation form.

In matrix form, Theorem 1 Section 10.7 is restated as follows:

Theorem 10.9. *The homogeneous equation $\underline{A}\underline{X} = \underline{0}$ has non-trivial solution if and only if the solution has at least one free variable.*

Example 10.39. *Describe all the solutions of the homogeneous system $2x_1 - 3x_2 + 4x_3 = 0 \dots (i)$. Give the geometrical interpretation also.*

Solution: A single linear equation can be treated as a simple system of equations. There is no need for matrix notation. Any of the three variables can be treated as a basic variable. Choosing x_1 as a basic variable, we get $x_1 = \frac{3}{2}x_2 - 2x_3$.

The general solution is

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1.5x_2 - 2x_3 \\ x_2 \\ x_3 \end{pmatrix}$$

$$x_2 \begin{pmatrix} 1.5 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = x_2 \underline{u} + x_3 \underline{v}$$

with x_2, x_3 as free variables,

$$\underline{u} = \begin{pmatrix} 1.5 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{v} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

Thus, every solution of (i) is a linear combination of the vectors \underline{u} and \underline{v} . Hence, the solution set is $\text{span} \{ \underline{u}, \underline{v} \}$. Since, \underline{u} is not a scalar multiple of \underline{v} , therefore, \underline{u} and \underline{v} are non-collinear. Hence, $\text{span} \{ \underline{u}, \underline{v} \}$ is a plane through the origin containing \underline{u} and \underline{v} .

$X = x_2 \underline{u} + x_3 \underline{v}$ where x_2, x_3 are parameters is the parametric form of the solution.

Example 10.40. Determine if the following homogeneous system has a non-trivial solution. Also describe the solution set.

$$2x_1 + 7x_2 - x_3 = 0$$

$$2x_1 - 5x_2 + 8x_3 = 0$$

$$4x_1 + 2x_2 + 7x_3 = 0$$

Solution: Here

$$A = \begin{pmatrix} 2 & 7 & -1 \\ 2 & -5 & 8 \\ 4 & 2 & 7 \end{pmatrix}$$

is the matrix of coefficients. We reduce $[A|0]$ to echelon form.

$$[A|0] = \left(\begin{array}{ccc|c} 2 & 7 & -1 & 0 \\ 2 & -5 & 8 & 0 \\ 4 & 2 & 7 & 0 \end{array} \right)$$

Apply $R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - 2R_1$

$$\sim \left(\begin{array}{ccc|c} 2 & 7 & -1 & 0 \\ 0 & -12 & 9 & 0 \\ 0 & -12 & 9 & 0 \end{array} \right)$$

Apply $R_3 \rightarrow R_3 - R_2$

$$\sim \left(\begin{array}{ccc|c} \mathbf{2} & 7 & -1 & 0 \\ 0 & \mathbf{-12} & 9 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

The pivots are bold-faced.

Number of variables = 3

Number of basic variables = Number of pivots = 2

Hence, number of free variables = $3 - 2 = 1$. Since, there is one free variable, therefore, the system has non-trivial solutions. To find the solution we obtain

the reduced echelon form of the augmented matrix.

Apply $R_2 \rightarrow \frac{1}{3}R_2$

$$[A|0] \sim \left(\begin{array}{ccc|c} 2 & 7 & -1 & 0 \\ 0 & -4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Apply $R_1 \rightarrow R_1 + \frac{7}{4}R_2$

$$\sim \left(\begin{array}{ccc|c} 2 & 0 & \frac{17}{4} & 0 \\ 0 & -4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Apply $R_1 \rightarrow \frac{1}{2}R_1$, $R_2 \rightarrow -\frac{1}{4}R_2$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & \frac{17}{8} & 0 \\ 0 & 1 & -\frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

The equivalent system is

$$\begin{aligned} x_1 + \frac{17}{8}x_3 &= 0 \\ x_2 - \frac{3}{4}x_3 &= 0 \end{aligned}$$

Thus,

$$\begin{aligned} x_1 &= -\frac{17}{8}x_3 \\ x_2 &= \frac{3}{4}x_3 \\ x_3 &\text{ is free.} \end{aligned}$$

In the vector form, the general solution is

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\frac{17}{8}x_3 \\ \frac{3}{4}x_3 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -\frac{17}{8} \\ \frac{3}{4} \\ 1 \end{pmatrix} = x_3 \underline{u}$$

where $\underline{u} = \begin{pmatrix} -17/8 \\ 3/4 \\ 1 \end{pmatrix}$ and x_3 is a parameter.

This is the parametric vector form of the solution. Here, x_3 is factored out of the expression for the general solution vector. Thus, every solution of $AX = 0$ is a scalar multiple of \underline{u} . The trivial solution is obtained by choosing $x_3 = 0$. In order to avoid fractions, we can write

$$X = x_3 \frac{1}{8} \begin{pmatrix} -17 \\ 6 \\ 8 \end{pmatrix} = k\underline{v}$$

where $k = x_3 \frac{1}{8}$, $\underline{v} = \begin{pmatrix} -17 \\ 6 \\ 8 \end{pmatrix}$. As x_3 is arbitrary, so k is also arbitrary.

Geometrically, the solution set is a line through origin in \mathbb{R}^3 .

In the above two examples, we observe that general solution of a homogeneous system is a linear combination of a certain set of p vectors, which themselves are solutions of the given system. Moreover, p is the number of free variables.

10.18.2 Non-homogeneous System

Before studying the general vector form of the solution of a non-homogeneous linear system, we will consider an example. In order to compare the relation between the solutions of a non-homogeneous system $\underline{A}\underline{X} = \underline{b}$ and the corresponding homogeneous system $\underline{A}\underline{X} = \underline{0}$, we consider the system $\underline{A}\underline{X} = \underline{b}$ for which the corresponding homogeneous system $\underline{A}\underline{X} = \underline{0}$ is that of Example 10.40.

Example 10.41. We describe the general solution of

$$\begin{aligned} 2x_1 + 7x_2 - x_3 &= -7 \\ 2x_1 - 5x_2 + 8x_3 &= 23 \\ 4x_1 + 2x_2 + 7x_3 &= 16 \end{aligned}$$

Solution: The reduced echelon form of the augmented matrix of the system is

$$\left(\begin{array}{ccc|c} 1 & 0 & \frac{17}{8} & \frac{21}{4} \\ 0 & 1 & -\frac{3}{4} & -\frac{5}{2} \\ 0 & 0 & 0 & 0 \end{array} \right)$$

The general solution is:

$$\begin{aligned} x_1 &= -\frac{17}{8}x_3 + \frac{21}{4} \\ x_2 &= \frac{3}{4}x_3 - \frac{5}{2} \\ x_3 &\text{ is free.} \end{aligned}$$

There is one free variable x_3 . In vector form, the general solution is written as

$$\begin{aligned} \underline{X} &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{21}{4} - \frac{17}{8}x_3 \\ -\frac{5}{2} + \frac{3}{4}x_3 \\ x_3 \end{pmatrix} \\ \underline{X} &= \begin{pmatrix} \frac{21}{4} \\ -\frac{5}{2} \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{17}{8}x_3 \\ \frac{3}{4}x_3 \\ x_3 \end{pmatrix} \end{aligned}$$

Hence, $\underline{X} = \underline{a} + x_3\underline{u}$, where

$$\underline{a} = \begin{pmatrix} \frac{21}{4} \\ -\frac{5}{2} \\ 0 \end{pmatrix}, \underline{u} = \begin{pmatrix} -\frac{17}{8} \\ \frac{3}{4} \\ 1 \end{pmatrix}$$

Thus $\underline{X} = \underline{a} + k\underline{u}$, where $x_3 = k$ is a parameter which can take any real value. \therefore the solution set of $A\underline{X} = \underline{b}$ is $\{\underline{a} + k\underline{u} \mid k \in \mathbb{R}\}$.

Note that

1. $\underline{X} = k\underline{u}$ is the general solution of the corresponding homogeneous system $A\underline{X} = \underline{0}$.
2. $\underline{X} = \underline{a}$ is a particular solution of $A\underline{X} = \underline{b}$ (This is the solution corresponding to $k = 0$).
3. The general solution of $A\underline{X} = \underline{b}$ is obtained by adding a particular solution of $A\underline{X} = \underline{b}$ to the general solution of $A\underline{X} = \underline{0}$.

Geometrically, $\underline{X} = \underline{a} + k\underline{u}$ where $k \in \mathbb{R}$ represents a line through the point ' \underline{a} ' and parallel to $\underline{X} = k\underline{u}$.

The following are the steps involved in writing the solution set of a consistent system $A\underline{X} = \underline{b}$ in parametric vector form.

Step 1 Reduce the augmented matrix $[A|b]$ to reduced echelon form.

Step 2 Express the basic variable in terms of the free variables (if any).

Step 3 Write a typical solution \underline{X} as a vector whose entries depend on the free variables.

Step 4 Express \underline{X} as a linear combination of vectors (with numeric entries) using the free variables as parameters.

The relation between the solution set of $A\underline{X} = \underline{0}$ as shown in the above illustration generalizes to any consistent equation $A\underline{X} = \underline{b}$. The solution set may be larger than a line when there are more than one free variables. The following theorem gives us the precise statement.

Theorem 10.10. *Suppose the equation $A\underline{X} = \underline{b}$ is consistent for some given \underline{b} , and let \underline{a} be a solution. Then the solution set of $A\underline{X} = \underline{b}$ is the solution set of all vectors of the form $\underline{w} = \underline{a} + u_h$ where u_h is any solution of the homogeneous system $A\underline{X} = \underline{0}$.*

Proof: Given that \underline{a} is the solution of $A\underline{X} = \underline{b}$. Then $A\underline{a} = \underline{b}$. Let u_h be a solution of $A\underline{X} = \underline{0}$. Then, $Au_h = \underline{0}$. Define $\underline{w} = \underline{a} + u_h$. Now, $A\underline{w} = A(\underline{a} + u_h) = A\underline{a} + Au_h = \underline{b} + \underline{0} = \underline{b}$. Thus, $A\underline{w} = \underline{b}$, so that \underline{w} is a solution of $A\underline{X} = \underline{b}$.

On the other hand, let \underline{w} be any solution of $A\underline{X} = \underline{b}$. Then, $A\underline{w} = \underline{b}$. Using $A\underline{a} = \underline{b}$, we get, $A\underline{w} - A\underline{a} = \underline{0}$, so that $A(\underline{w} - \underline{a}) = \underline{0}$. Hence, $\underline{w} - \underline{a}$ is a solution of $A\underline{X} = \underline{0}$. Thus, $\underline{w} - \underline{a} = u_h$, for some solution u_h of $A\underline{X} = \underline{0}$. Hence, $\underline{w} = \underline{a} + u_h$. \square

Example 10.42. *Describe the solution set in parametric vector form of the system $A\underline{X} = \underline{b}$, where*

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 7 & -21 & 14 \\ -3 & 9 & -6 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 28 \\ -12 \end{pmatrix}$$

Also give the geometrical interpretation.

Solution: *Step 1* The augmented matrix is

$$[A|b] = \left(\begin{array}{ccc|c} 1 & -3 & 2 & 4 \\ 7 & -21 & 14 & 28 \\ -3 & 9 & -6 & -12 \end{array} \right)$$

and

$$[A|b] \sim \left(\begin{array}{ccc|c} 1 & -3 & 2 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

which is the reduced echelon form.

Step 2 The solution is: $x_1 = 4 + 3x_2 - 2x_3$ x_2, x_3 are free.

Step 3 The solution in vector form is

$$\begin{aligned} \underline{X} &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 + 3x_2 - 2x_3 \\ x_2 \\ x_3 \end{pmatrix} \\ \underline{X} &= \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 3x_2 \\ x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} -2x_3 \\ 0 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} x_3 \end{aligned}$$

$= \underline{a} + k_1 \underline{u}_1 + k_2 \underline{u}_2$ where,

$$\underline{a} = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{u}_1 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad x_2 = k_1, x_3 = k_2$$

k_1, k_2 are parameters which can take any real value.

Thus, the solution set of $A\underline{X} = \underline{b}$ is $\{\underline{a} + k_1 \underline{u}_1 + k_2 \underline{u}_2 \mid k_1, k_2 \in \mathbb{R}\}$.

$\underline{X} = \underline{a} + k_1 \underline{u}_1 + k_2 \underline{u}_2$ where $k_1, k_2 \in \mathbb{R}$ represents a plane passing through \underline{a} and containing the vectors \underline{u}_1 and \underline{u}_2 .

Example 10.43. Describe the solution set in parametric vector form of $A\underline{X} = \underline{0}$, where

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 7 & -21 & 14 \\ -3 & 9 & -6 \end{pmatrix}$$

Also give the geometrical interpretation.

Solution: The matrix A is the same as in the previous illustration. Repeating steps 1 and 2 on $[A|0]$ we get the solution as

$$\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_2 - 2x_3 \\ x_2 \\ x_3 \end{pmatrix} = k_1 \underline{u}_1 + \underline{u}_2 k_2$$

where

$$\underline{u}_1 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad x_2 = k_1, x_3 = k_2$$

k_1, k_2 are parameters which can take any real values.

Thus, $\underline{X} = k_1 \underline{u}_1 + k_2 \underline{u}_2$, $k_1, k_2 \in \mathbb{R}$ describes the solution set of $A\underline{X} = \underline{0}$ in parametric vector form.

This represents a plane through the origin and containing the vectors \underline{u}_1 and \underline{u}_2 .

Example 10.44. Describe the solution set in parametric vector form of Example 10.24 of Section 10.13. Also give the geometrical interpretation.

Solution: The solution set is $x_1 = 2, x_2 = -1, x_3 = 3$. In vector form, the solution can be written as

$$\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \underline{a}(\text{say})$$

Since, there is no free variable, the solution contains no parameters. Geometrically, the solutions is a point ' \underline{a} ' in \mathbb{R}^3 .

Example 10.45. Describe the solution set in parametric vector form of Example 10.23 of Section 10.13. Also give the geometrical interpretation.

Solution: The solution set is

$$\begin{aligned} x_1 &= -5 - 2x_4 \\ x_2 &= 2 + 3x_4 \\ x_3 &= 3 + 2x_4 \end{aligned}$$

x_4 is free.

In vector form, it can be written as

$$\begin{aligned} \underline{X} &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -5 - 2x_4 \\ 2 + 3x_4 \\ 3 + 2x_4 \\ x_4 \end{pmatrix} \\ \underline{X} &= \begin{pmatrix} -5 \\ 2 \\ 3 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 3 \\ 2 \\ 1 \end{pmatrix} = \underline{a} + \underline{u}k \end{aligned}$$

where

$$\underline{a} = \begin{pmatrix} -5 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \underline{u} = \begin{pmatrix} -2 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

$x_4 = k$ is a parameter which can take any real value.

Thus, $\underline{X} = \underline{a} + k\underline{u}$, $k \in \mathbb{R}$ describes the solution set in parametric vector form. This represents a straight line passing through the point \underline{a} and parallel to \underline{u} in \mathbb{R}^4 .

Example 10.46. Describe the solution set in parametric vector form of Example 10.25 of Section 10.13.

Solution: The solution set is

$$\begin{aligned} x_1 &= -2x_2 + 3x_4 + x_6 \\ x_3 &= 1 - 2x_6 \\ x_5 &= 2 - x_6 \end{aligned}$$

x_2, x_4, x_6 are free.

In vector form, the solution set can be written as

$$\begin{aligned} X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} &= \begin{pmatrix} -2x_2 + 3x_4 + x_6 \\ x_2 \\ 1 - 2x_6 \\ x_4 \\ 2 - x_6 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} -2x_2 \\ x_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 3x_4 \\ 0 \\ 0 \\ x_4 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} x_6 \\ 0 \\ -2x_6 \\ 0 \\ -x_6 \\ x_6 \end{pmatrix} \\ &= \underline{a} + x_2\underline{u}_1 + x_4\underline{u}_2 + x_6\underline{u}_3, \end{aligned}$$

where

$$\underline{u}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \underline{u}_2 = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \underline{u}_3 = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

$\therefore \underline{X} = \underline{a} + k_1\underline{u}_1 + k_2\underline{u}_2 + k_3\underline{u}_3$, where $k_1 = x_2, k_2 = x_4, k_3 = x_6$. k_1, k_2, k_3 are parameters and can take any real value.

Example 10.47. Describe the general solution of $A\underline{X} = \underline{b}$ where

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 1 & 0 \end{pmatrix}, \underline{b} = \begin{pmatrix} 4 \\ 2 \\ 0 \\ 4 \end{pmatrix}, \underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

in vector form, and give the geometrical interpretation.

Solution: Reduce the augmented matrix $[A|\underline{b}]$ to echelon form.

$$\begin{aligned} A &= \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 1 & 1 & 1 & -1 & 2 \\ 1 & -1 & 1 & -1 & 0 \\ 1 & 2 & 1 & 0 & 4 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & -2 & 0 & -2 & -4 \\ 0 & 1 & 0 & -1 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & -2 & 0 & -2 & -4 \\ 0 & 0 & 0 & -2 & -2 \end{array} \right) \\ &\sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & -2 & -2 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{aligned}$$

There are three pivot columns and the pivot elements have been bold faced. Now, we reduce it to reduced echelon form.

$$[A|\underline{b}] \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

The equivalent system is

$$\begin{aligned} x_1 + x_3 &= 2 \\ x_2 &= 1 \\ x_4 &= 1 \end{aligned}$$

Thus, the solution is

$$\begin{aligned}x_1 &= -x_3 + 2 \\x_2 &= 1 \\x_4 &= 1\end{aligned}$$

x_3 is free, so that there is one free variable. In vector form, the general solution is written as

$$\begin{aligned}X &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 - x_3 \\ 1 \\ x_3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ &= \underline{u} + k\underline{v}\end{aligned}$$

$$\text{where, } \underline{u} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \underline{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

and $x_3 = k$ is a parameter which can take any real value.

Thus, $\underline{X} = \underline{u} + k\underline{v}$, $k \in \mathbb{R}$ describes the solution set of $A\underline{X} = \underline{b}$ in parametric vector form.

10.19 Exercise

For the following questions of Exercise 10.17, write the solution sets in the parametric vector form and interpret them geometrically:

1. Q-4
2. Q-5
3. Q-6
4. Q-7
5. Q-8
6. Q-9

10.20 Answers to Exercises

Exercise - 10.5

1. (i) $\begin{pmatrix} -1 & 3 & 1 & 1 \\ 3 & -2 & 5 & 6 \\ 2 & 0 & 4 & 2 \end{pmatrix}$
- (ii) $\begin{pmatrix} -2 & 0 & 4 & 2 \\ -9 & 6 & -15 & -18 \\ -1 & 3 & 1 & 1 \end{pmatrix}$

$$(iii) \begin{pmatrix} 5 & -9 & 1 & -1 \\ 3 & -2 & 5 & -6 \\ -1 & 3 & 1 & 1 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 2 & 0 & 4 & 2 \\ 0 & -1 & 2 & 5 \\ -1 & 3 & 1 & 1 \end{pmatrix}$$

2. (i) $R_1 \leftrightarrow R_3$
(ii) $R_3 \rightarrow -2R_3$
(iii) $R_3 \rightarrow \frac{-1}{2}R_3$
(iv) $R_3 \rightarrow R_3 + \frac{1}{2}R_1$
(v) $R_2 \rightarrow R_2 - 2R_1$
3. The application of any row operation gives a matrix which is row equivalent to the given matrix. Thus 3 possible answers are:
- (i) Apply $R_2 \leftrightarrow R_3$

$$\begin{pmatrix} 4 & 3 & -1 & 5 \\ 2 & -3 & 0 & -5 \\ -4 & 2 & -11 & 0 \end{pmatrix}$$
- (ii) Apply $R_2 \rightarrow R_2 + R_1$

$$\begin{pmatrix} 4 & 3 & -1 & 5 \\ 0 & 5 & -12 & 5 \\ 2 & -3 & 0 & -5 \end{pmatrix}$$
- (iii) Apply $R_3 \rightarrow \frac{1}{2}R_3$
$$\begin{pmatrix} 4 & 3 & -1 & 5 \\ -4 & 2 & 0 & -5 \\ 1 & \frac{-3}{2} & 0 & \frac{-5}{2} \end{pmatrix}$$
4. (i) $(3k - 2, \frac{10k-7}{3}, k)$ where k is any real number
(ii) $(2, 1)$ is the solution
(iii) No solution
(iv) Unique solution $(-19, 12, 2)$
5. (i) $x = -1, y = 4, z = -3$
(iii) $x = 0, y = 0, z = 0$
(vii) Inconsistent
6. (i) $(\frac{19}{2} - 9k, \frac{-5}{2} + \frac{17}{4}k, 2 - \frac{3}{2}k, k)$ where k is any real number.
(ii) $(2, -1.3)$
(iii) $(-5 - 2k, 2 + 3k, 3 + 2k, k)$ where k is any real number.
(iv) $x = k, y = 1, z = 2 - k, w = k$
(v) $(-3k_1 + \frac{19}{4}k_2, \frac{-17}{8}k_2, k_1, k_2)$, where k_1, k_2 are arbitrary constants.
(vi) $(3k_2 - 2k_1, k_2, k_1)$, where k_1, k_2 are arbitrary constants.
(vii) No solution exists
7. (i)(a) $k = \pm 2\sqrt{2}$

- (b) $k \neq \pm 2\sqrt{2}$
 (c) None
 (ii)(a) $k = -2$
 (b) $k \neq \pm 2$
 (c) $k = 2$
 (iii)(a) $k = \pm\sqrt{3}$
 (b) $k \neq \pm\sqrt{3}$
 (c) None

$$8. \text{ (i) } \left[\begin{array}{cc|c} 2 & 3 & -7 \\ 1 & -1 & 9 \end{array} \right]; \left[\begin{array}{cc|c} 3 & 2 & 2 \\ 1 & -1 & 9 \end{array} \right]; \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & -5 \end{array} \right]$$

$$\text{(ii) } \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{array} \right]; \left[\begin{array}{ccc|c} 1 & 1 & 1 & -2 \\ 0 & 2 & 0 & 4 \\ 0 & 2 & 3 & 1 \end{array} \right]$$

9. $\lambda u + (1 - \lambda)v$, where $\lambda \in \mathbb{R}$ is a solution

10. *Hint:*

$x = 3, y = 3$ is a solution in \mathbb{Z}_5 . There is no solution in \mathbb{R} . Other solutions are $x = 0, y = 2; x = 1, y = 4; x = 2, y = 1; x = 4, y = 0$.

11. $x = \frac{1}{2}, y = 0$ is a solution in \mathbb{R} . Solutions in \mathbb{Z}_5 are : $x = 0, y = 4; x = 1, y = 1; x = 2, y = 3; x = 3, y = 0; x = 4, y = 2$.
 Note that in \mathbb{Z}_5 the two equations reduce to a single equation $2X + 4Y = 1$, as $8 \cong 3 \pmod{5}$.

Exercise - 10.7

2. Possible answers are:

$$\text{(i) } \left(\begin{array}{ccccc} 1 & -3 & 2 & 1 & 2 \\ 0 & 0 & 4 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right), \left(\begin{array}{ccccc} 1 & -3 & 6 & 0 & 5 \\ 0 & 0 & 4 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

$$\text{(iii) } \left(\begin{array}{cccc} 1 & 3 & -1 & 2 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 7 & 26 \\ 0 & 0 & 0 & 1 \end{array} \right), \left(\begin{array}{cccc} 1 & 3 & -1 & 2 \\ 0 & -1 & 2 & 3 \\ 0 & 0 & 7 & 26 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

$$\text{(iv) } \left(\begin{array}{cccc} 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 4 & -7 \\ 0 & 0 & 0 & 1 \end{array} \right), \left(\begin{array}{cccc} 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

$$\text{(v) } \left(\begin{array}{ccccc} 3 & 0 & 3 & 0 & 2 \\ 0 & 3 & -6 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right), \left(\begin{array}{ccccc} -3 & 0 & -3 & 0 & -2 \\ 0 & 3 & -6 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right)$$

$$(vi) \begin{pmatrix} 0 & 2 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 4 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Exercise - 10.12

1. (i) $u + v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $2u - v = \begin{pmatrix} -4 \\ 7 \end{pmatrix}$, $-2u + 1/2v = \begin{pmatrix} 3 \\ -6.5 \end{pmatrix}$
2. (i) (6, -2) where X direction is along east and Y direction is along north
(ii) (3, -2)
4. (i)
$$\begin{aligned} 2x_1 - 1.5x_2 &= 4 \\ -x_1 + 6x_2 &= -3 \\ 3x_1 &= -2.5 \end{aligned}$$

(ii)
$$\begin{aligned} -x_1 + 2x_2 - x_3 &= 0 \\ 2x_1 - 3x_2 &= 0 \\ 3x_1 + x_2 - 4x_3 &= 0 \end{aligned}$$

(iii)
$$\begin{aligned} x_1 - x_3 &= 0 \\ -2x_1 - x_2 + x_3 &= 0 \\ 4x_2 + 5x_3 &= 0 \\ 3x_1 + -6x_2 &= 0 \end{aligned}$$
5. (i) $xu_1 + yu_2 + zu_3 = b$ where $u_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 3 \\ 0 \\ 14 \end{pmatrix}$,
$$u_3 = \begin{pmatrix} -4 \\ 11 \\ 8 \end{pmatrix}, b = \begin{pmatrix} -1 \\ 10 \\ 23 \end{pmatrix}$$

(ii) $xu_1 + yu_2 + zu_3 = b$ where $u_1 = \begin{pmatrix} -1 \\ 0 \\ -3 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \end{pmatrix}$,
$$u_3 = \begin{pmatrix} -3 \\ 5 \\ 0 \\ -4 \end{pmatrix}, b = \begin{pmatrix} 12 \\ 6 \\ 0 \\ -1 \end{pmatrix}$$

(iii) $xu_1 + yu_2 + zu_3 + wu_4 = b$ where
$$u_1 = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, u_3 = \begin{pmatrix} 3 \\ -2 \\ -13 \end{pmatrix},$$

$$u_4 = \begin{pmatrix} -4 \\ 0 \\ 14 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 11 \\ 18 \end{pmatrix}$$

6. (i) Yes $b = -4u_1 - u_2$
(ii) Infinitely many ways are possible. Two possible ways are:
 $b = -2u_1 + u_2 + u_3 + u_4$
 $b = -2u_1 - u_2 + 0u_3 + 0u_4$
(iii) Infinitely many ways are possible. Two possible ways are:
 $b = -3u_1 + 4u_2 + 0u_3$
 $b = -4u_1 + 6u_2 - u_3$
(iv) NO
7. (i) No
(ii) Yes $b = 3/2u_1 - 1/2u_2$
(iii) Yes $b = -6u_1 + 3u_2 + 2u_3$
(iv) Yes. Infinitely many ways are possible. Two possible ways are
 $b = -u_1 - u_2 + 3u_3 + 0u_4$ $b = 0u_1 - 3u_2 + 4u_3 + u_4$
8. (i) Yes $b = 0u_1 + u_2 + u_3$
(ii) Yes $b = 2u_1 - u_2$
(iii) Yes. Infinitely many ways are possible. Two possible ways are:
 $b = 4u_1 + 5u_2 + u_3 + 0u_4 + 0u_5$
 $b = 3u_1 + u_2 + u_3 + u_4 + u_5$
(iv) No
9. $u_1 + u_2 = \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix}$, $u_1 - u_2 = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}$, $2u_1 + u_2 = \begin{pmatrix} 1 \\ -4 \\ 7 \end{pmatrix}$
Many more answers are possible.
10. $u_1 + u_2 + u_3 = \begin{pmatrix} -3 \\ 2 \\ 1 \\ 3 \end{pmatrix}$, $u_1 + u_2 - u_3 = \begin{pmatrix} 5 \\ 4 \\ -1 \\ 1 \end{pmatrix}$, $u_1 - u_2 = \begin{pmatrix} 9 \\ 3 \\ -2 \\ -4 \end{pmatrix}$,
 $-u_1 - u_2 - u_3 = \begin{pmatrix} 3 \\ -2 \\ -1 \\ -3 \end{pmatrix}$
Many more are possible
11. (i) Yes
(ii) Yes
(iii) Yes
13. (i) No, 3 vectors
(ii) No. Infinitely many.
(iii) $u_2 = 0u_1 + 1u_2 + 0u_3$ so $u_2 \in W$
(iv) Yes. $u_3 = 1/2u_1 + 1/2u_2$

14. (i) Yes
 (ii) Yes
 (iii) No

15. For all values of h and k

16. The echelon form of $\left(A \ :b \right)$ is

$$\begin{pmatrix} 1 & 0 & 0 & :1 \\ 0 & 1 & 0 & :2 \\ 0 & 0 & 0 & :1 \end{pmatrix}$$

Apply row operations to get a matrix A with non zero entries.

17. The echelon form of $\left(A \ :b \right)$ is

$$\begin{pmatrix} 1 & 0 & 2 & :1 \\ 0 & 1 & 2 & :1 \\ 0 & 0 & 0 & :1 \\ 0 & 0 & 0 & :0 \end{pmatrix}$$

Apply row operations to get a matrix A with non zero entries

18. (i) and (ii)
 19. (ii) and (iii)
 20. (i) and (iv)
 21. No

Exercise - 10.17

1. (i) No
 (ii) Yes
 (iii) Yes
 (iv) No
2. (i)

$$\begin{aligned} x_1 &= 4 + 5x_2 \\ x_2 &= 5 + 6x_3 \end{aligned}$$

x_3 is free.

- (ii) $x_1 = \frac{-5}{2} - \frac{3}{2}x_2 + 2x_3$ x_2, x_3 are free.

(iii)

$$\begin{aligned}x_1 &= -9 - 7x_3 \\x_2 &= 1 + 6x_3 + 2x_4 \\x_5 &= 0\end{aligned}$$

 x_3, x_4 are free.

(iv)

$$\begin{aligned}x_1 &= 1 \\x_2 &= \frac{2}{3} \\x_3 &= \frac{-2}{3}\end{aligned}$$

(v) No solution.

(vi)

(vii) $x_1 = x_4 - 2$

$x_2 = -1$

$x_3 = -2x_4 + 8$

 x_4 is free.

3. (i) Unique solution (1, 2, -1).

(ii) Inconsistent.

(iii) Infinitely many solutions.

(iv) Inconsistent.

4. (i)

$w = -1$

$x = -4$

$y = 2$

$z = -1$

(ii) Inconsistent.

(iii)

$w = 2$

$x = 0$

$y = 1$

$z = 3$

(iv)

$x = \frac{1}{2}z + \frac{5}{2}$

$y = \frac{1}{2}z + \frac{1}{2}$

z is free.

$$(v) \quad \begin{aligned} x &= 1 - 2z \\ y &= 3z - 1 \end{aligned}$$

z is free.

5. (i)

$$\begin{aligned} x_1 &= 2x_3 - 3x_4 - 24 \\ x_2 &= 2x_3 - 2x_4 - 7 \\ x_5 &= 4 \end{aligned}$$

x_3, x_4 are free.

(ii)

$$\begin{aligned} x_1 &= x_3 \\ x_2 &= -x_3 \end{aligned}$$

x_3 is free.

(iii) No solution.

(iv) Unique solution $(-1, 0, 1, 2)$

(v) Unique solution $(7, 6, -1)$

6. (i) $(-\frac{1}{12}k - \frac{7}{12}, \frac{23}{12} + \frac{5}{12}k, -\frac{5}{4} + \frac{1}{4}k, k)$ where k is a parameter.

(ii) $(1 - k, 2, 1, k)$ where k is a parameter.

(iii) $(5 + 3k_1, 1 + 4k_1, k_2, 4 - 9k_1, k_1)$ where k_1, k_2 are parameters.

7. $(x = \frac{1}{11}(7 - 16z), y = \frac{1}{11}(z + 3))$
 z is free.

8. Inconsistent.

9. (i)

$$\begin{aligned} x_1 &= 11x_2 \\ x_3 &= -7x_2 \end{aligned}$$

x_2 is free.

(ii) $x_1 = x_2 = x_3 = 0$

10. (i) $(5k, 6k, k)^t$ where k is any real number.

(ii) $(-k, 0, k)^t$ where k is any real number.

11. (i) Non-trivial.

(ii) Trivial.

12. (i) $3a + b - c = 0$.

(ii) $a + b - c = 0$.

(iii) For any a, b, c .

13. Consistent if and only if $(7b_1 + 2b_3 - 8b_4 = 0)$.
General solution is:

$$\begin{aligned}x_1 &= \frac{1}{8}b_1 - x_2 \\x_3 &= \frac{1}{8}(6b_3 - 16b_2 - 6b_1) \\x_5 &= \frac{1}{8}(-2b_3 + 8b_2 + 2b_1)\end{aligned}$$

14. Do yourself

15. (i) Consistent, as there is no row of the form $[0\ 0\ 0\ 0\ 0|b]$, with $b \neq 0$.
(ii) Inconsistent, as there is no row of the form $[0\ 0\ 0\ 0\ 0|b]$, with $b \neq 0$.
(iii) Consistent if 6th column is not a pivot column, inconsistent if 6th column is a pivot column.
(iv) Consistent as there is no row of the form $[0\ 0\ 0\ 0\ 0|b]$
(v) Consistent if the pivot is not in the 6th column. Inconsistent if the pivot is in the 6th column.

16. Do yourself

17. Consistent if the augmented column is not a pivot column. Unique solution when the coefficient matrix has a pivot in every row.

18. Here represents a pivot and * any real number.

(i)

$$\left(\begin{array}{cccc|c} r & * & * & * & * \\ 0 & 0 & r & * & * \\ 0 & 0 & 0 & 0 & r \end{array} \right)$$

(ii)

$$\left(\begin{array}{cccc|c} r & * & * & * & * \\ 0 & 0 & r & * & * \\ 0 & 0 & 0 & r & r \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

(iii)

$$\left(\begin{array}{ccc|c} r & * & * & * \\ 0 & 0 & r & * \end{array} \right)$$

(iv)

$$\left(\begin{array}{cc|c} r & * & * \\ 0 & r & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

Other solutions are also possible.

19. The echelon form of the augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & k-1 \\ 0 & 0 & 0 & k^2 - 3k + 2 \end{array} \right)$$

is inconsistent if $k \neq 1, 2$.

is consistent if $k = 1, 2$.

$k = 1$ solution is $x = 2z + 1, y = -3z$, z is free. $k = 2$ solution is $x = 2z, y = 1 - 3z$, z is free.

20. Echelon form is

$$\left(\begin{array}{ccc|c} 2 & 3k & 3k+4 & 0 \\ 0 & 3 & 2k+1 & 0 \\ 0 & 0 & k^2 - 4 & 0 \end{array} \right)$$

Unique trivial solution if $k \neq \pm 2$.

Infinitely many non-trivial solutions if $k = \pm 2$.

21. Echelon form is

(i) Unique solution for all values of d . It is
 $x = \frac{(d-b)(d-c)}{(a-b)(a-c)}, y = \frac{(d-a)(d-c)}{(b-a)(b-c)}, z = \frac{(d-b)(d-a)}{(b-c)(a-c)}$.

(ii) Inconsistent when $d \neq a$ or b .

consistent and infinitely many solutions when $d=a$ or $d=b$. Solution is:

If $c = a$, solution is $x = \frac{b-d}{b-a} - z, y = \frac{d-a}{b-a}$, z is free. If $c = b$, solution is $x = \frac{b-d}{b-a}, y = \frac{d-a}{b-a} - z$, z is free.

(iii) Inconsistent when $d \neq a$.

Consistent and infinitely many solutions when $d = a$. The solution is $x = 1 - y - z$, y is free, z is free.

Exercise - 10.19

1. (i) $X = [w \ x \ y \ z]^t = [-1 \ -4 \ 2 \ -1]^t$, a point in \mathbb{R}^4
 (ii) Null set
 (iii) $X = [2 \ 0 \ 1 \ 3]^t$
 (iv) $X = [x \ y \ z]^t = [\frac{5}{2} \ \frac{1}{2} \ 0]^t + z[1 \ 1 \ 2]^t, z \in \mathbb{R}$, a line in \mathbb{R}^3 through the point $[\frac{5}{2}, \frac{1}{2}, 0]^t$ and parallel to the vector $[1 \ 1 \ 2]^t$
 (v) $X = [x \ y \ z]^t = [-1 \ -1 \ 0]^t + z[-2 \ 3 \ 1]^t, z \in \mathbb{R}$, a line in \mathbb{R}^3 through the point $[-1 \ -1 \ 0]^t$ and parallel to the vector $[-2 \ 3 \ 1]^t$.
2. (i) $X = [x_1 \ x_2 \ x_3 \ x_4 \ x_5]^t = [-24 \ -7 \ 2 \ 0 \ 0 \ 4]^t + x_3[2 \ 2 \ 1 \ 0 \ 0] + x_4[-3 \ -2 \ 0 \ 1 \ 0]$ a plane in \mathbb{R}^5 through the point $[-24 \ -7 \ 0 \ 0 \ 4]^t$ and containing vectors $[2 \ 2 \ 1 \ 0 \ 0]^t, [-3 \ -2 \ 0 \ 1 \ 0]^t$
 (ii) $X = [x_1 \ x_2 \ x_3]^t = x_3[1 \ -1 \ 1]^t, x_3$ is a parameter, a line in \mathbb{R}^3 through origin, parallel to vector $[1 \ -1 \ 1]$.
 (iii) Null set

- (iv) $X = [-1 \ 0 \ 1 \ 2]^t$, a point in \mathbb{R}^4
- (v) $X = [7 \ 6 \ -1]^t$, a point in \mathbb{R}^3
3. (i) $X = [x_1 \ x_2 \ x_3 \ x_4]^t = [\frac{-7}{12} \ \frac{23}{12} \ -\frac{5}{4} \ 0]^t + x_4[\frac{-1}{12} \ \frac{5}{12} \ \frac{1}{4} \ 1]^t$, $x_4 \in \mathbb{R}$ it represents a line through point $[\frac{-7}{12} \ \frac{23}{12} \ -\frac{5}{4} \ 0]^t$ and parallel to vector $[\frac{-1}{12} \ \frac{5}{12} \ \frac{1}{4} \ 1]^t$
- (ii) $X = [x_1 \ x_2 \ x_3 \ x_4]^t = [1 \ 2 \ 1 \ 0]^t + x_4[-1 \ 0 \ 0 \ 1]^t$, x_4 is a parameter.
- (iii) $X = [x_1 \ x_2 \ x_3 \ x_4 \ x_5]^t = [5 \ 1 \ 0 \ 4 \ 0]^t + x_3[0 \ 0 \ 1 \ 0 \ 0]^t + x_5[3 \ 4 \ 0 \ -9 \ 1]^t$, x_3, x_5 are parametric; It represents a plane in \mathbb{R}^5 passing through point $[5 \ 1 \ 0 \ 4 \ 0]^t$ and containing two vectors $[0 \ 0 \ 1 \ 0 \ 0]^t$, $[3 \ 4 \ 0 \ -9 \ 1]^t$.
4. $X = [x \ y \ z]^t = [\frac{7}{11} \ \frac{3}{11} \ 0]^t + z[\frac{-16}{11} \ \frac{1}{11} \ 1]$, z is a parameter. It represents a line in \mathbb{R}^3 through the point $[\frac{7}{11} \ \frac{3}{11} \ 0]^t$ and parallel to vector $[\frac{-16}{11} \ \frac{1}{11} \ 1]$
5. Null set
6. (i) $X = [x_1 \ x_2 \ x_3]^t = x_2[11 \ 1 \ -7]^t$, x_2 is a parameter, a line in \mathbb{R}^3 passing through origin and parallel to vector $[11 \ 1 \ -7]^t$.
- (ii) $X = [0 \ 0 \ 0]^t$ a point, viz origin in \mathbb{R}^3 .

Chapter 11

Matrices

In this chapter we shall study matrices in detail. Apart from studying elementary operations on matrices, namely addition and multiplication, to find the product and inverse of large order matrices, the technique of partitioning is explained. Properties of special types of matrices have been discussed. We end with a study of eigen values and eigen vectors of square matrices.

11.1 Matrix of Numbers

We begin by defining a matrix of numbers as follows:

Definition 11.1. A $m \times n$ matrix A is a rectangular array of mn numbers arranged in m horizontal rows and n vertical columns:

$$\begin{pmatrix} a_{11} & a_{12} & \cdot & a_{1j} & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & a_{2j} & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdot & a_{ij} & \cdot & a_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdot & a_{mj} & \cdot & a_{mn} \end{pmatrix}$$

Notation Matrices are generally denoted by capital letters.

We write $A = (a_{ij})_{m \times n}$. The (i, j) th element of A is also denoted by $[A]_{ij}$

Note

1. The i th row is $(a_{i1} \quad a_{i2} \quad \cdot \quad \cdot \quad a_{in})$

2. The j th column is $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ a_{mj} \end{pmatrix}$

3. a_{ij} denotes the element at the intersection of the i th row and j th column. For instance a_{23} is the element in the 2nd row and 3rd column.

4. The diagonal elements are those for which the row suffix is equal to the column suffix, that is, $a_{11}, a_{22}, \dots, a_{nn}$.

Example 11.1.

1. $A = \begin{pmatrix} -1 & 2 & 3 \\ 0 & -2 & 1 \end{pmatrix}$ is a 2×3 matrix. It has 6 elements. The (1,2)th element is 2. The diagonal elements are $a_{11} = -1$, $a_{22} = -2$.

2. $B = \begin{pmatrix} -1 & 2 & 5 \\ 10 & & 3 \end{pmatrix}$ is not a matrix as there is no (2,2)th element.

Types of matrices

Matrices can be classified into various types.

On the basis of size

1. *Square matrix*

If the number of rows and columns of a matrix are equal, then the matrix is called a square matrix.

A $n \times n$ square matrix is also called a n -rowed square matrix or a square matrix of order n .

For example, $\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 & 1 \\ 2 & 3 & 4 \\ 5 & -1 & 6 \end{pmatrix}$ are square matrices of orders

2 and 3 respectively.

2. *Row matrix*

A matrix having only one row is called a row matrix or row vector.

For example, $(-3 \ 2)$, $(0 \ -1 \ 4)$ are 1×2 and 1×3 row matrices respectively.

3. *Column matrix*

A matrix having only one column is called a column matrix or a column vector.

For example, $\begin{pmatrix} -3 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$ are 2×1 and 3×1 column matrices

respectively.

On the Basis of Elements

4. *Diagonal matrix*

A square matrix having non-diagonal elements zero is called a diagonal matrix.

$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ are diagonal matrices.

$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ is not a diagonal matrix as the (2,1)th element is non-zero.

5. *Scalar matrix*

A diagonal matrix having all the diagonal elements equal is called a scalar

matrix. $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ is a scalar matrix. $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ is not a scalar matrix as $(1, 1)$ th element $= 3 \neq 4 = (2, 2)$ th element.

6. *Identity matrix (or unit matrix)*

A scalar matrix having each diagonal entry 1 is called an identity matrix or unit matrix.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is a unit matrix of order 2. A $n \times n$ unit matrix is denoted by I_n .

7. *Zero matrix (or Null matrix)*

A $m \times n$ matrix having all elements zero is called a zero matrix or null matrix. $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is a 2×3 null matrix. A $m \times n$ null matrix is written as $O_{m \times n}$ or simply O .

8. *Upper triangular matrix*

A $m \times n$ matrix is called an upper triangular matrix if all the entries below the diagonal are zero.

For example, $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 3 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$

9. *Lower triangular matrix*

A $m \times n$ matrix is called a lower triangular matrix if all the entries above the diagonal are zero.

For example, $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ -1 & 3 & 4 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & -1 \\ -1 & 1 & 1 \end{pmatrix}$

10. *Triangular matrix*

A matrix which is either upper triangular or lower triangular is called a triangular matrix.

11.2 Operations on Matrices

Definition 11.2. (Comparable matrices):

Let A be a $m \times n$ matrix and B a $p \times q$ matrix. A and B are said to be comparable matrices if $m = p$ and $n = q$, that is, they have the same number of rows and the same number of columns.

Definition 11.3. (Equality of matrices):

Let $A = (a_{ij})_{m \times n}$ and $B = (b_{ij})_{p \times q}$ be two matrices. Then $A = B$ if

(i) $m = p$ and $n = q$

(ii) $a_{ij} = b_{ij}$, for all $i = 1, \dots, m; j = 1, \dots, n$

condition (i) says that A and B are comparable matrices, whereas condition (ii), says that the corresponding elements are equal.

Example 11.2. 1. Let $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 4 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 4 & 6 \end{pmatrix}$. A and B are each 2×3 matrices and therefore they are comparable. Also their corresponding elements are equal, so that $A = B$.

2. Let $A = \begin{pmatrix} 1 & 4 \\ 3 & -6 \end{pmatrix}$, $B = (1 \ 4 \ 3 \ -6)$. A is a 2×2 matrix whereas B is a 1×4 matrix. Thus they are not comparable matrices so that $A \neq B$.

3. Let $A = \begin{pmatrix} -1 & 0 \\ 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 0 \\ 3 & -2 \end{pmatrix}$. A and B are comparable matrices as each is a 2×2 matrix. But $(2, 2)$ th element of $A = 2 \neq -2 = (2, 2)$ th element of B .

\therefore Corresponding elements of A and B are not equal, so that $A \neq B$.

11.2.1 Matrix Addition

Let us now see how we can add two matrices.

Definition 11.4. (Sum of two matrices):

Let $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$ be two comparable matrices. Their sum is the matrix $C = (c_{ij})_{m \times n}$, where $c_{ij} = a_{ij} + b_{ij}$, $i = 1, \dots, m$; $j = 1, \dots, n$. We write $C = A + B$.

Thus only comparable matrices can be added and their sum is a matrix of the same order, whose elements are obtained by adding the corresponding elements.

Example 11.3. 1. Let $A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 3 & -3 \end{pmatrix}$, $B = \begin{pmatrix} -4 & 1 & 3 \\ -3 & 2 & -1 \end{pmatrix}$. Since A and B are both 3×3 matrices, \therefore their sum $A + B$ is a 3×3 matrix given by

$$A + B = \begin{pmatrix} -3 & 0 & 5 \\ -5 & 5 & -4 \end{pmatrix}$$

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Since A is a 3×2 matrix and B is a 2×3 matrix, therefore they are not comparable matrices. Hence they cannot be added.

The following theorem gives the properties of matrix addition.

Theorem 11.1. Let A , B and C be $m \times n$ matrices. Then

- (i) $A + (B + C) = (A + B) + C$ (Associativity)
- (ii) $A + B = B + A$ (Commutativity)
- (iii) If O is the $m \times n$ null matrix, then $A + O = O + A = A$
- (iv) If A is any matrix, then there exists a matrix B such that $A + B = B + A = O$. B is called an additive inverse of A .

Proof: Let $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$ and $C = (c_{ij})_{m \times n}$.

(i) Since A , B , C are $m \times n$ matrices

$\therefore A + B$, $B + C$ are also $m \times n$ matrices. Consequently $(A + B) + C$ and $A + (B + C)$ are $m \times n$ matrices and therefore comparable.

We now show that their corresponding elements are equal.

For $i = 1, \dots, m$; $j = 1, \dots, n$;

$$\begin{aligned}
[(A + B) + C]_{ij} &= [A + B]_{ij} + [C]_{ij} \\
&= (a_{ij} + b_{ij}) + c_{ij} \\
&= a_{ij} + (b_{ij} + c_{ij}) \\
&= [A]_{ij} + [B + C]_{ij} \\
&= [A + (B + C)]_{ij}
\end{aligned}$$

Hence $(A + B) + C = A + (B + C)$

(ii) Clearly $A + B$, $B + A$ are comparable matrices each being of type $m \times n$.

For $i = 1, \dots, m; j = 1, \dots, n$

$$\begin{aligned}
[A + B]_{ij} &= a_{ij} + b_{ij} \\
&= b_{ij} + a_{ij} \\
&= [B + A]_{ij}
\end{aligned}$$

Hence $A + B = B + A$.

(iii) Since $A + O$ is a $m \times n$ matrix, therefore $A + O$ and A are comparable. For

$i = 1, \dots, m; j = 1, \dots, n$

$$a_{ij} = a_{ij} + 0$$

$\therefore (i, j)$ th element of $A = (i, j)$ th element of $A + O$.

Hence $A = A + O$

Using (ii), $A + O = O + A = A$.

(iv) Let $A = (a_{ij})_{m \times n}$ be the given matrix. Define $B = (b_{ij})_{m \times n}$, such that

$$b_{ij} = -a_{ij}.$$

Then $A + B$ and the null matrix O are both $m \times n$ matrices.

$$\begin{aligned}
\text{Also } [A + B]_{ij} &= a_{ij} + b_{ij} \\
&= a_{ij} + (-a_{ij}) \\
&= 0 \\
&= (i, j)\text{th element of } O
\end{aligned}$$

$$\therefore A + B = O$$

Using (ii) we get $A + B = B + A = O$. □

The additive inverse B of A is usually denoted by $-A$. Thus $-A = (-a_{ij})_{m \times n}$. The null matrix plays the role of the number O .

Now we give another type of operation on a matrix namely multiplication of a matrix by a scalar.

Definition 11.5. (Scalar multiplication):

Let $A = (a_{ij})_{m \times n}$ and k be any complex number. Then, the matrix $kA = (ka_{ij})_{m \times n}$ is called the scalar multiple of A by k .

Example 11.4. Let $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & -1 & 4 \end{pmatrix}$, then $2A = \begin{pmatrix} 2 & 4 & -6 \\ 4 & -2 & 8 \end{pmatrix}$,
 $-A = \begin{pmatrix} -1 & -2 & 3 \\ -2 & 1 & -4 \end{pmatrix}$

The following theorem gives the properties of scalar multiplication.

Theorem 11.2. Let A and B be $m \times n$ matrices and k, l any complex numbers.

Then

- (i) $k(A + B) = kA + kB$
- (ii) $(k + l)A = kA + lA$
- (iii) $(kl)A = k(lA)$
- (iv) $1A = A$

Proof: Let $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$. Then $A + B$, kA , kB are $m \times n$ matrices so that $k(A+B)$ and $kA+kB$ are also $m \times n$ matrices. For $i = 1, \dots, m$;
 $j = 1, \dots, n$

$$\begin{aligned} [k(A+B)]_{ij} &= k(i, j)\text{th element of } (A+B) \\ &= k(a_{ij} + b_{ij}) \\ &= ka_{ij} + kb_{ij} \\ &= [kA]_{ij} + [kB]_{ij} \\ &= [kA + kB]_{ij} \end{aligned}$$

Hence $k(A+B) = kA + kB$.

Proof of the other parts are left to the reader. \square

11.2.2 Matrix Multiplication

As we have seen that only matrices of the same order are added and the sum is a matrix of the same order. But when multiplying two matrices, their orders can be different and the order of the product can also be different from the order of either of the factors.

To multiply two matrices, we proceed as follows:

Definition 11.6. (Product of vectors):

Let $A = (a_1 \ a_2 \ \dots \ a_n)$ be a $1 \times n$ row vector, and $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ be a $n \times 1$ column vector. The product AB is a 1×1 matrix (or just a number) defined

$$\begin{aligned} \text{by } AB &= (a_1 \ a_2 \ \dots \ a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = (a_1b_1 + a_2b_2 + \dots + a_nb_n) \\ &= (\sum_{i=1}^n a_ib_i) \end{aligned}$$

The number of elements in A is equal to the number of elements in B so that component-wise multiplication is possible.

Definition 11.7. Let $A = (a_{ij})_{m \times n}$ and $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ be a $n \times 1$ column vector.

Write $A = (C_1 \ C_2 \ \dots \ C_n)$, where C_j is the j th column of A .

$$\begin{aligned} \text{Then, define } AB &= (C_1 \ C_2 \ \dots \ C_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\ &= C_1b_1 + C_2b_2 + \dots + C_nb_n. \end{aligned}$$

Definition 11.8. Let $A = (a_{ij})_{m \times n}$ and $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ be a $n \times 1$ column vector.

Write $A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}$, where R_i is the i th row of A .

Define, $AB = \begin{pmatrix} R_1 B \\ R_2 B \\ \vdots \\ R_m B \end{pmatrix}$.

Definition 11.9. Let $A = (a_{ij}) = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}$ be a $m \times n$ matrix with rows

R_1, R_2, \dots, R_m and $B = (b_{ij}) = (C_1 \ C_2 \ \dots \ C_p)$ be a $n \times p$ matrix with columns C_1, C_2, \dots, C_p .

The product AB is a $m \times p$ matrix given by

$$AB = (AC_1 \ AC_2 \ \dots \ AC_p) = \begin{pmatrix} R_1 C_1 & R_1 C_2 & \dots & R_1 C_p \\ R_2 C_1 & R_2 C_2 & \dots & R_2 C_p \\ \vdots & \vdots & \ddots & \vdots \\ R_m C_1 & R_m C_2 & \dots & R_m C_p \end{pmatrix}$$

$$\begin{aligned} [AB]_{ij} &= R_i C_j \\ &= a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} \\ &= \sum_{k=1}^n a_{ik}b_{kj}. \end{aligned}$$

In the product AB , A is called the pre-factor and B is called the post-factor. Two matrices can be multiplied only when the number of columns of the pre-factor = number of rows of the post-factor.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \hline a_{i1} & a_{i2} & \dots & a_{in} \\ \hline \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ b_{21} & \dots & b_{2j} & \dots & b_{2p} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{pmatrix}$$

(i, j) th element of AB i.e.

To obtain the $[AB]_{ij}$, multiply the i th row of the pre-factor A by the j th column of the post-factor B .

Example 11.5. If $A = \begin{pmatrix} 1 & 2 & -1 \\ -2 & 3 & 1 \\ 2 & 0 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & 3 \\ 3 & 4 & -1 & 0 \end{pmatrix}$ find AB .

If we write $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = (a_{ij})_{3 \times 3}$, $B = (b_1 \ b_2 \ b_3 \ b_4) = (b_{ij})_{3 \times 4}$

$$\begin{aligned} \text{then } AB &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} (b_1 \ b_2 \ b_3 \ b_4) \\ &= \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 & a_1 b_4 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 & a_2 b_4 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 & a_3 b_4 \end{pmatrix} \text{ where } a_i b_j = a_{i1} b_{1j} + a_{i2} b_{2j} + a_{i3} b_{3j} \\ &= \begin{pmatrix} -2 & -1 & 3 & 7 \\ 1 & 12 & 2 & 7 \\ 14 & 14 & -4 & 2 \end{pmatrix} \end{aligned}$$

If A is a 2×3 matrix and B is a 3×4 matrix, then AB is defined and is a 2×4 matrix but the product BA is not even defined as the number of columns in $B = 4 \neq 3 =$ number of rows in A .

If C is a 3×2 matrix then AC is a 2×2 matrix whereas CA is a 3×3 matrix, so that AC and CA cannot be equal. But if A and B are both square matrices of the same order say n , then AB and BA are both square matrices of order n . Will they be equal? The answer is no, as is shown by the following example.

Example 11.6. Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

Then $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

so that $AB \neq BA$

Thus matrix multiplication is not commutative. In this way matrices behave differently from numbers. In the above example we also see that $A \neq O$, $B \neq O$ but $AB = O$ i. e. the product of two non-zero matrices can be the zero matrix. This is a unique property of matrices which is not possessed by numbers.

Theorem 11.3. Let A, B, C be matrices of suitable sizes for which the matrix operations below can be defined, and let k be any complex number. Then

- (i) $(AB)C = A(BC)$ (Associative law)
- (ii) $A(B + C) = AB + AC$, $(A + B)C = AC + BC$ (Distributive law)
- (iii) $I_n A = A I_n = A$ (Identity)
- (iv) $k(AB) = (kA)B = A(kB)$.

Proof:

- (i) Let $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times p}$, $C = (c_{ij})_{p \times r}$

Then AB is a $m \times p$ matrix and $(AB)C$ is a $m \times r$ matrix. $A(BC)$ is also a $m \times r$ matrix so that $(AB)C$ and $A(BC)$ are comparable matrices. Also,

$$[AB]_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \quad \dots (1)$$

$$[BC]_{ij} = \sum_{s=1}^p b_{is} c_{sj} \quad \dots (2)$$

$$\begin{aligned} [(AB)C]_{ij} &= \sum_{l=1}^p [AB]_{il} C_{lj} \\ &= \sum_{l=1}^p \left(\sum_{k=1}^n a_{ik} b_{kl} \right) c_{lj} \quad \text{using (1)} \\ &= \sum_{l=1}^p \sum_{k=1}^n a_{ik} b_{kl} c_{lj} \quad \dots (3) \end{aligned}$$

$$\begin{aligned}
[A(BC)]_{ij} &= \sum_{k=1}^n a_{ik}[BC]_{kj} \\
&= \sum_{k=1}^n a_{ik} \sum_{s=1}^p b_{ks}c_{sj} && \text{using (2)} \\
&= \sum_{k=1}^n \sum_{s=1}^p a_{ik}b_{ks}c_{sj} \\
&= \sum_{s=1}^p \sum_{k=1}^n a_{ik}b_{ks}c_{sj} \quad \dots (4)
\end{aligned}$$

From (3) and (4) we get $[(AB)C]_{ij} = [A(BC)]_{ij}$

So that $(AB)C = A(BC)$

(ii) Let $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times p}$, $C = (c_{ij})_{n \times p}$ so that $B + C$ exists. It is a $n \times p$ matrix.

$\therefore A(B + C)$ is a $m \times p$ matrix. Also $AB + AC$ is a $m \times p$ matrix. Thus $A(B + C)$ and $AB + AC$ are comparable matrices.

$$\begin{aligned}
[A(B + C)]_{ij} &= \sum_{k=1}^n a_{ik}[B + C]_{kj} \\
&= \sum_{k=1}^n a_{ik}(b_{kj} + c_{kj}) \\
&= \sum_{k=1}^n (a_{ik}b_{kj} + a_{ik}c_{kj}) \\
&= \sum_{k=1}^n a_{ik}b_{kj} + \sum_{k=1}^n a_{ik}c_{kj} \\
&= [AB]_{ij} + [AC]_{ij} \\
&= [AB + AC]_{ij}
\end{aligned}$$

Hence $A(B + C) = AB + AC$

Similarly, it can be proved that $(A + B)C = AC + BC$

The proofs of the other parts can be shown by direct computation of each entry on both sides of the equalities, and are left to the reader. \square

The unit matrix I_n plays the same role as the number 1 does in the set of numbers.

Definition 11.10. (Transpose of a matrix):

Let $A = (a_{ij})$ be a $m \times n$ matrix. The $n \times m$ matrix $B = (b_{ij})$ defined such that $b_{ij} = a_{ji}$, that is (i, j) th element of $B = (j, i)$ th element of A , is called the transpose of A . It is denoted by A^T or A' or A^t . We shall use A^t .

Example 11.7. If $A = \begin{pmatrix} 2 & 1 & 3 & 4 \\ -5 & 0 & 1 & 2 \\ 2 & -3 & -4 & 0 \end{pmatrix}$, then $A^t = \begin{pmatrix} 2 & -5 & 2 \\ 1 & 0 & -3 \\ 3 & 1 & -4 \\ 4 & 2 & 0 \end{pmatrix}$

Theorem 11.4. If A and B are matrices of suitable sizes, then

- (i) $(A^t)^t = A$
- (ii) $(A + B)^t = A^t + B^t$
- (iii) $(kA)^t = kA^t$, where k is a complex number
- (iv) $(AB)^t = B^tA^t$ (Reversal law for transpose).

Proof: The proofs of (i) to (iii) are simple. We shall prove (iv).

(iv) Let $A = (a_{ij})_{m \times n}$ and $B = (b_{ij})_{n \times p}$. Then AB is a $m \times p$ matrix so that $(AB)^t$ is a $p \times m$ matrix. A^t , B^t are $n \times m$ and $p \times n$ matrices respectively, so that B^tA^t is a $p \times m$ matrix. Hence $(AB)^t$ and B^tA^t are comparable matrices. We now prove that their corresponding elements are equal.

$$\begin{aligned}
\text{If } AB &= (c_{ij})_{m \times p}, \text{ then } c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}, \text{ for } i = 1, \dots, p; j = 1, \dots, m \\
(i, j)\text{th element of } (AB)^t &= \sum_{k=1}^n ((i, k)\text{th element of } B^t)((k, j)\text{th element of } A^t) \\
&= \sum_{k=1}^n b_{ki}a_{jk} \\
&= \sum_{k=1}^n a_{jk}b_{ki}
\end{aligned}$$

$$= (j, i)\text{th element of } AB$$

$$= (i, j)\text{th element of } (AB)^t$$

Thus the corresponding elements of $B^t A^t$ and $(AB)^t$ are equal, so that $(AB)^t = B^t A^t$. \square

Definition 11.11. (Conjugate of a matrix):

Let A be an $m \times n$ matrix. The $m \times n$ matrix B obtained by taking the complex conjugate of each element of A is called the conjugate of A . It is denoted by \overline{A} . Symbolically, if $A = (a_{ij})_{m \times n}$, then $\overline{A} = (\overline{a_{ij}})_{m \times n}$. If a matrix has real entries then $\overline{A} = A$.

Example 11.8. If $A = \begin{pmatrix} 2+3i & -4i \\ 0 & 5 \end{pmatrix}$,

then $\overline{A} = \begin{pmatrix} \overline{2+3i} & \overline{-4i} \\ \overline{0} & \overline{5} \end{pmatrix} = \begin{pmatrix} 2-3i & 4i \\ 0 & 5 \end{pmatrix}$.

Theorem 11.5. Let A and B be matrices of suitable sizes. Then

- (i) $\overline{(\overline{A})} = A$
- (ii) $\overline{(A+B)} = \overline{A} + \overline{B}$
- (iii) $\overline{kA} = \overline{k} \overline{A}$, where k is a complex number
- (iv) $\overline{AB} = \overline{A} \overline{B}$
- (v) $\overline{(A)^t} = (\overline{A})^t$.

Proof left to the reader. \square

Definition 11.12. (Tranjugate of a matrix):

Let A be a $m \times n$ matrix. The $n \times m$ matrix B obtained by taking the conjugate of the transpose of A is called the transposed conjugate or tranjugate of A . It is denoted by A^θ . Symbolically, if $A = (a_{ij})_{m \times n}$, then $A^\theta = (\overline{a_{ji}})_{n \times m} = \overline{(A^t)}$.

Example 11.9. Let $A = \begin{pmatrix} 1+2i & 3i & 4 \\ -5i & -6 & -1+i \end{pmatrix}$, $A^t = \begin{pmatrix} 1+2i & -5i \\ 3i & -6 \\ 4 & -1+i \end{pmatrix}$,

$A^\theta = \overline{(A^t)} = \begin{pmatrix} 1-2i & 5i \\ -3i & -6 \\ 4 & -1-i \end{pmatrix}$.

Theorem 11.6. Let A and B be matrices of suitable sizes. Then

- (i) $(A^\theta)^\theta = A$
- (ii) $(A+B)^\theta = A^\theta + B^\theta$
- (iii) $(kA)^\theta = \overline{k} A^\theta$, where k is a complex number
- (iv) $(AB)^\theta = B^\theta A^\theta$.

Proof:

(i) $A^\theta = \overline{(A^t)} = (\overline{A})^t$
 $(A^\theta)^t = ((\overline{A})^t)^t = \overline{A}$
 $(A^\theta)^\theta = \overline{(A^\theta)^t} = \overline{\overline{A}} = A$
 $\therefore (A^\theta)^\theta = A$.

(ii) Let A and B be matrices of the same order.

$$(A+B)^\theta = \overline{(A+B)^t}$$

$$= \overline{A^t + B^t}$$

$$= \overline{A^t} + \overline{B^t}$$

$$= A^\theta + B^\theta.$$

$$\begin{aligned}
\text{(iii)} \quad (kA)^\theta &= \overline{(kA)^t} \\
&= \overline{kA^t} \\
&= \overline{k(A^t)} \\
&= \overline{k}A^\theta \\
\therefore (kA)^\theta &= \overline{k}A^\theta.
\end{aligned}$$

(iv) Let A and B be matrices such that AB exists. Then

$$\begin{aligned}
(AB)^\theta &= \overline{(AB)^t} \\
&= \overline{B^t A^t} \text{ using reversal law for transposes} \\
&= \overline{B^t} \overline{A^t} \\
&= B^\theta A^\theta \\
\therefore (AB)^\theta &= B^\theta A^\theta. \quad \square
\end{aligned}$$

Definition 11.13. (Trace of a matrix):

Let A be a square matrix of order n . The sum of the diagonal elements of A is called the trace of A . It is denoted by $tr A$. Hence if $A = (a_{ij})_{m \times n}$, then $tr(A) = a_{11} + a_{22} + \cdots + a_{nn}$.

Theorem 11.7. If A and B are square matrices of order n , and λ is a scalar, then

- (i) $tr(A + B) = tr(A) + tr(B)$
- (ii) $tr(kA) = k tr(A)$
- (iii) $tr(AB) = tr(BA)$.

Proof: Let $A = (a_{ij})_{n \times n}$ and $B = (b_{ij})_{n \times n}$. Then $tr(A) = \sum_{i=1}^n a_{ii}$, $tr(B) = \sum_{i=1}^n b_{ii}$

$$\begin{aligned}
\text{(i)} \quad A + B &= (a_{ij} + b_{ij})_{n \times n} \\
\therefore tr(A + B) &= \sum_{i=1}^n (a_{ii} + b_{ii}) \\
&= \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} \\
&= tr(A) + tr(B).
\end{aligned}$$

$$\begin{aligned}
\text{(ii)} \quad kA &= (ka_{ij})_{n \times n} \\
\therefore tr(kA) &= \sum_{i=1}^n ka_{ii} \\
&= k \sum_{i=1}^n a_{ii} \\
&= k tr(A).
\end{aligned}$$

$$\begin{aligned}
\text{(iii)} \quad \text{If } AB &= (c_{ij})_{n \times n} \text{ and } BA = (d_{ij})_{n \times n} \text{ then } c_{ij} = \sum_{k=1}^n a_{ik}b_{kj} \text{ and } d_{ij} = \\
&\sum_{k=1}^n b_{ik}a_{kj} \\
tr(AB) &= \sum_{i=1}^n c_{ii} \\
&= \sum_{i=1}^n \left(\sum_{k=1}^n a_{ik}b_{ki} \right) \\
&= \sum_{i=1}^n \sum_{k=1}^n b_{ki}a_{ik} \\
&= \sum_{k=1}^n \sum_{i=1}^n b_{ki}a_{ik}, \text{ interchanging the order of summation.} \\
&= \sum_{k=1}^n d_{kk} \\
&= tr(BA). \quad \square
\end{aligned}$$

11.3 Partitioning of Matrices

Matrices of large orders often arise in practical problems and it is required to find the inverse of these matrices. It is convenient to partition such matrices to find the inverse in order to speed up the calculation.

A partition of a matrix is done by drawing lines parallel to rows and columns. For example,

$$\text{If } A = \begin{pmatrix} 1 & 2 & -1 & 3 & 4 & 5 \\ 2 & 1 & 8 & 3 & 4 & 0 \\ -1 & 0 & 2 & 1 & -1 & 3 \\ 4 & 8 & 9 & 5 & 6 & 2 \\ 1 & 1 & 0 & 1 & 2 & 0 \\ 0 & -1 & 1 & -2 & -3 & 4 \\ 4 & 8 & 9 & 5 & 1 & 3 \end{pmatrix} \text{ then partition of } A \text{ can be}$$

$$A = \left(\begin{array}{cc|ccc} 1 & 2 & -1 & 3 & 4 & 5 \\ 2 & 1 & 8 & 3 & 4 & 0 \\ \hline -1 & 0 & 2 & 1 & -1 & 3 \\ 4 & 8 & 9 & 5 & 6 & 2 \\ 1 & 1 & 0 & 1 & 2 & 0 \\ \hline 0 & -1 & 1 & -2 & -3 & 4 \\ 4 & 8 & 9 & 5 & 1 & 3 \end{array} \right).$$

Thus the matrix A can be written as:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} \text{ where } A_{11}, A_{12}, A_{21}, A_{22}, A_{31} \text{ and } A_{32} \text{ are sub-} \\ \text{matrices of } A \text{ of order } 2 \times 2, 2 \times 4, 3 \times 2, 3 \times 4, 2 \times 2, 2 \times 4 \text{ respectively.}$$

Example 11.10. Given a 7×8 matrix A , partition it in the form $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ so that A_{11} is a 3×2 matrix. What are the sizes of A_{12} , A_{21} and A_{22} ? Since A_{11} is a 3×2 matrix, $\therefore A_{12}$ is a $3 \times (8 - 2)$ i.e. 3×6 matrix. A_{21} is a $(7 - 3) \times 2$ i.e. 4×2 matrix. A_{22} is a $(7 - 3) \times (8 - 2)$ i.e. 4×6 matrix.

Addition and Scalar Multiplication of Partitioned Matrices

Suppose A and B are two matrices conformable to addition. A and B are partitioned as $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ so that for each submatrix A_{ij} , $i = 1, 2; j = 1, 2$ the corresponding submatrix B_{ij} is of the same order. This will be so provided A and B are partitioned in precisely the same way.

$$\text{Then } A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}.$$

This also provides a partition of $A + B$. Scalar multiplication of a partitioned matrix is obtained by taking the scalar multiple of each block.

$$\text{Thus } kA = \begin{pmatrix} kA_{11} & kA_{12} \\ kA_{21} & kA_{22} \end{pmatrix}.$$

Example 11.11. Let $A = \begin{pmatrix} 1 & 2 & -1 & 3 & 4 \\ 2 & 3 & 0 & 1 & -2 \\ 4 & -6 & 5 & 8 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 4 & 5 & 3 & 2 \\ 0 & 1 & -1 & 1 & 3 \\ -3 & 8 & 0 & -7 & 4 \end{pmatrix}$
 A and B are of the same order, so they can be added. Partitioning A and B in the same way, let $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, where $A_{11} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \end{pmatrix}$,
 $A_{12} = \begin{pmatrix} 3 & 4 \\ 1 & -2 \end{pmatrix}$, $A_{21} = \begin{pmatrix} 4 & -6 & 5 \end{pmatrix}$, $A_{22} = \begin{pmatrix} 8 & 1 \end{pmatrix}$

and $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, where $B_{11} = \begin{pmatrix} -1 & 4 & 5 \\ 0 & 1 & -1 \end{pmatrix}$, $B_{12} = \begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix}$,
 $B_{21} = \begin{pmatrix} -3 & 8 & 0 \end{pmatrix}$, $B_{22} = \begin{pmatrix} -7 & 4 \end{pmatrix}$

Then $A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix} = \left(\begin{array}{ccc|cc} 0 & 6 & 4 & 6 & 6 \\ 2 & 4 & -1 & 2 & 1 \\ \hline 1 & 2 & 5 & 1 & 5 \end{array} \right)$

Also $2A = \begin{pmatrix} 2A_{11} & 2A_{12} \\ 2A_{21} & 2A_{22} \end{pmatrix} = \left(\begin{array}{ccc|cc} 2 & 4 & -2 & 6 & 8 \\ 4 & 6 & 0 & 2 & -4 \\ \hline 8 & -12 & 10 & 16 & 2 \end{array} \right)$.

Example 11.12. Let $A = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & -1 & 0 & 2 \\ \hline -1 & -2 & -3 & -4 \end{array} \right)$,

$B = \left(\begin{array}{cc|cc} -1 & 2 & -3 & 4 \\ -4 & 2 & 3 & -1 \\ \hline 0 & 3 & 2 & 5 \end{array} \right)$

be matrices each of order 3×4 . Being of the same order, they can be added. Suppose A and B are partitioned as shown by the lines. Since they are partitioned in the same way, their sum is obtained by adding the corresponding submatrices.

Thus $A + B = \left(\begin{array}{cc|cc} 0 & 4 & 0 & 8 \\ 1 & 1 & 3 & 1 \\ \hline -1 & 1 & -1 & 1 \end{array} \right)$.

$-4A = \left(\begin{array}{cc|cc} -4 & -8 & -12 & -16 \\ -20 & 4 & 0 & -8 \\ \hline 4 & 8 & 12 & 16 \end{array} \right)$.

11.3.1 Multiplication of Partitioned Matrices

The multiplication of partitioned matrices can be done in the usual way as if the block entries are scalars. To obtain AB , the partition of A and B have to be done in such a way that the column partition of A matches the row partition of B .

Example 11.13. Let $A = \begin{pmatrix} 1 & 2 & -1 & 0 & 1 & 3 \\ 2 & 1 & 4 & 1 & 1 & -1 \\ 1 & 3 & -1 & 2 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 & 4 & 1 \\ -1 & 0 & 1 & 3 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & 2 & 1 \end{pmatrix}$

A partition of A and B is shown by lines

$A = \left(\begin{array}{cc|cccc} 1 & 2 & -1 & 0 & 1 & 3 \\ 2 & 1 & 4 & 1 & 1 & -1 \\ \hline 1 & 3 & -1 & 2 & 1 & 1 \end{array} \right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$

$B = \left(\begin{array}{cc|cc} 1 & -1 & 4 & 1 \\ -1 & 0 & 1 & 3 \\ \hline 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & 2 & 1 \end{array} \right) = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$

$$\begin{aligned}
\text{Then } AB &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\
&= \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} \\
&= \left(\begin{array}{cc|cc} 0 & 2 & 11 & 9 \\ 9 & 4 & 7 & 8 \\ \hline -1 & 4 & 10 & 10 \end{array} \right)
\end{aligned}$$

Considering another partition of A and B , namely

$$\begin{aligned}
A &= \left(\begin{array}{cc|cc|cc} 1 & 2 & -1 & 0 & 1 & 3 \\ 2 & 1 & 4 & 1 & 1 & -1 \\ 1 & 3 & -1 & 2 & 1 & 1 \end{array} \right) = \begin{pmatrix} A'_{11} & A'_{12} & A'_{13} \\ A'_{21} & A'_{22} & A'_{23} \end{pmatrix} \\
B &= \left(\begin{array}{c|ccc} 1 & -1 & 4 & 1 \\ \hline -1 & 0 & 1 & 3 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ \hline 0 & 1 & -1 & 0 \\ 1 & 1 & 2 & 1 \end{array} \right) = \begin{pmatrix} B'_{11} & B'_{12} \\ B'_{21} & B'_{22} \\ B'_{31} & B'_{32} \end{pmatrix} \\
\therefore AB &= \begin{pmatrix} A'_{11}B'_{11} + A'_{12}B'_{21} + A'_{13}B'_{31} & A'_{11}B'_{12} + A'_{12}B'_{22} + A'_{13}B'_{32} \\ A'_{21}B'_{11} + A'_{22}B'_{21} + A'_{23}B'_{31} & A'_{21}B'_{12} + A'_{22}B'_{22} + A'_{23}B'_{32} \end{pmatrix} \\
&= \left(\begin{array}{c|ccc} 0 & 2 & 11 & 9 \\ \hline 9 & 4 & 7 & 8 \\ -1 & 4 & 10 & 10 \end{array} \right).
\end{aligned}$$

Thus, by considering different partitions of A and B , different partitions of AB will be obtained. However, the matrix AB remains the same.

In the above example corresponding to a partition of A , the matrix B is partitioned, so that the various products are defined. We now explain this in detail.

Let A be a $m \times n$ matrix and B a $n \times p$ matrix over the same field F . Let A be partitioned in any manner. Corresponding to this partition of A , partition B is as follows:

To each partition line of A parallel to the columns, associate a partition line of B parallel to its rows such that the number of rows of B between two adjacent partition lines is the same as the number of columns of A between the corresponding adjacent partition lines.

Two matrices A and B which are conformable for multiplication and partitioned in the manner described above are said to be conformably partitioned for multiplication. With such partitionings, it is possible to multiply the two matrices in the usual manner, as if the submatrices are elements. We have the following theorem.

Theorem 11.8. *Let A and B be two matrices which are conformable for multiplication. Suppose A and B are conformably partitioned as $A = (A_{ij})_{u \times v}$, $B = (B_{jk})_{v \times w}$. If the product $C = AB$ is partitioned according to the row partition of A and the column partition of B , so that $C = (C_{ik})_{u \times w}$ then $C_{ik} = \sum_{j=1}^v A_{ij}B_{jk}$*

Proof: We will prove the result for the case $u = v = w = 2$ only.

Let $A = (a_{ij})_{m \times n}$ and $B = (b_{ij})_{n \times p}$ are two matrices. Suppose A and B are

conformably partitioned as follows:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

If A_{11} is a $s \times t$ matrix then A_{12} , A_{21} and A_{22} are $s \times (n-t)$, $(m-s) \times t$ and $(m-s) \times (n-t)$ matrices respectively.

Since the partitions of A and B are conformable for multiplication.

\therefore Number of rows of B_{11} = number of columns of $A_{11} = t$.

Let number of columns of $B_{11} = k$ then B_{11} , B_{12} , B_{21} , B_{22} are $t \times k$, $t \times (p-k)$, $(n-t) \times k$, $(n-t) \times (p-k)$ matrices respectively. (i, j) th element of $AB = \sum_{l=1}^n a_{il}b_{lj}$
 $= \sum_{l=1}^t a_{il}b_{lj} + \sum_{l=t+1}^n a_{il}b_{lj}$

The range of l from 1 to n has been split into 2 ranges, one from 1 to t and the other from $t+1$ to n . Hence each row of A as well as each column of B has been broken up into two parts. Thus we write $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = (E_1 \ E_2)$

(say), where $E_1 = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix}$, $E_2 = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}$

Thus E_1 is a $m \times t$ matrix and E_2 a $m \times (n-t)$ matrix

and $B = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$, where $F_1 = (B_{11} \ B_{12})$, $F_2 = (B_{21} \ B_{22})$ F_1 is a $t \times p$, matrix and F_2 is a $(n-t) \times p$ matrix.

Then from the definition of multiplication

$$AB = (E_1 \ E_2) \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$$

$\therefore AB = E_1F_1 + E_2F_2$

$$E_1F_1 = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix} (B_{11} \ B_{12}) = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{21}B_{11} & A_{21}B_{12} \end{pmatrix}$$

$$E_2F_2 = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} (B_{21} \ B_{22}) = \begin{pmatrix} A_{12}B_{21} & A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}$$

Hence $AB = E_1F_1 + E_2F_2$. □

11.4 Special Types of Matrices

In many practical problems we come across matrices having special forms. We now study these special types of matrices.

Symmetric and Skew Symmetric Matrices

Definition 11.14. A square matrix $A = (a_{ij})_{n \times n}$ is said to be symmetric if $a_{ij} = a_{ji}$, for all $i, j = 1, 2, \dots, n$.

Definition 11.15. A square matrix $A = (a_{ij})_{n \times n}$ is said to be skew symmetric if $a_{ij} = -a_{ji}$, for all $i, j = 1, 2, \dots, n$.

Example 11.14.

$$1. \text{ Let } A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -5 & 4 \\ 3 & 4 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 2 & -8 \\ -2 & 0 & 4 \\ 8 & -4 & 0 \end{pmatrix}$$

In A , $a_{ij} = a_{ji}$, $\forall i, j = 1, 2, 3$

$\therefore A$ is a symmetric matrix. In B , $a_{ij} = -a_{ji}$, $\forall i, j = 1, 2, 3$

$\therefore B$ is a skew symmetric matrix.

2. Let $C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Then C is a symmetric as well as a skew symmetric matrix.

3. Let $D = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -3 \\ 0 & 3 & 2 \end{pmatrix}$. Since $d_{33} \neq -d_{33}$

$\therefore D$ is not a skew symmetric matrix.

Theorem 11.9. Let A be an n -rowed square matrix. Then

(i) A is symmetric if and only if $A^t = A$.

(ii) A is skew symmetric if and only if $A^t = -A$.

Proof: Let $A = (a_{ij})_{n \times n}$. Then A^t is also a $n \times n$ matrix.

(i) A is symmetric

$$\Leftrightarrow a_{ji} = a_{ij}, \text{ for all } i, j = 1, \dots, n$$

$$\Leftrightarrow (i, j)\text{th element of } A^t = (i, j)\text{th element of } A \text{ for all } i, j = 1, 2, \dots, n$$

$$\Leftrightarrow A^t = A.$$

(ii) A is skew symmetric

$$\Leftrightarrow a_{ji} = -a_{ij}, \text{ for all } i, j = 1, 2, \dots, n$$

$$\Leftrightarrow (i, j)\text{th element of } A^t = -(i, j)\text{th element of } A, \text{ for all } i, j = 1, 2, \dots, n$$

$$\Leftrightarrow (i, j)\text{th element of } A^t = (i, j)\text{th element of } (-A), \text{ for all } i, j = 1, 2, \dots, n$$

$$\Leftrightarrow A^t = -A. \quad \square$$

Remark 11.1. If $A = (a_{ij})_{n \times n}$ is a skew symmetric matrix, then $a_{ji} = -a_{ij}$ for all $i, j = 1, 2, \dots, n$. In particular, $a_{ii} = -a_{ii}$, for all $i, j = 1, 2, \dots, n$

$$\Rightarrow 2a_{ii} = 0, \text{ for all } i, j = 1, 2, \dots, n$$

$$\Rightarrow a_{ii} = 0, \text{ for all } i = 1, 2, \dots, n$$

\Rightarrow diagonal elements are zero.

Thus the diagonal elements of a skew symmetric matrix are zero. But the converse of this result is not true, that is, a matrix having diagonal elements zero need not be a skew symmetric.

Theorem 11.10. Every square matrix can be uniquely expressed as the sum of a symmetric and a skew symmetric matrix.

Proof: Let A be a square matrix

Existence

$$\text{Let } P = \frac{1}{2}(A + A^t), Q = \frac{1}{2}(A - A^t)$$

$$\begin{aligned} \text{Then } P^t &= \left[\frac{1}{2}(A + A^t) \right]^t \\ &= \frac{1}{2}(A + A^t)^t, \text{ using } (kA)^t = kA^t \\ &= \frac{1}{2}(A^t + (A^t)^t), \text{ using } (A + B)^t = A^t + B^t \\ &= \frac{1}{2}(A^t + A) \\ &= P \therefore P^t = P, \text{ so that } P \text{ is symmetric.} \end{aligned}$$

$$\begin{aligned} Q^t &= \left[\frac{1}{2}(A - A^t) \right]^t \\ &= \frac{1}{2}(A - A^t)^t \\ &= \frac{1}{2}(A^t - A) \\ &= -\frac{1}{2}(A - A^t) \\ &= -Q \end{aligned}$$

$\therefore Q^t = -Q$, so that Q is a skew symmetric matrix.

Also $A = P + Q$.

Thus A is expressible as the sum of a symmetric and skew symmetric matrix.

Uniqueness

Let $A = X + Y \dots (1)$

where X is a symmetric matrix and Y a skew symmetric matrix.

Then $A^t = (X + Y)^t$

or $A^t = X^t + Y^t \dots (2)$

Thus (1) and (2) gives $X = \frac{1}{2}(A + A^t)$ $Y = \frac{1}{2}(A - A^t)$

This gives unique values of X and Y so that the expression in (1) is unique. \square

Hermitian and Skew Hermitian Matrices

Definition 11.16. Let $A = (a_{ij})_{n \times n}$. Then A is said to be Hermitian if $a_{ij} = \overline{a_{ji}}$, for all $i, j = 1, 2, \dots, n$.

Definition 11.17. Let $A = (a_{ij})_{n \times n}$. Then A is said to be skew Hermitian if $a_{ij} = -\overline{a_{ji}}$, for all $i, j = 1, 2, \dots, n$.

Example 11.15. Let $A = \begin{pmatrix} 2 & -i \\ i & -2 \end{pmatrix} = (a_{pq})_{2 \times 2}$.

Then $\overline{a_{pq}} = a_{qp}$, $p, q = 1, 2$ so that A is a Hermitian matrix.

Let $B = \begin{pmatrix} 0 & 2 + 3i \\ -2 + 3i & i \end{pmatrix}$ Then $b_{pq} = -\overline{b_{qp}}$, for $p, q = 1, 2$

so that B is a skew Hermitian matrix.

The following theorem gives a characterization of Hermitian and skew Hermitian matrices which is simpler to use.

Theorem 11.11. Let A be a square matrix. Then

(i) A is Hermitian if and only if $A^\theta = A$

(ii) A is skew Hermitian if and only if $A^\theta = -A$.

Proof: Left to the reader. \square

Remark 11.2. The diagonal elements of a skew Hermitian matrix are zero or pure imaginary. For if $A = (a_{pq})$, is a skew Hermitian matrix then $a_{pq} = -\overline{a_{qp}}$, for all $p, q = 1, 2, \dots, n$.

In particular $a_{pp} = -\overline{a_{pp}}$, for all $p = 1, 2, \dots, n$.

$\therefore a_{pp} + \overline{a_{pp}} = 0$

$\Rightarrow 2\text{Re } a_{pp} = 0$

$\Rightarrow \text{Re } a_{pp} = 0$, for all $p = 1, 2, \dots, n$

$\therefore a_{pp} = iy$ for some real number y

$\Rightarrow a_{pp} = 0$ or a pure imaginary number.

Theorem 11.12. A square matrix A is Hermitian if and only if A can be uniquely expressed as $B + iC$ where B is a real symmetric matrix and C is a real skew symmetric matrix.

Proof: Let A be a Hermitian matrix. Then $A = A^\theta \dots (1)$

Also $A = (\overline{A})^t$

so that $A^t = \overline{A} \dots (2)$

Existence

Let $P = \frac{1}{2}(A + \bar{A})$, $Q = \frac{1}{2i}(A - \bar{A})$. Then P and Q are matrices with real entries. Also

$$\begin{aligned} P^t &= \left(\frac{1}{2}(A + \bar{A})\right)^t \\ &= \frac{1}{2}(A^t + (\bar{A})^t) \\ &= \frac{1}{2}(\bar{A} + A), \text{ using (1) and (2)} \\ &= P. \end{aligned}$$

Thus P is a symmetric matrix.

$$\begin{aligned} Q^t &= \left(\frac{1}{2i}(A - \bar{A})\right)^t \\ &= \frac{1}{2i}(A^t - (\bar{A})^t) \\ &= \frac{1}{2i}(\bar{A} - A), \text{ using (1) and (2)} \\ &= -\frac{1}{2i}(A - \bar{A}) \\ &= -Q \end{aligned}$$

$\therefore Q^t = -Q$, so that Q is a skew symmetric matrix.

Also $P + iQ = A$

Thus A is expressed as $P + iQ$, where P, Q are real matrices, P is symmetric and Q is skew symmetric.

Uniqueness

Let $A = X + iY \dots\dots(3)$

where X is a real symmetric matrix and Y is a real skew symmetric matrix.

Then $\bar{X} = X$ and $\bar{Y} = -Y$.

Taking conjugate in (3) we get $\bar{A} = \bar{X} + i\bar{Y}$

$$\bar{A} = \bar{X} - i\bar{Y}$$

or $\bar{A} = X - iY \dots(4)$

$$(3) \text{ and } (4) \Rightarrow X = \frac{1}{2}(A + \bar{A}), Y = \frac{1}{2i}(A - \bar{A}).$$

Thus, the values of X and Y are expressed in terms of A and \bar{A} and are therefore unique.

Hence the expression is unique.

Conversely, let A be uniquely expressible as $A = B + iC \dots\dots(5)$

where B is a real symmetric matrix and C a real skew symmetric matrix.

Then $B = B^t$, $C = -C^t \dots(6)$

$$(5) \Rightarrow A^\theta = B^\theta + (iC)^\theta$$

$$= B^t - iC^t, \because B, C \text{ are real}$$

$$= B + iC$$

$$= A \quad \text{using (5)}$$

Hence A is Hermitian. □

Theorem 11.13. Every square matrix can be expressed uniquely as $P + iQ$ where P and Q are Hermitian matrices.

Proof: Let A be a square matrix.

Existence

$$\text{Let } P = \frac{1}{2}(A + A^\theta) \quad \dots(1)$$

$$Q = \frac{1}{2i}(A - A^\theta) \quad \dots(2)$$

Then $P + iQ = A$.

$$\begin{aligned} P^\theta &= \left(\frac{1}{2}(A + A^\theta)\right)^\theta \\ &= \frac{1}{2}(A + A^\theta)^\theta, \because (kA)^\theta = \bar{k}A^\theta \\ &= \frac{1}{2}(A^\theta + (A^\theta)^\theta), \because (A + B)^\theta = A^\theta + B^\theta \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2}(A^\theta + A), \because (A^\theta)^\theta = A \\ &= P \end{aligned}$$

$\therefore P$ is a Hermitian matrix.

Similarly $Q^\theta = Q$ so that Q is a Hermitian matrix.

Thus $A = P + iQ$ where P, Q given by (1) and (2) respectively are Hermitian matrices.

Uniqueness

Let $A = X + iY \dots (3)$

where X, Y are Hermitian matrices. Then $X^\theta = X$ and $Y^\theta = Y \dots (4)$

$$\begin{aligned} A^\theta &= (X + iY)^\theta \\ &= X^\theta + (iY)^\theta \\ &= X^\theta - iY^\theta \\ &= X - iY \quad \text{using (4)} \end{aligned}$$

$\therefore A^\theta = X - iY \dots (5)$

(3) and (5) $\Rightarrow X = \frac{1}{2}(A + A^\theta) \dots (6)$

$Y = \frac{1}{2i}(A - A^\theta) \dots (7)$

Hence the matrices X and Y are unique and are given by (6) and (7). \square

Corollary 11.14. *Every square matrix can be expressed uniquely as sum of a Hermitian and skew Hermitian matrix.*

Problem 11.1. *Prove that if A is a square matrix then A is Hermitian $\Leftrightarrow iA$ is skew Hermitian.*

Solution: Let A be a Hermitian matrix.

Then $A = A^\theta \dots (1)$

$$\begin{aligned} (iA)^\theta &= \bar{i}A^\theta \quad (\because (kA)^\theta = \bar{k}A^\theta) \\ &= -iA^\theta \\ &= -iA \quad \text{using (1)} \end{aligned}$$

Hence $(iA)^\theta = -iA$ so that iA is skew Hermitian.

Conversely, let iA be a skew Hermitian matrix.

Then $(iA)^\theta = -(iA)^\theta$

$$\Rightarrow iA = -\bar{i}A^\theta$$

$$\Rightarrow iA = iA^\theta$$

$$\Rightarrow A = A^\theta$$

$\Rightarrow A$ is a Hermitian.

Problem 11.2. *If A and B are Hermitian matrices such that $A^2 + B^2 = 0$, show that $A = B = 0$.*

Solution: Let $A = (a_{ij})_{n \times n}$, $B = (b_{ij})_{n \times n}$ be Hermitian matrices. Then $A = A^\theta$ and $B = B^\theta$.

Let $A^\theta = (c_{ij})_{n \times n}$, then $c_{ij} = \bar{a}_{ji}$

(i, j) th element of A^2

$= (i, j)$ th element of AA^θ

$$= \sum_{k=1}^n a_{ik}c_{kj}$$

$$= \sum_{k=1}^n a_{ik}\bar{a}_{jk}$$

$$\therefore (i, i)\text{th element of } A^2 = \sum_{k=1}^n a_{ik}\bar{a}_{ik}$$

$$= \sum_{k=1}^n |a_{ik}|^2$$

(i, i) th element of $(A^2 + B^2)$

$$= (i, i)\text{th element of } A^2 + (i, i)\text{th element of } B^2$$

$$= \sum_{k=1}^n |a_{ik}|^2 + \sum_{k=1}^n |b_{ik}|^2$$

$$\text{Now, } A^2 + B^2 = 0$$

$$\Rightarrow (i, i)\text{th element of } (A^2 + B^2) = 0, \forall i = 1, 2, \dots, n$$

$$\Rightarrow \sum_{k=1}^n |a_{ik}|^2 + \sum_{k=1}^n |b_{ik}|^2 = 0, \forall i = 1, 2, \dots, n.$$

Since a sum of non-negative quantities is zero if and only if each term is zero,

$$\therefore |a_{ik}|^2 = 0 = |b_{ik}|^2, \quad \forall i, k = 1, 2, \dots, n$$

$$\text{i.e. } a_{ik} = b_{ik} = 0, \quad \forall i, k = 1, 2, \dots, n$$

$$\text{i.e. } A = B = 0.$$

11.5 Exercise

1. Let $A = \begin{pmatrix} 1+i & 1-3i \\ -2 & -1+i \end{pmatrix}$, $B = \begin{pmatrix} 2i & 1-2i \\ 0 & -3+i \end{pmatrix}$, $C = \begin{pmatrix} 2-i \\ i \end{pmatrix}$.

Compute the following and express the entries in the form $a + ib$

- (i) $A + B$
- (ii) $(4 + i)A$
- (iii) AB
- (iv) BC
- (v) $A - 2iI_2$
- (vi) A^θ
- (vii) $B^\theta C$
- (viii) $(\overline{A} + \overline{B})C$
- (ix) $C^t A$
- (x) $B^\theta A^\theta$.

2. If $A = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$, compute all matrices B with complex entries such that $B^2 = A$.

3. For $f(x) = 2x^2 - 3x + 5$, compute $f(A)$ for each of the following:

$$\begin{array}{ll} \text{(i)} & A = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} & \text{(ii)} & A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \\ \text{(iii)} & A = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} & \text{(iv)} & A = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \end{array}$$

4. If A is a 2×2 matrix, the sum of whose diagonal elements is 0 and $|A| = 1$, then show that $f(A) = O_2$, where $f(x) = x^2 + 1$.

5. Find all 2×2 scalar matrices A which satisfy the relation $f(A) = 0$, for $f(x) = x^2 + 1$.

6. Determine the nature (symmetric, skew-symmetric, Hermitian or skew Hermitian) of the following matrices

$$\begin{aligned}
\text{(i)} & \begin{pmatrix} 1+i & 2 & i \\ 2 & 0 & 4-i \\ i & 4-i & 5i \end{pmatrix} \\
\text{(ii)} & \begin{pmatrix} 2i & -4+i & 2+3i \\ 4+i & 0 & 5i \\ -2+3i & 5i & -2i \end{pmatrix} \\
\text{(iii)} & \begin{pmatrix} 4 & 2i & 1+i \\ 2i & 5i & -3i \\ 1+i & -3i & 6 \end{pmatrix} \\
\text{(iv)} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\text{(v)} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\text{(vi)} & \begin{pmatrix} 2 & 1+i & -2+i \\ 1-i & 3 & 6i \\ -2-i & -6i & 4 \end{pmatrix} \\
\text{(vii)} & \begin{pmatrix} 0 & -2+i & 4 \\ 2-i & 0 & 6i \\ -4 & -6i & 0 \end{pmatrix} \\
\text{(viii)} & \begin{pmatrix} 0 & 2 & -3 \\ -2 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix}.
\end{aligned}$$

7. Let D_1, D_2 be diagonal matrices and A be a matrix such that $A^2 = I$. If $P = AD_1A$ and $Q = AD_2A$. Prove that P and Q commute.

8. Prove that if A is a skew Hermitian matrix, then iA and $-iA$ are Hermitian matrices.

9. If $A = \begin{pmatrix} 2+i & 3 & -4 \\ 5+i & 2 & 6+3i \\ 3i & -1+4i & 6i \end{pmatrix}$, express A as

- (i) sum of a symmetric and a skew symmetric matrix.
- (ii) sum of a Hermitian and a skew Hermitian matrix.
- (iii) $P + iQ$, where P, Q are Hermitian matrices.

10. Express A in the form $P + iQ$, where P is a real symmetric matrix and Q is a real skew symmetric matrix, where

$$\begin{aligned}
\text{(i)} \quad A &= \begin{pmatrix} 1 & 2-3i & 4i+3 \\ 2+3i & 0 & 4-5i \\ 3-4i & 4+5i & 2 \end{pmatrix} \\
\text{(ii)} \quad A &= \begin{pmatrix} 2 & 5+6i & -1+2i \\ 5-6i & 3 & 3-4i \\ -1-2i & 3+4i & i \end{pmatrix}.
\end{aligned}$$

11. How many independent elements are there in the following matrices?
- a $m \times n$ matrix
 - a $n \times n$ matrix
 - a matrix of order n with trace zero
 - a diagonal matrix of order n
 - a scalar matrix of order n
 - a symmetric matrix of order n
 - a skew symmetric matrix of order n
 - a Hermitian matrix of order n
 - a skew Hermitian matrix of order n
12. If A is a symmetric (skew symmetric) matrix, show that for any matrix B , B^tAB is symmetric (skew symmetric).
13. Prove that every skew symmetric matrix of odd order is singular.
14. Prove that if A and B are symmetric (Hermitian) matrices then AB is symmetric (Hermitian) if and only if $AB = BA$.
15. Prove that all the positive integral powers of a symmetric (Hermitian) matrix is symmetric (Hermitian).
16. If A is a Hermitian matrix, prove that AA^θ and $A^\theta A$ are also Hermitian.
17. If A is a Hermitian matrix such that $A^2 = 0$, show that $A = 0$.
18. Show that every Hermitian matrix is normal. Is the converse true?
19. Given a 9×12 matrix, partition it in the form $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ where A_{11} is (i) 4×5 (ii) 7×6 (iii) 3×4 (iv) 6×9 .
State the orders of the matrices A_{12} , A_{21} and A_{22} in each case.
20. If a $m \times n$ matrix A is partitioned as $\begin{pmatrix} P & Q \\ R & S \end{pmatrix}$, prove that $A^t = \begin{pmatrix} P^t & R^t \\ Q^t & S^t \end{pmatrix}$.

21. Find AB using partitioning of matrices, where

$$A = \left(\begin{array}{cc|cc|c} 1 & 0 & 2 & -1 & 3 \\ -1 & 2 & 1 & 3 & 5 \\ \hline 4 & 6 & 3 & 2 & 1 \\ 2 & 1 & 3 & 4 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{array} \right), \quad B = \left(\begin{array}{ccc|cc|c} 0 & 1 & 2 & 1 & 2 & 1 \\ 1 & 3 & -1 & 2 & 0 & 1 \\ \hline 2 & 3 & 1 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 & 1 & 2 \\ 4 & 5 & 0 & 1 & 2 & 3 \end{array} \right)$$

Also find AB by (i) direct multiplication

(ii) by considering another partition of A and B .

22. Find AB using partitioning of matrices, where

$$A = \begin{pmatrix} 1 & -1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & -1 & 3 & 0 \\ -1 & 0 & 1 & -2 & 0 & 3 \\ 1 & 1 & 2 & -1 & 1 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & 2 \\ -1 & 0 & 1 & 1 \\ 2 & 1 & -1 & -1 \\ -2 & -1 & 0 & 0 \\ 3 & 2 & 1 & 1 \end{pmatrix}.$$

23. If A and B are skew symmetric matrices, prove that
- $AB + BA$ is symmetric.
 - $AB - BA$ is skew symmetric.
24. If A and B are symmetric matrices prove that
- $AB + BA$ is symmetric
 - $AB - BA$ is skew symmetric.
25. If A and B are skew Hermitian matrices, prove that AB is skew Hermitian if and only if $AB = -BA$.
26. If A and B are symmetric matrices, prove that AB is symmetric if and only if $AB = BA$.

11.6 Inverse of a Matrix

So far we have added and multiplied matrices. However, division is not defined for arbitrary matrices. We now study the matrix analogue of the reciprocal of a number.

Definition 11.18. Let A be a $m \times n$ matrix. Then

- an $n \times m$ matrix B is called a left inverse of A if $BA = I_n$.
- an $n \times m$ matrix C is called a right inverse of A if $AC = I_m$.

Example 11.16. Let $A = \begin{pmatrix} 1 & 2 & 1 \\ -2 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -3 \\ -1 & 4 \\ 2 & -5 \end{pmatrix}$

$$\begin{aligned} &\text{Then } AB = I_2 \\ BA &= \begin{pmatrix} 7 & 2 & -2 \\ -9 & -2 & 3 \\ 12 & 4 & -3 \end{pmatrix} \neq I_3 \end{aligned}$$

Hence B is a right inverse of A but not a left inverse of A . Also A is a left inverse of B but not a right inverse of B .

Since matrix multiplication is not commutative, a matrix may have one sided inverse only. But if a square matrix has a right and a left inverse then they must be equal, as is proved in the following theorem.

Theorem 11.15. If a square matrix A has a right inverse B and a left inverse C then $B = C$.

Proof: Since B is a right inverse

$$\therefore AB = I_n \quad \dots(1)$$

Since C is a left inverse

$$\therefore CA = I_n \quad \dots(2)$$

Since matrix multiplication is associative

$$\begin{aligned} \therefore C(AB) &= (CA)B \\ \text{using (1) and (2)} & \\ CI_n &= I_n B \\ \text{or } C &= B. \end{aligned}$$

□

If a square matrix A has both left and right inverses, then any two left inverses must be equal to the right inverse B , and hence to each other. Thus left inverse is unique.

Similarly right inverse is unique. So there exist only one left and right inverses and they must be equal to each other. But a non-square matrix can have only one sided inverse, and it may not be unique as is shown by the following example.

Example 11.17. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, if $x, y \in \mathbb{R}$, then $B = \begin{pmatrix} 1 & 0 \\ x & y \\ 0 & 1 \end{pmatrix}$ are

such that $AB = I_2$.

Thus A has infinitely many right inverses.

Definition 11.19. A $n \times n$ square matrix A is said to be invertible if there exists a $n \times n$ square matrix B such that $AB = BA = I_n$. B is called an inverse of A . If a matrix does not have an inverse it is said to be non-invertible.

By the definition, A has a right and left inverse B . Since right and left inverses are unique, therefore it follows that inverse of a matrix is unique so that we can talk of ‘the inverse’ instead of ‘an inverse’. The inverse of matrix A is denoted by A^{-1} .

Not every square matrix is invertible. In fact the null matrix is not invertible. Also a square matrix having a row (or column) of zeros does not have an inverse.

11.7 Adjoint of a Matrix

Definition 11.20. (Minor):

Let A be any $m \times n$ matrix. The determinant of any p -rowed square submatrix of A obtained by deleting $m-p$ rows and $n-p$ columns is called a p -rowed minor of A .

Definition 11.21. If $A = (a_{ij})_{n \times n}$ is an n -rowed matrix. The minor of a_{ij} is the determinant of the submatrix obtained by deleting the i th row and j th column.

The minor of a_{ij} is denoted by M_{ij} .

Definition 11.22. (Cofactor):

Let $A = (a_{ij})_{n \times n}$. The cofactor of an element a_{ij} of A is $(-1)^{i+j}$ times the determinant of the sub-matrix of A obtained by deleting the i th row and the j th column. It is denoted by A_{ij} .

Thus $A_{ij} = (-1)^{i+j} M_{ij}$.

Definition 11.23. (Adjoint of a Matrix)

Let $A = (a_{ij})_{n \times n}$. The matrix $B = (b_{ij})$, where $b_{ij} = \text{cofactor of } a_{ji}$, is called the adjoint of A and is denoted by $\text{adj}A$.

To obtain the adjoint of a matrix A , take the transpose A^t of the matrix A and then replace each element of A^t by its cofactor.

Example 11.18. Find the adjoint of the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 5 & 12 \end{pmatrix}$.

Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 5 & 12 \end{pmatrix}$. Then $A^t = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 3 & 5 & 12 \end{pmatrix}$

The cofactors of the elements of the 1st row of A^t are 11, -9, 1.

Cofactors of the elements of the 2nd row of A^t are -7, 9, -2.

Cofactors of the elements of the 3rd row of A^t are 2, -3, 1.

$$\text{Hence } \text{adj}A = \begin{pmatrix} 11 & -9 & 1 \\ -7 & 9 & -2 \\ 2 & -3 & 1 \end{pmatrix}.$$

The following theorem gives a relation between a matrix A and its adjoint.

Theorem 11.16. If A is an n -rowed square matrix, then

$$A(\text{adj}A) = (\text{adj}A)A = |A|I$$

Proof: Since A and $\text{adj}A$ are both n -rowed square matrices, therefore $A(\text{adj}A)$ and $(\text{adj}A)A$ are both n -rowed square matrices. Let $A = (a_{ij})_{n \times n}$ and $\text{adj}A = (b_{ij})_{n \times n}$ where $b_{ij} = A_{ij}$.

We know that $\sum_{j=1}^n a_{ij}A_{ij} = |A|$, $\sum_{j=1}^n a_{ij}A_{kj} = 0$, if $k \neq i$

Now, (i, j) th element of $A(\text{adj}A)$

$$\begin{aligned} &= \sum_{k=1}^n a_{ik}b_{kj} \\ &= \sum_{k=1}^n a_{ik}A_{jk} \\ &= \begin{cases} 0, & \text{if } j \neq i; \\ |A|, & \text{if } j = i. \end{cases} \end{aligned}$$

$$\therefore A(\text{adj}A) = |A|I_n$$

$$\text{Similarly } (\text{adj}A)A = |A|I_n$$

$$\text{Hence } A(\text{adj}A) = (\text{adj}A)A = |A|I_n. \quad \square$$

Corollary 11.17. If A is a matrix, such that $|A| \neq 0$, then A is invertible and $A^{-1} = \frac{1}{|A|}(\text{adj}A)$. Since A is non-singular, $|A| \neq 0$. Thus

$$A\left(\frac{1}{|A|}\text{adj}A\right) = \frac{1}{|A|}A(\text{adj}A) = \frac{1}{|A|}|A|I_n = I_n.$$

so that A^{-1} exists and $A^{-1} = \frac{1}{|A|}\text{adj}A. \quad \square$

Remark 11.3. From the theorem, it follows that $|\text{adj}A| = |A|^{n-1}$, if $|A| \neq 0$.

It has been proved later that $|A| = 0 \Leftrightarrow |\text{adj}A| = 0$, so that we can always say that $|\text{adj}A| = |A|^{n-1}$.

Definition 11.24. (Singular matrix):

A square matrix A is said to be singular if $|A| = 0$, and non-singular if $|A| \neq 0$.

Theorem 11.18. (Existence of the Inverse):

A necessary and sufficient condition for a square matrix A to be invertible is that it should be non-singular.

Proof: Let A be invertible. Then there exists a matrix B such that

$$\begin{aligned} AB &= BA = I_n \\ \therefore |AB| &= |I_n| \\ \Rightarrow |A||B| &= 1 \\ \Rightarrow |A| &\neq 0 \\ \Rightarrow A &\text{ is non-singular.} \end{aligned}$$

Conversely, let A be non-singular. By corollary 11.17, A is invertible. \square

11.8 Negative Integral Powers of a Non-singular Matrix

Let A be a non-singular matrix. If p is any positive integer, then we define $(A^{-p}) = (A^p)^{-1}$.

Theorem 11.19. *If A is any n -rowed non-singular matrix, then $(A^{-k}) = (A^{-1})^k$, for all $k \in \mathbb{N}$.*

Proof: By definition, if $k \in \mathbb{N}$

$$\begin{aligned} A^{-k} &= (A^k)^{-1} \\ &= (A.A \dots k \text{ times})^{-1} \\ &= A^{-1}A^{-1} \dots k \text{ times} \\ &= (A^{-1})^k \\ \therefore A^{-k} &= (A^k)^{-1} = (A^{-1})^k. \end{aligned} \quad \square$$

Theorem 11.20. *Let A, B be invertible matrices. Then*

- (i) $(A^{-1})^{-1} = A$
- (ii) $(kA)^{-1} = k^{-1}A^{-1}, 0 \neq k \in \mathbb{C}$
- (iii) $(AB)^{-1} = B^{-1}A^{-1}$
- (iv) $(A^t)^{-1} = (A^{-1})^t$
- (v) $(A^\theta)^{-1} = (A^{-1})^\theta$.

Proof:

- (i) Since A is invertible, therefore there exists a matrix $B = A^{-1}$ such that $AB = BA = I$

$$\text{Then } B^{-1} = A \text{ by definition of inverse i.e. } (A^{-1})^{-1} = A.$$

- (ii) Let $C = k^{-1}A^{-1}$. Then,

$$(kA)C = (kA)(k^{-1}A^{-1}) = kk^{-1}AA^{-1} = I.$$

Similarly

$$C(kA) = I.$$

Thus

$$\begin{aligned} (kA)C &= C(kA) = I \\ \Rightarrow (kA)^{-1} &= C = k^{-1}A^{-1}. \end{aligned}$$

- (iii) Let $C = B^{-1}A^{-1}$. Then,

$$\begin{aligned} (AB)C &= (AB)(B^{-1}A^{-1}) \\ &= A(B(B^{-1}A^{-1})) \\ &= A(BB^{-1}A^{-1}) \\ &= A(IA^{-1}) = AA^{-1} = I. \end{aligned}$$

Similarly

$$\begin{aligned} C(AB) &= I, \text{ so that} \\ (AB)C &= C(AB) = I. \end{aligned}$$

$$\text{Hence } (AB)^{-1} = C = B^{-1}A^{-1}.$$

(iv) We know that

$$AA^{-1} = A^{-1}A = I$$

Taking transpose, we get

$$\begin{aligned} & (AA^{-1})^t = (A^{-1}A)^t = I^t \\ \Rightarrow & (A^{-1})^t A^t = A^t (A^{-1})^t = I, \text{ using } (AB)^t = B^t A^t \\ \Rightarrow & (A^t)^{-1} = (A^{-1})^t. \end{aligned}$$

(v) Similar to (iv) but take transposed conjugate instead of transpose. \square

11.9 Inverse of Partitioned Matrices

If the partitioned matrices are of special types, then it is easy to find their inverses. However the inverse of any invertible matrix can be obtained by partitioning.

Definition 11.25. A matrix having the blocks below(above) the main diagonal blocks, as blocks of zero is called a block upper triangular (lower triangular) matrix.

$$\text{Such a matrix is of the form } \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ A_{21} & A_{22} \end{pmatrix}.$$

Definition 11.26. A partitioned matrix having zero blocks off the main diagonal blocks is called a block diagonal matrix.

Example 11.19. $A = \left(\begin{array}{cc|ccc} 1 & 2 & 3 & 4 & 1 \\ 2 & 1 & 5 & 8 & -1 \\ 0 & 0 & 4 & 8 & 0 \\ 0 & 0 & 2 & 3 & 2 \end{array} \right) = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$ is a block upper triangular matrix.

$B = \left(\begin{array}{ccc|cc} 2 & 1 & 0 & 0 & 0 \\ 2 & 3 & 5 & 0 & 0 \\ 1 & 5 & 8 & 2 & 3 \end{array} \right) = \begin{pmatrix} B_{11} & O \\ B_{12} & B_{22} \end{pmatrix}$ is a block lower triangular matrix.

$$C = \left(\begin{array}{cc|ccc} 2 & 1 & 0 & 0 & 0 \\ 5 & 4 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 8 \end{array} \right) = \begin{pmatrix} C_{11} & O \\ O & C_{22} \end{pmatrix} \text{ is a block diagonal matrix.}$$

It is important to observe that every partition of A may not give it the form of a block upper triangular matrix. For instance, partitioning A as

$$\left(\begin{array}{cc|ccc} 1 & 2 & 3 & 4 & 1 \\ 2 & 1 & 5 & 8 & -1 \\ 0 & 0 & 4 & 8 & 0 \\ 0 & 0 & 2 & 3 & 2 \end{array} \right) \text{ does not make it a block upper triangular matrix.}$$

It is easy to find the inverse of block upper (lower) triangular and block diagonal matrices.

Theorem 11.21. If $A = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$ is a block upper triangular matrix such that A_{11} and A_{22} are square matrices of orders p and q respectively, then A is invertible if and only if A_{11} and A_{22} are invertible and

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{pmatrix}.$$

Proof: Let A be invertible and let $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ be the inverse of A .

Then $AB = I \dots (1)$

$$\Rightarrow \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} I_p & O \\ O & I_q \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{pmatrix} = \begin{pmatrix} I_p & O \\ O & I_q \end{pmatrix}$$

$$\Rightarrow A_{11}B_{11} + A_{12}B_{21} = I_p \dots (2)$$

$$A_{11}B_{12} + A_{12}B_{22} = 0 \dots (3)$$

$$A_{22}B_{21} = 0 \dots (4)$$

$$A_{22}B_{22} = I_q \dots (5)$$

(5) $\Rightarrow A_{22}$ has a right inverse B_{22} . Since A_{22} is a square matrix, $\therefore B_{22}$ is the inverse of A_{22} by 11.15

$$\therefore A_{22}^{-1} = B_{22} \dots (6)$$

Pre-multiplying (4) by A_{22}^{-1} , we get

$$IB_{21} = A_{22}^{-1}0$$

$$\text{or } B_{21} = 0 \dots (7)$$

Thus (2) and (7) $\Rightarrow A_{11}B_{11} = I_p$

Thus B_{11} is a right inverse of the square matrix A_{11} , so that $A_{11}^{-1} = B_{11}$.

$$\therefore (3) \text{ gives } A_{11}B_{12} + A_{12}A_{22}^{-1} = 0$$

$$\Rightarrow B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

$$\text{Hence } A^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{pmatrix}.$$

Conversely let A_{11} and A_{22} be invertible matrices.

$$\text{Let } B = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{pmatrix}$$

$$\begin{aligned} \text{Then } AB &= \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} A_{11}A_{11}^{-1} & -A_{11}A_{11}^{-1}A_{12}A_{22}^{-1} + A_{12}A_{22}^{-1} \\ O & A_{22}A_{22}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} I_p & -A_{12}A_{22}^{-1} + A_{12}A_{22}^{-1} \\ O & I_q \end{pmatrix} \\ &= I \end{aligned}$$

Similarly $BA = I$.

Hence $AB = BA = I$.

So that A is invertible and B is the inverse of A . □

Example 11.20. Find the inverse of $\begin{pmatrix} 1 & 2 & 1 \\ -1 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

Solution: Let $A = \left(\begin{array}{cc|c} 1 & 2 & 1 \\ -1 & 2 & 1 \\ 0 & 0 & 3 \end{array} \right) = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$

Where $A_{11} = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}$, $A_{12} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $A_{22} = (3)$

Since $|A_{11}| \neq 0$ and $|A_{22}| \neq 0$, therefore A_{11} and A_{22} are non-singular and therefore invertible. Also

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{pmatrix}$$

Now $A_{11}^{-1} = \frac{1}{4} \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}$, $A_{22}^{-1} = \left(\frac{1}{3} \right) = \frac{1}{3} (1)$

$$A_{11}^{-1}A_{12}A_{22}^{-1} = \frac{1}{4} \times \frac{1}{3} \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1)$$

$$= \frac{1}{12} \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\therefore A^{-1} = \left(\begin{array}{cc|c} \frac{2}{4} & \frac{-2}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{-1}{6} \\ \hline 0 & 0 & \frac{1}{3} \end{array} \right) = \frac{1}{12} \begin{pmatrix} 6 & -6 & 0 \\ 3 & 3 & -2 \\ 0 & 0 & 4 \end{pmatrix}$$

Theorem 11.22. If $A = \begin{pmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{pmatrix}$ is a block lower triangular matrix such that A_{11} and A_{22} are square matrices, then A is invertible if and only if A_{11} and A_{22} are invertible and $A^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{12}A_{11}^{-1} & A_{22}^{-1} \end{pmatrix}$

Proof: Similar to the proof of Theorem 11.21 □

Corollary 11.23. If $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$ is a block diagonal matrix such that A_{11} and A_{22} are square matrices, then A is invertible if and only if A_{11} and A_{22} are invertible and $A^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{pmatrix}$ □

Theorem 11.24. If $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & O \end{pmatrix}$ is a matrix such that A_{12} and A_{21} are square matrices, then A is invertible if and only if A_{12} and A_{21} are invertible, and $A^{-1} = \begin{pmatrix} O & A_{21}^{-1} \\ A_{12}^{-1} & -A_{12}^{-1}A_{11}A_{21}^{-1} \end{pmatrix}$

Proof: Similar to the proof of Theorem 11.21. □

Corollary 11.25. If $A = \begin{pmatrix} 0 & A_{12} \\ A_{21} & 0 \end{pmatrix}$, such that A_{12} and A_{21} are square matrices then A is invertible if and only if A_{12} and A_{21} are invertible and $A^{-1} = \begin{pmatrix} 0 & A_{21}^{-1} \\ A_{12}^{-1} & 0 \end{pmatrix}$. □

Theorem 11.26. If $A = \begin{pmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ is a matrix such that A_{12} and A_{21} are square matrices, then A is invertible if and only if A_{12} and A_{21} are invertible, and $A^{-1} = \begin{pmatrix} -A_{21}^{-1}A_{22}A_{12}^{-1} & A_{21}^{-1} \\ A_{12}^{-1} & 0 \end{pmatrix}$

Proof: Similar to the proof of Theorem 11.21. □

11.10 Solved Problems

Problem 11.3. Prove that A is non-singular if and only if $\text{adj}A$ is non-singular.

Solution: We know that for any matrix A , $A(\text{adj}A) = (\text{adj}A)A = |A|I \dots (1)$

Let A be non-singular. Then $|A| \neq 0$

$$A(\text{adj}A) = |A|I$$

$$\Rightarrow |A|\text{adj}A = |A|^n, \text{ where } A \text{ is a } n \times n \text{ matrix}$$

$$\Rightarrow |A||\text{adj}A| = |A|^n$$

$$\Rightarrow |\text{adj}A| = |A|^{n-1}$$

$$\Rightarrow |\text{adj}A| \neq 0$$

$$\Rightarrow \text{adj}A \text{ is non singular.}$$

Conversely, let $(\text{adj}A)$ be non-singular.

Then there exists a matrix B such that

$$(\text{adj}A)B = B(\text{adj}A) = I$$

$$(\text{adj}A)B = I$$

$$\Rightarrow A(\text{adj}A)B = AI \text{ (premultiplying by } A)$$

$$\Rightarrow |A|IB = A \text{ using (1)}$$

$$\Rightarrow |A|B = A.$$

$$\text{If } |A| = 0 \text{ then } A = 0$$

$$\Rightarrow \text{adj}A = 0$$

$$\Rightarrow |\text{adj}A| = 0, \text{ a contradiction.}$$

$\therefore |A| \neq 0$ so that A is non-singular.

Problem 11.4. If A is any square matrix then $\text{adj}A^t = (\text{adj}A)^t$.

Solution: Let $A = (a_{ij})_{n \times n}$. Then $\text{adj}A = (A_{ij})_{n \times n}$, where A_{ij} is the cofactor of a_{ij} .

Since $\text{adj}A^t$ and $(\text{adj}A)^t$ are $n \times n$ matrices, therefore they are comparable. For $i, j = 1, 2, \dots, n$ (i, j)th element of $(\text{adj}A)^t$

$$= (j, i)\text{th element of } \text{adj}A$$

$$= \text{cofactor of } a_{ij} \text{ in } A$$

$$= \text{cofactor of } a_{ji} \text{ in } A^t$$

$$= (i, j)\text{th element of } (\text{adj}A^t)$$

Thus, the corresponding elements of $(\text{adj}A)^t$ and $\text{adj}A^t$ are equal, so that $(\text{adj}A)^t = \text{adj}A^t$.

Problem 11.5. If A and B are square matrices of the same order, then $\text{adj}(AB) = (\text{adj}B)(\text{adj}A)$.

Solution: We know that

$$AB(\text{adj}AB) = (\text{adj}AB)(AB) = |AB|I \dots (1)$$

$$\text{Now, } (AB)(\text{adj}B)(\text{adj}A) = A(B\text{adj}B)\text{adj}A$$

$$= A(|B|I)\text{adj}A$$

$$= |B|A(\text{adj}A)$$

$$= |B||A|I$$

$$= |A||B|I$$

$$= |AB|I$$

$$\text{Similarly, } (\text{adj}B)(\text{adj}A)(AB) = |AB|I$$

$$\text{Hence, } AB(\text{adj}B\text{adj}A) = (\text{adj}B)(\text{adj}A)AB = |AB|I \dots (2)$$

Comparing (1) and (2), we get $(\text{adj}AB) = (\text{adj}B)(\text{adj}A)$.

Problem 11.6. Find a right inverse of the matrix $\begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$. Is it unique? Can you find all the right inverses?

Solution: Let $A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$.

If $B = (c_1 \ c_2)$ is a right inverse A then $AB = I_2 = (e_1 \ e_2)$ say
 $\Rightarrow Ac_1 = e_1$ and $Ac_2 = e_2$

To find B , we need to solve $AX = e_1$, and $AX = e_2$

$$(A \mid e_1 \mid e_2) = \left(\begin{array}{ccc|c|c} 1 & -1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 0 & 1 \end{array} \right)$$

Using E -row operations, we get $(A \mid e_1 \mid e_2) \sim \left(\begin{array}{ccc|c|c} 1 & 0 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{array} \right)$

$$\therefore \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} X = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \text{ where } X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\therefore x + 2z = 2$$

$$y + z = 1$$

Solution is $x = 2 - 2k$

$$y = 1 - k$$

$z = k$, where k is any real number.

$$\therefore c_1 = X = \begin{pmatrix} 2 - 2k \\ 1 - k \\ k \end{pmatrix}$$

$$\text{Also } \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

gives $c_2 = \begin{pmatrix} 1 - 2k_1 \\ 1 - k_1 \\ k_1 \end{pmatrix}$, where k_1 is any real number.

Hence $B = (c_1 \ c_2)$

$$= \begin{pmatrix} 2 - 2k & 1 - 2k_1 \\ 1 - k & 1 - k_1 \\ k & k_1 \end{pmatrix}$$

Taking $k = 0, k_1 = 1$, we get $B_1 = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

B_1 is a particular right inverse of A . Right inverse is not unique. In fact any

matrix of the form $\begin{pmatrix} 2 - 2k & 1 - 2k_1 \\ 1 - k & 1 - k_1 \\ k & k_1 \end{pmatrix}$ where k and k_1 are any real numbers

is a right inverse.

Problem 11.7. Find the inverse of $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ using partitioning of

matrices.

Solution: Let $X = \left(\begin{array}{c|ccc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$

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

where $A = [1], C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Then, $X^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & C^{-1} \end{pmatrix}$

since, $C = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array} \right) = \begin{pmatrix} A_1 & 0 \\ 0 & C_1 \end{pmatrix}$

$$\therefore C^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ 0 & C_1^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ so that } X^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Problem 11.8. Find the inverse of the matrix $\begin{pmatrix} 1 & 0 & 1 & 1 & -1 \\ -1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$ using

partitioning of matrices. Do the partition in two different ways.

Solution: Let $P = \begin{pmatrix} 1 & 0 & 1 & 1 & -1 \\ -1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$

1st way of partitioning

Partition P so that on the diagonals there are 3×3 and 2×2 matrices

$$\therefore P = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

Where $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

Then $P^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ 0 & C^{-1} \end{pmatrix}$

Now C being a 2×2 matrix $C^{-1} = \frac{-1}{2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

To find A^{-1} we again use partitioning.

Let $A = \left(\begin{array}{cc|c} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) = \begin{pmatrix} A_1 & B_1 \\ 0 & C_1 \end{pmatrix}$ (say)

$$\therefore A^{-1} = \begin{pmatrix} A_1^{-1} & -A_1^{-1}B_1C_1^{-1} \\ 0 & C_1^{-1} \end{pmatrix}$$

$$A_1^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, C_1^{-1} = [1], A_1^{-1}B_1C_1^{-1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} \therefore A^{-1} &= \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \\ A^{-1}BC^{-1} &= \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \therefore P^{-1} &= \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{-1}{2} \end{pmatrix} \end{aligned}$$

2nd way of partitioning

Partition P so that on the main diagonal there are 2×2 and 3×3 matrices respectively. Then $P = \begin{pmatrix} X & Y \\ O & W \end{pmatrix}$,

$$\text{where } X = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, Y = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}, W = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Then $P^{-1} = \begin{pmatrix} X^{-1} & -X^{-1}YW^{-1} \\ O & W^{-1} \end{pmatrix}$. To obtain W^{-1} we can partition it

$$\text{as } W = \left(\begin{array}{c|cc} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{array} \right) \text{ Then we obtain } P^{-1} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{-1}{2} \end{pmatrix}$$

which is the same as before.

11.11 Exercise

1. Find a right inverse of the matrix $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$
2. Find a left inverse of the matrix $\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$. Is it unique? Can you find all the left inverses?
3. Let A be a matrix such that there exists a non-zero matrix B satisfying $AB = 0$. Prove that A does not have a left inverse.
4. Prove that if a matrix A has two distinct right inverses, then there exists a non-zero matrix B such that $AB = 0$.
5. Let A be a matrix such that there exists a non-zero matrix C satisfying $CA = 0$. Prove that A does not have a right inverse.
6. Prove that a matrix A has a right inverse if and only if A^t has a left inverse.
7. Prove that a singular matrix does not have an inverse.

8. If A is a non-singular matrix, prove that
- $(adj A)^{-1} = adj A^{-1}$.
 - $adj(adj A) = |A|^{n-2}A$.
9. If A is a skew symmetric matrix of order n , then $adj A$ is symmetric or skew symmetric according as n is odd or even.
10. For the following matrices verify $A(adj A) = (adj A)A = |A|I_n$

$$(i) \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix}$$

$$(ii) \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 0 \\ 2 & 3 & 6 \end{pmatrix}$$

$$(iii) \quad \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}$$

$$(iv) \quad \begin{pmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \\ -3 & 1 & 2 \end{pmatrix}$$

11. Prove that
- the adjoint of a diagonal matrix is a diagonal matrix.
 - the adjoint of a scalar matrix is a scalar matrix.
 - the adjoint of a triangular matrix is a triangular matrix
 - the adjoint of a symmetric matrix is a symmetric matrix.
 - the adjoint of a Hermitian matrix is also Hermitian.

12. If A is any square matrix, prove that $(adj kA) = k^{n-1}(adj A)$.

13. Prove that A is invertible if and only if $adj A$ is invertible.

14. Find the inverse of the following matrices

$$(i) \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 5 & 12 \end{pmatrix} \quad (iv) \quad \begin{pmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{pmatrix}$$

$$(ii) \quad \begin{pmatrix} 1 & -1 & 1 \\ 4 & 1 & 0 \\ 8 & 1 & 1 \end{pmatrix} \quad (v) \quad \begin{pmatrix} 1 & x & 0 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}$$

$$(iii) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 3 & -1 \end{pmatrix} \quad (vi) \quad \begin{pmatrix} 1 & 0 & 0 \\ y & 1 & 0 \\ 0 & y & 1 \end{pmatrix}$$

15. If $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 4 & 6 & 2 \end{pmatrix}$, find A^{-2} . Also verify that $(A^2)^{-1} = (A^{-1})^2$.

16. If $A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$, find A^{-3} . Also verify that $(A^3)^{-1} = (A^{-1})^3$.
17. Prove the following:
- The inverse of a diagonal matrix is diagonal matrix.
 - The inverse of a scalar matrix is a scalar matrix.
 - The inverse of a symmetric matrix is a symmetric matrix.
 - The inverse of a skew Hermitian matrix is a skew Hermitian matrix.
 - The inverse of a triangular matrix is triangular.
18. Obtain by partitioning, the inverse of following matrices
- $\begin{pmatrix} 2 & 3 & 3 \\ 2 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$
 - $\begin{pmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & \beta & 1 \end{pmatrix}$
19. Show that the inverse of the matrix $\begin{pmatrix} A & 0 \\ B & C \end{pmatrix}$ is $\begin{pmatrix} A^{-1} & 0 \\ -C^{-1}BA^{-1} & C^{-1} \end{pmatrix}$ where A and C are non-singular matrices.
20. If A and B are invertible matrices, prove that $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}$
21. If B and C are invertible matrices, prove that
- $\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & C^{-1} \\ B^{-1} & -B^{-1}AC^{-1} \end{pmatrix}$
 - $\begin{pmatrix} 0 & B \\ C & A \end{pmatrix}^{-1} = \begin{pmatrix} -C^{-1}AB^{-1} & C^{-1} \\ B^{-1} & 0 \end{pmatrix}$
22. Find the inverse of the following matrices, using partitioning of matrices. Also verify your answer.
- $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & -2 \end{pmatrix}$
 - $\begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \\ 2 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 \end{pmatrix}$
 - $\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

$$(iv) \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$(v) \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & -1 & -1 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$(vi) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

23. Using partitioning of matrices, obtain the inverse of the matrix

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

11.12 Orthogonal and Unitary Matrices

It is known that multiplication by a $m \times n$ matrix defines a mapping from \mathbb{C}^n to \mathbb{C}^m . In this section we study those matrices which give rise to length preserving mappings.

Definition 11.27. (Dot product): If X and Y are two vectors in \mathbb{C}^n , $X = (x_1, \dots, x_n)^t$, $Y = (y_1, \dots, y_n)^t$, then $X^\theta Y = \bar{x}_1 y_1 + \bar{x}_2 y_2 + \dots + \bar{x}_n y_n$ is called the dot product of X by Y , and is denoted by $X \cdot Y$. If $X, Y \in \mathbb{R}^n$, then $X^\theta Y = X^t Y = x_1 y_1 + \dots + x_n y_n$.

Definition 11.28. (Orthogonal vectors): Two vectors X and Y in \mathbb{C}^n are said to be orthogonal if $X^\theta Y = 0$. If $X, Y \in \mathbb{R}^n$ then X and Y are orthogonal if $x_1 y_1 + x_2 y_2 + \dots + x_n y_n = 0$.

Definition 11.29. (Norm or length of a vector):

For any vector $X \in \mathbb{C}^n$, the positive square root of $X^\theta X$ is called the norm of X and is denoted by $\|X\|$. Thus if $X = (x_1, x_2, \dots, x_n)^t$,

then $\|X\|^2 = X^\theta X$

$= |x_1|^2 + \dots + |x_n|^2$

so that $\|X\| = \sqrt{|x_1|^2 + \dots + |x_n|^2}$. Clearly $\|X\|$ is a non-negative real number

and $\|X\| = 0 \iff X = 0$. If $X \in \mathbb{R}^n$, then $\|X\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$.

Definition 11.30. (Normal vector):

A vector X for which $\|X\| = 1$ is called a normal vector.

Definition 11.31. (Angle between two vectors):

If X and Y are two vectors in \mathbb{R}^n . $X = (x_1, x_2, \dots, x_n)^t$, $Y = (y_1, y_2, \dots, y_n)^t$ then the angle θ between X and Y is defined by $\cos \theta = \frac{X \cdot Y}{\|X\| \|Y\|}$

Definition 11.32. (Orthogonal vectors):

A set of vectors X_1, X_2, \dots, X_k in \mathbb{C}^n is said to be an orthogonal set of vectors if $X_i^\theta X_j = 0$ for $i \neq j; i, j = 1, 2, \dots, k$.

Definition 11.33. (Orthonormal vectors):

A set of vectors X_1, X_2, \dots, X_k in \mathbb{C}^n is said to be an orthonormal set of vectors if

- (i) $\|X_i\|=1, i = 1, 2, \dots, k$
- (ii) $X_i^\theta X_j = 0, i \neq j, i, j = 1, 2, \dots, k$.

In particular $X, Y \in \mathbb{R}^n$ are orthogonal if $X.Y = 0$. A set of vectors $X_1, X_2, \dots, X_n \in \mathbb{R}^n$ is said to be orthonormal if

- (i) $X_i.X_i = 1$ for all $i = 1, 2, \dots, k$
- (ii) $X_i.X_j = 0$ for $i \neq j; i, j = 1, 2, \dots, k$.

Theorem 11.27. Let A be a $m \times n$ matrix whose columns are orthonormal, and $X, Y \in \mathbb{R}^n$. Then

- (i) $\|AX\| = \|X\|$
- (ii) $(AX.AY) = (X.Y)$
- (iii) $(AX.AY) = 0$ if and only if $X.Y = 0$
- (iv) angle between AX and AY is the same as the angle between X and Y .

Proof: Let $A = [c_1, c_2, \dots, c_n], c_i \in \mathbb{R}^m$. Then

$$c_i.c_j = \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j. \end{cases} \dots (1)$$

as the columns form an orthonormal set. $A^t A = \begin{pmatrix} c_1^t \\ c_2^t \\ \vdots \\ c_n^t \end{pmatrix} (c_1 \ c_2 \ \dots \ c_n)$

$$= \begin{pmatrix} c_1^t c_1 & c_1^t c_2 & \dots & \dots & c_1^t c_n \\ c_2^t c_1 & c_2^t c_2 & \dots & \dots & c_2^t c_n \\ * & * & * & * & * \\ c_n^t c_1 & c_n^t c_2 & \dots & \dots & c_n^t c_n \end{pmatrix} = I_n \text{ using (1)}$$

$\therefore A^t A = I \dots (2)$

(i) $\|AX\|^2 = (AX)^t(AX)$
 $= X^t A^t AX$
 $= X^t I X$, using (2)
 $= X^t X$
 $= \|X\|^2$
 $\therefore \|AX\|^2 = \|X\|^2$
 $\Rightarrow \|AX\| = \|X\|$, as norm function is non-negative.

(ii) $AX.AY = (AX)^t(AY)$
 $= X^t A^t AY$
 $= X^t Y \dots$ Using(2)
 $= X.Y$
 $\therefore AX.AY = X.Y$

(iii) Follows from (ii).

(iv) If θ is the angle between AX and AY , then
 $\cos\theta = \frac{AX \cdot AY}{\|AX\| \|AY\|} = \frac{X \cdot Y}{\|X\| \|Y\|}$ using (i) and (ii)
 = cosine of angle between X and Y
 \therefore angle between X and Y is also θ . □

11.13 Length Preserving Mapping

Let A be an n -rowed square matrix and $X \in \mathbb{C}^n$. Then $Y = AX \in \mathbb{C}^n$. Thus multiplication by A defines a mapping from \mathbb{C}^n to \mathbb{C}^n . We are interested in knowing the condition on A so that lengths are preserved. This is shown in the following theorem.

Theorem 11.28. *A necessary and sufficient condition for a mapping $Y = AX$ to preserve lengths is that $A^\theta A = I$*

Proof:

$$\begin{aligned} Y &= AX \\ \Rightarrow Y^\theta &= X^\theta A^\theta \\ \therefore Y^\theta Y &= X^\theta A^\theta AX \\ \Rightarrow Y^\theta Y &= X^\theta (A^\theta A) X \end{aligned}$$

The condition is necessary.

Suppose the length is preserved, i.e. $X^\theta X = Y^\theta Y$.

Then (1) $\Rightarrow X^\theta X = X^\theta A^\theta AX$
 $\Rightarrow X^\theta (A^\theta A - I)X = 0$, for all $X \in \mathbb{C}^n$
 $\Rightarrow X^\theta BX = 0$, where $B = A^\theta A - I$.

Let $B = (b_{ij})_{n \times n}$, $X = (x_1, \dots, x_n)^t$

Then $\sum_{i,j=1}^n \bar{x}_i b_{ij} x_j = 0$
 $\Rightarrow \sum_{j=1}^n \sum_{i=1}^n b_{ij} \bar{x}_i x_j = 0 \dots (2)$

Taking $X = e_i = (0, 0, \dots, 1, 0, \dots, 0)$ here 1 is in i th position
 for $i = 1, 2, \dots, n$; (2) gives

$b_{ii} = 0$ (3)

Taking $X = e_k + e_l$, for $k, l = 1, 2, \dots, n; k \neq l$

(2) $\Rightarrow b_{kl} + b_{lk} = 0 \dots (4)$

Taking $X = e_k - e_l$, for $k, l = 1, 2, \dots, n; k \neq l$

(2) $\Rightarrow b_{kl} - b_{lk} = 0 \dots (5)$

(4) and (5) $\Rightarrow b_{kl} = 0$, where $k, l = 1, 2, \dots, n; k \neq l$

This, together with (3) $\Rightarrow B = 0$

Hence $A^\theta A = I$.

Condition is sufficient.

Let $A^\theta A = I$

Substituting in (1), we get $Y^\theta Y = X^\theta X$

so that length is preserved. □

The mapping which preserves length is called a unitary mapping and the matrix associated with it is called a unitary matrix. Thus we have the following definition.

Definition 11.34. (Unitary Matrix):

A matrix P is said to be unitary if $P^\theta P = I$.

Clearly $P^\theta P = I$
 $\Rightarrow |P| \neq 0$
 $\Rightarrow P$ is non-singular
 $\Rightarrow P$ is invertible
 $P^\theta P = I \Rightarrow P^{-1} = P^\theta$

Thus, the inverse of a unitary matrix P is P^θ .

Definition 11.35. (Orthogonal matrix):

A real unitary matrix is called an orthogonal matrix, i.e., a real matrix A is orthogonal if $A^t A = I$. If A is an orthogonal matrix, then A is invertible and $A^{-1} = A^t$.

Theorem 11.29. The columns of a unitary matrix form an orthonormal set.

Proof: Let $P = [C_1 C_2 \dots C_n]$ be a unitary matrix, where C_1, C_2, \dots, C_n are the columns of P .

$$\begin{aligned} \therefore P^\theta P &= I \\ \Rightarrow \begin{pmatrix} \overline{C_1^t} \\ \overline{C_2^t} \\ \vdots \\ \overline{C_n^t} \end{pmatrix} (C_1 \ C_2 \ \dots \ C_n) &= I \\ \Rightarrow \begin{pmatrix} \overline{C_1^t} C_1 & \overline{C_1^t} C_2 & \dots & \dots & \overline{C_1^t} C_n \\ \overline{C_2^t} C_1 & \overline{C_2^t} C_2 & \dots & \dots & \overline{C_2^t} C_n \\ * & * & * & * & * \\ * & * & * & * & * \\ \overline{C_n^t} C_1 & \overline{C_n^t} C_2 & \dots & \dots & \overline{C_n^t} C_n \end{pmatrix} &= I \\ \Rightarrow \overline{C_i^t} C_j &= \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j. \end{cases} \end{aligned}$$

Hence C_1, C_2, \dots, C_n is an orthonormal set. \square

Corollary 11.30. The row vectors of a unitary matrix form an orthonormal set.

Corollary 11.31. The column vectors of an orthogonal matrix form an orthonormal set.

Corollary 11.32. The row vectors of an orthogonal matrix form an orthonormal set.

Problem 11.9. The matrix $\begin{pmatrix} i & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ 0 & \frac{-i}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$ is a unitary matrix. Moreover, the columns of this matrix forms an orthonormal set of vectors.

Solution: Let $A = \begin{pmatrix} i & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ 0 & \frac{-i}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$, so that $A^\theta = \begin{pmatrix} -i & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ 0 & \frac{-i}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$

Then $AA^\theta = I_3$, so that A is a unitary matrix. If $A = (C_1 \ C_2 \ C_3)$, then $S = \{C_1, C_2, C_3\}$ is such that

$$\|C_1\|^2 = |-i|^2 + 0^2 + 0^2 = 1$$

$$\therefore \|C_1\| = 1$$

Similarly $\|C_2\| = \|C_3\| = 1$.

Also $C_1^\theta C_2 = 0$. Similarly, $C_2^\theta C_3 = 0$, $C_1^\theta C_3 = 0$ can be checked. In fact

$$(-i \ 0 \ 0) \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} \end{pmatrix} = 0$$

Hence S forms an orthonormal set i.e., the columns of A form an orthonormal set.

Problem 11.10. The matrix $\begin{pmatrix} p & q & r \\ 0 & q & -2r \\ -p & q & r \end{pmatrix}$, where $p = \frac{1}{\sqrt{2}}$, $q = \frac{1}{\sqrt{3}}$, $r = \frac{1}{\sqrt{6}}$ is orthogonal. Moreover the columns form an orthonormal set.

Solution: Here $A = \begin{pmatrix} p & q & r \\ 0 & q & -2r \\ -p & q & r \end{pmatrix}$

$$AA^t = \begin{pmatrix} p & q & r \\ 0 & q & -2r \\ -p & q & r \end{pmatrix} \begin{pmatrix} p & 0 & -p \\ q & q & q \\ r & -2r & r \end{pmatrix}$$

$$= \begin{pmatrix} p^2 + q^2 + r^2 & q^2 - 2r^2 & -p^2 + q^2 + r^2 \\ q^2 - 2r^2 & q^2 + 4r^2 & q^2 - 2r^2 \\ -p^2 + q^2 + r^2 & q^2 - 2r^2 & p^2 + q^2 + r^2 \end{pmatrix}$$

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

$\therefore AA^t = I$ so that A is an orthogonal matrix.

If we write $A = (C_1 \ C_2 \ C_3)$

Then $\|C_1\|^2 = p^2 + p^2 = 1 \therefore \|C_1\| = 1$

$\|C_2\|^2 = 3q^2 = 1 \therefore \|C_2\| = 1$

$\|C_3\|^2 = 6r^2 = 1 \therefore \|C_3\| = 1$

$C_1^t \cdot C_2 = pq + 0(q) - pq = 0$

Similarly $C_i^t C_j = 0$, for $i \neq j$; $i, j = 1, 2, 3$

Hence the set $\{C_1, C_2, C_3\}$, consisting of the column vectors of A forms an orthonormal set.

Problem 11.11. If A is skew Hermitian then $(I - A)(I + A)^{-1}$ is unitary, assuming that $I + A$ is non-singular.

Solution: Since A is skew Hermitian, $\therefore A^\theta = -A \dots (1)$

Let $B = (I - A)(I + A)^{-1}$

$$\begin{aligned} \text{Then } B^\theta &= ((I - A)(I + A)^{-1})^\theta \\ &= ((I + A)^{-1})^\theta (I - A)^\theta, \therefore (AB)^\theta = B^\theta A^\theta \\ &= ((I + A)^\theta)^{-1} (I^\theta - A^\theta), \therefore (A^{-1})^\theta = (A^\theta)^{-1}, (X - Y)^\theta = X^\theta - Y^\theta \\ &= (I + A^\theta)^{-1} (I - A^\theta) \\ &= (I - A)^{-1} (I + A), \text{ using (1)} \end{aligned}$$

$$\begin{aligned}
B^\theta B &= (I - A)^{-1}(I + A)(I - A)(I + A)^{-1} \\
&= (I - A)^{-1}(I - A^2)(I + A)^{-1} \\
&= (I - A)^{-1}(I - A)(I + A)(I + A)^{-1} \\
&= II \\
&= I
\end{aligned}$$

$\therefore B^\theta B = I$ so that B is unitary.

11.14 Exercise

1. Show that the matrix $\frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & -1 & -\sqrt{3} \\ \sqrt{2} & 2 & 0 \\ \sqrt{2} & -1 & \sqrt{3} \end{pmatrix}$ is an orthogonal matrix.

2. Show that the matrix A is an orthogonal matrix, where

$$(i) \quad A = \begin{pmatrix} \sin \theta & \cos \theta \\ \cos \theta & -\sin \theta \end{pmatrix}$$

$$(ii) \quad A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$(iii) \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$(iv) \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \sin \theta & \cos \theta \\ 0 & 0 & \cos \theta & -\sin \theta \end{pmatrix}$$

3. For what values of x are the vectors $u = (1, 1, -2)^t$ and $v = (x, -1, 2)^t$ orthogonal?

4. For what values of x and y is the set $\{u, v\}$ an orthonormal set, where $u = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})^t$ and $v = (x, \frac{1}{\sqrt{2}}, -y)^t$?

5. Let $X_1 = \frac{1}{\sqrt{2}}(1, -1, 0)^t$, $X_2 = \frac{1}{\sqrt{3}}(1, 1, 1)^t$. Find a vector X_3 so that the matrix $A = (X_1 \ X_2 \ X_3)$ is an orthogonal matrix.

6. Let $X_1 = \frac{1}{\sqrt{2}}(1, 1, 0, 0)^t$, $X_2 = \frac{1}{\sqrt{3}}(1, -1, 1, 0)^t$, $X_3 = \frac{1}{\sqrt{42}}(-1, 1, 2, 6)^t$. Find a vector X_4 such that the matrix $A = (X_1 \ X_2 \ X_3 \ X_4)$ is an orthogonal matrix.

7. Show that the matrix $\begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix}$ is unitary if and only if $a^2 + b^2 + c^2 + d^2 = 1$ where $x = a + ib$, $y = -c + id$.

8. Show that every unitary matrix is normal. Is the converse true?

9. Prove that
- the tranjugate of a unitary matrix is unitary.
 - the inverse of a unitary matrix is unitary.
 - the product of two unitary matrices is a unitary matrix.
10. Prove that
- the transpose of an orthogonal matrix is orthogonal.
 - the inverse of an orthogonal matrix is orthogonal.
 - the product of two orthogonal matrices is orthogonal.
11. Let $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{-2}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$, $X = \begin{pmatrix} -\sqrt{2} \\ 3 \end{pmatrix}$. Verify that
- columns of P are orthogonal.
 - $P^t P = I_2$
 - $\|PX\| = \|X\|$.
12. Prove that the matrix $A = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_3 & x_1 & x_2 \\ x_2 & x_3 & x_1 \end{pmatrix}$ is orthogonal if and only if x_1, x_2, x_3 are the roots of the equation $x^3 + x^2 + p = 0$ or $x^3 - x^2 + q = 0$, where p and q are any real numbers. What happens to A when $q = 0$.
13. Let A be any 3-rowed orthogonal matrix. Prove that the mapping of $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by $X \rightarrow AX$ preserves the angle between any two vectors in \mathbb{R}^3 .
14. If A is a skew symmetric matrix then show that $(I - A)(I + A)^{-1}$ is an orthogonal matrix (assuming that $I + A$ is singular).
15. If A, B, C are the vertices of a right-angled triangle in \mathbb{R}^2 , show that the transformation $X \rightarrow PX$ transforms triangle ABC to a right-angled triangle, where $P = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.
16. If A_k is a $k \times k$ orthogonal matrix, prove that
- $$A_{k+1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_k & \\ \vdots & & & \\ 0 & & & \end{pmatrix}$$
- is an orthogonal matrix. (This provides a way to obtain higher order orthogonal matrices, from a given orthogonal matrix).
17. If A_k is a $k \times k$ unitary matrix, then A_{k+1} and B_{k+1} are also unitary matrices, where $A_{k+1} = \begin{pmatrix} i & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_k & \\ \vdots & & & \\ 0 & & & \end{pmatrix}$ and $B_{k+1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_k & \\ \vdots & & & \\ 0 & & & \end{pmatrix}$.

18. If $X, Y \in \mathbb{R}^n$, define the distance between X and Y as $\|X - Y\|$ i.e., $dist(X, Y) = \|X - Y\|$. Prove that
- (i) $dist(X, Y) = \|X\|^2 + \|Y\|^2 - 2X \cdot Y$
 $= \|X\|^2 + \|Y\|^2 - 2\|X\|\|Y\|\cos\theta$
- (ii) Deduce from (a) that if A is an orthogonal matrix then $dist(AX, AY) = dist(X, Y)$.

11.15 Eigenvalues and Eigenvectors

Given a square matrix A , $X \rightarrow AX$ defines a mapping on \mathbb{C}^n . For this mapping, some vectors are special, as these vectors are mapped to a collinear vector, i.e. $X \rightarrow \alpha X$ for some scalar α . These vectors play a very important role in linear algebra so we study them in this chapter.

Definition 11.36. If A is a square matrix over \mathbb{C} , then a non zero vector $X \in \mathbb{C}^n$ is said to be an eigenvector of A , if there exists $\lambda \in \mathbb{C}$ such that $AX = \lambda X$. λ is called an eigenvalue (or characteristic root or latent root) of A corresponding to the eigenvector (or characteristic vector or latent vector) X .

If $AX = \lambda X$, then $A(kX) = \lambda(kX)$, so that an eigenvector corresponding to an eigenvalue is not unique.

Also, corresponding to two different eigenvalues, there can not be the same eigenvector, for if A is a matrix with distinct eigenvalues λ_1, λ_2 and X is an eigenvector, associated with λ_1 as well as λ_2 , then

$$\begin{aligned} AX &= \lambda_1 X \\ AX &= \lambda_2 X \\ \therefore (\lambda_1 - \lambda_2)X &= 0 \\ \Rightarrow X &= 0 \\ \text{as } \lambda_1 &\neq \lambda_2 \end{aligned}$$

which is not possible as X is an eigenvector.

The question that now arises is: given a square matrix A , how do we go about finding an eigenvalue and a corresponding eigenvector.

Determination of eigenvalues and eigenvectors

Suppose A is a square matrix and λ an eigenvalue of A then there exists a non-zero vector X such that

$$\begin{aligned} AX &= \lambda X \\ \Rightarrow (A - \lambda I)X &= 0 \end{aligned}$$

since $X \neq 0$, therefore $A - \lambda I$ is singular, so that $|A - \lambda I| = 0$ since A is an $n \times n$ matrix

$\therefore |A - xI| = 0$ is a polynomial of degree n in x , and is satisfied by $x = \lambda$. Hence any eigenvalue of A satisfies the n^{th} degree equation

$$|A - xI| = 0$$

Thus an $n \times n$ matrix has n eigenvalues. They may be real or complex, distinct or repeated. $|A - xI|$ is called the characteristic polynomial of A . The equation $|A - xI| = 0$ is called the characteristic equation of A and its roots are called the characteristic roots (or eigenvalues or latent roots) of A . A characteristic (or eigen)vector $X \neq 0$ associated with a characteristic root λ is a non-zero

solution of $(A - \lambda I)X = 0$. The set of all vectors X satisfying $(A - \lambda I)X = 0$ is called the eigenspace of A associated with the eigenvalue λ . This eigenspace may contain one or more linearly independent vectors X .

Let A be an n -rowed square matrix with characteristic polynomial

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} \dots + a_0$$

$$\therefore |A - xI| = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} \dots + a_0$$

putting $x = 0$, we get

$$|A| = a_0$$

The characteristic equation of A is

$$a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} \dots + a_0 = 0, \text{ where } a_n = (-1)^n$$

Since this is an equation of degree n in x , \therefore it has n roots. Product of roots = $a_0/a_n = (-1)^n a_0$

$$\therefore \text{Product of the characteristic roots} = (-1)^n |A|$$

Example 11.21. Find the eigenvalues and the corresponding eigenvectors of the matrix

$$\begin{pmatrix} 2 & 1 & 0 \\ 9 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}. \text{ Also, normalize the eigenvectors.}$$

Solution: Let $A = \begin{pmatrix} 2 & 1 & 0 \\ 9 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

The characteristic equation of A is $|A - xI| = 0$

$$\begin{vmatrix} 2-x & 1 & 0 \\ 9 & 2-x & 1 \\ 0 & 0 & 2-x \end{vmatrix} \Rightarrow (2-x)[(2-x)^2 - 9] = 0$$

$$\Rightarrow (x-2)(x-5)(x+1) = 0$$

$$\Rightarrow x = 2, 5, -1.$$

Hence the eigenvalues of A are 2, 5, -1.

Let us now find the eigenvectors corresponding to the different eigenvalues.

If $X = (x_1, x_2, x_3)$ is an eigenvector corresponding to eigenvalue λ , then

$$(A - \lambda I)X = 0$$

$$\Rightarrow \begin{pmatrix} 2-\lambda & 1 & 0 \\ 9 & 2-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{pmatrix} X = 0 \dots (1)$$

Let $\lambda = 2$. Equation (1) gives

$$\begin{pmatrix} 0 & 1 & 0 \\ 9 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X = 0$$

$$\Rightarrow x_2 = 0$$

$$9x_1 + x_3 = 0$$

$$\Rightarrow x_1 = -\frac{1}{9}k$$

$x_3 = k$ where k is any real number.

Taking $k = 9$, $X_1(-1, 0, 9)^t$ is an eigenvector corresponding to $\lambda = 2$. $\|X_1\| = \sqrt{82}$

$\therefore Y_1 = \frac{1}{\sqrt{82}}(-1 \ 0 \ 9)$ is such that $\|Y_1\| = 1$. Y_1 is the normalised eigen vector associated with $\lambda = 2$.

Let $\lambda = 5$, Equation (1) gives

$$\begin{pmatrix} -3 & 1 & 0 \\ 9 & -3 & 1 \\ 0 & 0 & -3 \end{pmatrix} X = 0$$

$$\Rightarrow \begin{aligned} -3x_1 + x_2 &= 0 \\ 9x_1 - 3x_2 + x_3 &= 0 \\ -3x_3 &= 0 \end{aligned}$$

Thus $\begin{aligned} x_1 &= \frac{1}{3}k \\ x_2 &= k \\ x_3 &= 0 \end{aligned}$ where k is any real number.

Take $k = 3$
 $\therefore X_2 = (1 \ 3 \ 0)^t$ is an eigenvector corresponding to $\lambda = 5, \|X_2\| = \sqrt{10}$
 $Y_2 = \frac{1}{\sqrt{10}}(1 \ 3 \ 0)^t$ is such that $\|Y_2\| = 1$

Let $\lambda = -1$, Equation (1) gives

$$\begin{pmatrix} 3 & 1 & 0 \\ 9 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} X = 0$$

$$\begin{aligned} 3x_1 + x_2 &= 0 \\ 9x_1 + 3x_2 + x_3 &= 0 \\ 3x_3 &= 0 \end{aligned}$$

Hence $\begin{aligned} x_1 &= -\frac{1}{3}k \\ x_2 &= k \\ x_3 &= 0, \text{ where } k \text{ is any real number.} \end{aligned}$

Take $k = 3$. Then, $X_3 = (-1 \ 3 \ 0)^t$ is an eigenvector corresponding to $\lambda = -1$ and $\|X_3\| = \sqrt{10}$.

$Y_3 = \frac{1}{\sqrt{10}}(-1 \ 3 \ 0)^t$ is such that $\|Y_3\| = 1$

The eigenvalues are 2,5 and -1 and the corresponding normalised eigenvectors are $\frac{1}{\sqrt{82}}(-1, 0, 9)^t, \frac{1}{\sqrt{10}}(1 \ 3, 0)^t$ and $\frac{1}{\sqrt{10}}(-1, 3, 0)^t$

Theorem 11.33. *If A is an $n \times n$ matrix then A and A^t have the same characteristic roots.*

Proof: Let λ be any complex number then $(A - \lambda I)^t = (A^t - \lambda I)$

So that $\begin{aligned} |A - \lambda I| &= |(A - \lambda I)^t| \\ &= |A^t - \lambda I| \end{aligned}$

Hence $\begin{aligned} |A - \lambda I| &= 0 \\ \Leftrightarrow |A^t - \lambda I| &= 0 \end{aligned}$

Hence A and A^t have the same characteristic equation and therefore the same characteristic roots. □

Theorem 11.34. *The characteristic roots of a triangular matrix are the diagonal elements.*

Proof: Let A be a triangular matrix. If A is upper triangular, then $A = (a_{ij})_{n \times n}$ where $a_{ij} = 0$ if $i > j$
 Let λ be any complex number

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{1n} \\ 0 & a_{22} - \lambda & a_{2n} \\ 0 & 0 & a_{nn} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda).$$

The characteristic roots are the roots of the equation

$$|A - \lambda I| = 0$$

$$\Rightarrow \lambda = a_{11}, a_{22}, \dots, a_{nn}.$$

Hence the characteristic roots of A are the diagonal elements of A . If A is a lower triangular matrix, it can be similarly proved that its characteristic roots are the diagonal elements. \square

Corollary 11.35. *The characteristic roots of a diagonal matrix are the diagonal elements.*

Corollary 11.36. *The characteristic roots of a n -rowed scalar matrix is a diagonal element, with multiplicity n .*

Theorem 11.37. *The characteristic roots of A^θ are the conjugates of the characteristic roots of A , and conversely.*

Proof: Let λ any complex number

$$\begin{aligned} \overline{(A - \lambda I)} &= \overline{A} - \overline{\lambda}I, & \text{as } \overline{I} &= I \\ \therefore (A - \lambda I)^\theta &= (\overline{A} - \overline{\lambda}I)^t \\ &= (\overline{A} - \overline{\lambda}I)^t \\ &= (\overline{A}^t) - \overline{\lambda}I^t \\ \text{Now } &= A^\theta - \overline{\lambda}I \end{aligned}$$

$$|(A - \lambda I)^\theta| = |A^\theta - \overline{\lambda}I| \quad \dots (1)$$

$$\begin{aligned} \text{Also } |(A - \lambda I)^\theta| &= |(\overline{A} - \overline{\lambda}I)^t| \\ &= |\overline{A} - \overline{\lambda}I| \\ &= |\overline{(A - \lambda I)}| \quad \dots (2) \end{aligned}$$

$$\text{Thus (1) and (2) } \Rightarrow |A^\theta - \overline{\lambda}I| = |A - \lambda I|$$

$$\text{Hence } |A - \lambda I| = 0$$

$$\Leftrightarrow |A^\theta - \overline{\lambda}I| = 0$$

So that λ is a characteristic root of A if and only if $\overline{\lambda}$ is a characteristic root of A^θ . \square

Theorem 11.38. *The characteristic roots of kA are k times the characteristic roots of A , where k is any complex number.*

Proof: If λ is any complex number,

$$\text{then } (kA - k\lambda I) = k(A - \lambda I)$$

$$\begin{aligned} \therefore |kA - k\lambda I| &= |k(A - \lambda I)| \\ &= k^n |A - \lambda I| \end{aligned}$$

$$\text{Thus } |A - \lambda I| = 0$$

$$\Leftrightarrow |kA - k\lambda I| = 0$$

So that if λ is a characteristic root of A , then $k\lambda$ is a characteristic root of kA . \square

Theorem 11.39. *If λ is a characteristic root of A then λ^p is a characteristic root of A^p .*

Proof: Let λ be a characteristic root of A , and X is a corresponding characteristic vector. Then

$$AX = \lambda X \quad (1)$$

We prove that $A^n X = \lambda^n X$, for all $n \in \mathbb{N}$ by induction on n .

The result clearly holds for $n = 1$. Let the result hold for $n = k$

$$\text{i.e. } A^k X = \lambda^k X$$

pre-multiplying by A , we get $AA^k X = A(\lambda^k X)$

$$\Rightarrow A^{k+1} X = \lambda^k (AX)$$

$$= \lambda^k (\lambda X)$$

$$= \lambda^{k+1} X$$

$$\therefore A^{k+1} X = \lambda^{k+1} X$$

so that the result holds for $n = k + 1$. Hence, by the principle of induction the result holds for all $n \in \mathbb{N}$

$$\therefore A^p X = \lambda^p X$$

So that λ^p is a characteristic root of A^p . □

Theorem 11.40. *The characteristic roots of a Hermitian matrix are real.*

Proof: Let A be a Hermitian matrix.

$$\therefore A = A^\theta \quad \dots (1)$$

Let λ be a characteristic root of A and X a corresponding characteristic vector.

Then

$$AX = \lambda X \quad \dots (2)$$

pre-multiplying by X^θ , we get

$$X^\theta AX = \lambda X^\theta X \quad \dots (3)$$

Taking transposed conjugate on both sides of (3), we get

$$(X^\theta AX)^\theta = (\lambda X^\theta X)^\theta$$

$$\Rightarrow (X^\theta)^\theta A^\theta (X^\theta)^\theta = \bar{\lambda} (X^\theta X)^\theta$$

$$\Rightarrow X^\theta AX = \bar{\lambda} X^\theta X \text{ using (1) and } (X^\theta)^\theta = X$$

$$\Rightarrow \lambda X^\theta X = \bar{\lambda} X^\theta X \text{ using (2)}$$

$$\Rightarrow (\lambda - \bar{\lambda}) X^\theta X = 0$$

Since $X \neq 0 \therefore X^\theta X \neq 0$, so that

$$\lambda - \bar{\lambda} = 0$$

$$\Rightarrow \lambda = \bar{\lambda}$$

$$\Rightarrow \lambda \text{ is real.} \quad \square$$

Corollary 11.41. *The characteristic roots of a real symmetric matrix are real.*

Proof: Since a real symmetric matrix is Hermitian, result follows by the above theorem. □

Corollary 11.42. *The characteristic roots of a skew Hermitian matrix are either zero or pure imaginary.*

Proof: Let A be a skew Hermitian matrix and λ a characteristic root of A . Then iA is a Hermitian matrix, and $i\lambda$ is a characteristic root of iA . By the above theorem, $i\lambda$ is real so that λ must be zero or purely imaginary. □

Corollary 11.43. *The characteristic roots of a real skew symmetric matrix are either zero or pure imaginary.*

Theorem 11.44. *The characteristic roots of a unitary matrix are of unit modulus.*

Proof: Let A be a unitary matrix. Then $AA^\theta = A^\theta A = I \dots (1)$

Let λ be a characteristic root of A and X a characteristic vector corresponding to λ . Then

$$AX = \lambda X \dots (2)$$

Taking transposed conjugate in (2) we get

$$X^\theta A^\theta = \bar{\lambda} X^\theta$$

$$\therefore X^\theta A^\theta AX = \bar{\lambda} X^\theta \lambda X$$

$$\Rightarrow X^\theta X = \lambda \bar{\lambda} X^\theta X$$

$$\Rightarrow X^\theta X(1 - |\lambda|^2) = 0$$

Since $X \neq 0 \therefore X^\theta X \neq 0$, so that

$$1 - |\lambda|^2 = 0$$

$$\Rightarrow |\lambda| = 1 \quad \square$$

Corollary 11.45. *The characteristic roots of an orthogonal matrix are of unit modulus.*

Theorem 11.46. *The characteristic vectors corresponding to distinct characteristic roots of a matrix are linearly independent.*

Proof: Let A be an n -rowed square matrix and $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_k$ are distinct characteristic roots of A . Let X_1, X_2, \dots, X_k be the corresponding characteristic vectors. Then

$$AX_i = \lambda_i X_i \quad \dots \quad \dots (1) \text{ for } i = 1, 2, 3 \dots k.$$

For $j \neq i$,

$$\begin{aligned} (A - \lambda_i I)X_j &= AX_j - \lambda_i IX_j \\ &= \lambda_j X_j - \lambda_i X_j, \text{ using (1)} \\ &= (\lambda_j - \lambda_i)X_j \end{aligned}$$

$$\therefore (A - \lambda_i I)X_j = (\lambda_j - \lambda_i)X_j \quad \dots (2)$$

Consider a linear relation of the type

$$c_1 X_1 + c_2 X_2 + c_3 X_3 + \dots + c_k X_k = 0 \quad \dots (3)$$

where c_i 's are scalars. Pre multiplying (3) by $(A - \lambda_2 I)$,

$$\text{we get } (A - \lambda_2 I)(c_1 X_1 + c_2 X_2 + c_3 X_3 + \dots + c_k X_k) = 0$$

$$\Rightarrow c_1 (A - \lambda_2 I)X_1 + c_2 (A - \lambda_2 I)X_2 + \dots + c_k (A - \lambda_2 I)X_k = 0$$

$$\Rightarrow c_1 (\lambda_1 - \lambda_2)X_1 + c_3 (\lambda_3 - \lambda_2)X_3 + \dots + c_k (\lambda_k - \lambda_2)X_k = 0$$

Note that in this relation the vector X_2 is missing. Proceeding in a similar way, and pre-multiplying by $(A - \lambda_3 I)$, $(A - \lambda_4 I)$, \dots , $(A - \lambda_k I)$ in succession we eliminate X_3, X_4, \dots, X_k in turn, and arrive at

$$c_1 (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \dots (\lambda_1 - \lambda_k)X_1 = 0$$

since λ_i 's are distinct and $X_1 \neq 0$

$$\therefore c_1 = 0$$

In a similar way we can, by pre-multiplying (3) by $(A - \lambda_1 I)$, $(A - \lambda_3 I)$, \dots , $(A - \lambda_k I)$, and eliminating in turn X_1, X_3, \dots, X_k , we obtain

$$c_2 (\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3) \dots (\lambda_2 - \lambda_k)X_2 = 0$$

giving $c_2 = 0$.

Proceeding in this way, we get

$$c_3 = c_4 = \dots = c_k = 0. \text{ Hence } X_1, X_2, \dots, X_k \text{ are linearly independent.} \quad \square$$

Theorem 11.47. *The characteristic vectors corresponding to two distinct characteristic roots of a Hermitian matrix are orthogonal.*

Proof: Let A be a Hermitian matrix. Then $A = A^\theta \dots (1)$

Let λ, μ be two distinct characteristic roots of A and X, Y the corresponding

characteristic vectors, Since A is Hermitian, therefore λ, μ are real. Then $AX =$

$$\lambda X \quad \dots (2)$$

$$AY = \lambda Y \quad \dots (3)$$

Pre-multiplying (2) by Y^θ we get

$$Y^\theta AX = \lambda Y^\theta X \quad \dots (4)$$

Taking transposed conjugate of (3), we get

$$Y^\theta A^\theta = \bar{\mu} Y^\theta$$

$$\Rightarrow Y^\theta A = \mu Y^\theta \text{ using (1) and } \mu \text{ is real.}$$

Post-multiplying by X , we get $Y^\theta AX = \mu Y^\theta X$

$$\Rightarrow \lambda Y^\theta X = \mu Y^\theta X \text{ using (4)}$$

$$\Rightarrow (\lambda - \mu) Y^\theta X = 0$$

$$\Rightarrow Y^\theta X = 0, \text{ as } \lambda \neq \mu$$

$$\Rightarrow X \text{ and } Y \text{ are orthogonal.} \quad \square$$

Corollary 11.48. *The characteristic vectors corresponding to distinct characteristic roots of a real symmetric matrix are orthogonal.*

Theorem 11.49. *If $\lambda_1, \lambda_2, \dots, \lambda_n$ are eigen values of an n -rowed matrix A , and X_1, X_2, \dots, X_n are the corresponding linearly independent eigenvectors, then*

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}, \text{ where } P = [X_1, X_2, \dots, X_n]$$

Proof: Since X_i is an eigenvector associated with the eigenvalue λ_i . $\therefore AX_i = \lambda_i X_i, i = 1, 2, \dots, n$

$$AP = A[X_1, X_2, \dots, X_n]$$

$$= [AX_1, AX_2, \dots, AX_n]$$

$$= [\lambda_1 X_1, \lambda_2 X_2, \dots, \lambda_n X_n]$$

$$= [X_1, X_2, \dots, X_n] \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = P \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Since the columns of P are linearly independent, therefore P^{-1} exists, so that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad \square$$

Definition 11.37. (Similar Matrices): *Let A be a square matrix. A square matrix B is said to be similar to A if there exists an invertible matrix P such that $B = P^{-1}AP$. The matrices A, B and P are assumed to be over the same field. We write $B \sim A$.*

The relation of similarity is an equivalence relation. Many properties of A are carried over to B . For instance,

- *They have the same determinant.*
- *They have the same characteristic equations.*
- *They have the same rank.*

- They have the same nullity.
- B is invertible if and only if A is invertible.

Thus, if a given matrix A is similar to a simpler matrix, say a diagonal matrix D , then those properties of A which are invariant under similarity can be studied for D rather than for A . In fact, if $A = P^{-1}DP$, for some diagonal matrix then $A^n = P^{-1}D^nP$. Since it is easy to calculate D^n , so the calculation of A^n is also simplified.

Definition 11.38. (Diagonalizable Matrix): A matrix is said to be diagonalizable if it is similar to a diagonal matrix.

Theorem 11.49 can be restated as follows:

Theorem 11.50. If an n -rowed square matrix has n linearly independent eigenvectors, then it is similar to a diagonal matrix.

It is important to note that n linearly independent eigenvectors may not always exist. If all the eigenvalues are distinct, linearly independent eigenvectors always exist, but when some of the eigenvalues are repeated, then n linearly independent eigenvectors may not always exist. A detailed discussion of this is not our aim here. An interested reader we refer to Algebra II by the same authors.

11.16 Cayley Hamilton Theorem and Its Applications

A very important and useful result regarding characteristic equation is the Cayley Hamilton Theorem. It is used to find inverse of a matrix and helps us in expressing a matrix polynomial of any degree in terms of a matrix polynomial of degree less than n , where n is the order of the matrix.

Theorem 11.51. (Cayley Hamilton theorem): Every square matrix satisfies its characteristic equation. Equivalently, if A is a square matrix, and $|A - \lambda I| = a_0 + a_1\lambda + \dots + a_n\lambda^n = 0$ is the characteristic equation of A , then $a_0I + a_1A + \dots + a_nA^n = 0$

Proof: The elements of $A - \lambda I$ are at most of degree one in λ , So that the elements of $\text{adj}(A - \lambda I)$ are at most of degree $(n - 1)$ in λ . If B_k is the matrix whose $(i, j)^{\text{th}}$ element is the coefficient of λ^k in the $(i, j)^{\text{th}}$ element of $\text{adj}(A - \lambda I)$ then, $\text{adj}(A - \lambda I)$ can be written as $B_0 + B_1\lambda + \dots + B_{n-1}\lambda^{n-1}$ where B_0, B_1, \dots, B_{n-1} are matrices of order n , and these elements depend upon the elements of A . Since

$$(A - \lambda I)\text{adj}(A - \lambda I) = |A - \lambda I|$$

$$(A - \lambda I)(B_0 + B_1\lambda + \dots + B_{n-1}\lambda^{n-1}) = (a_0 + a_1\lambda + \dots + a_n\lambda^n)I.$$

Equating coefficients of $\lambda^0, \lambda^1, \lambda^2, \dots, \lambda^n$ we get:

$$AB_0 = a_0I$$

$$AB_1 - IB_0 = a_1I$$

...

$$AB_{n-1} - IB_{n-2} = a_{n-1}I$$

$$-IB_{n-1} = a_nI$$

Pre-multiplying the above equation by $I, A, A^2, A^3, A^4, \dots, A^n$ respectively and adding we get,

$$0 = a_0I + a_1A + \dots + a_nA^n \text{ (as the terms on the left hand side cancel in pairs).}$$

Thus $\lambda = A$ satisfies the characteristic equation

$$a_0 + a_1\lambda + \dots + a_n\lambda^n = 0. \quad \square$$

Sometimes we state Cayley Hamilton theorem as:

Every Square matrix satisfies its characteristic equation.

Example 11.22. Verify that $A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$ satisfies its characteristic equation, and hence obtain A^{-1} .

Solution: The characteristic equation of A is $|A - \lambda I| = 0$

$$\begin{vmatrix} 2 - \lambda & -1 & 1 \\ -1 & 2 - \lambda & -1 \\ 1 & -1 & 2 - \lambda \end{vmatrix} = 0 \quad \dots (1)$$

$$\Rightarrow \lambda^3 + 6\lambda^2 + 9\lambda - 4 = 0$$

To verify that A satisfies (1), we have to verify that

$$A^3 - 6A^2 + 9A - 4I = 0 \quad \dots (2)$$

$$A^2 = \begin{pmatrix} 6 & -5 & 5 \\ -5 & 6 & 6 \\ 5 & -5 & 6 \end{pmatrix},$$

$$A^3 = \begin{pmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{pmatrix}$$

On substituting the values of I, A, A^2 and A^3 , it can be seen that (2) is satisfied.

To find A^{-1}

Pre-multiplying (2) by A^{-1} , we get

$$A^2 - 6A + 9I - 4A^{-1} = 0$$

$$\Rightarrow 4A^{-1} = A^2 - 6A + 9I$$

$$\therefore 4A^{-1} = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$$

$$\text{so that } A^{-1} = \frac{1}{4} \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$$

Example 11.23. Find A^3 and A^5 using Cayley Hamilton theorem, where $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$.

Solution: The characteristic equation of A is $|A - xI| = 0$.

$$\begin{vmatrix} 1 - x & 2 \\ 4 & 3 - x \end{vmatrix} = 0 \quad \dots (1)$$

By Cayley Hamilton theorem,

$$A^2 - 4A - 5I = 0 \quad \dots (2)$$

$$\therefore A^2 = 4A + 5I \quad \dots (3)$$

Pre-multiplying (3) by A , we get

$$\begin{aligned} A^3 &= 4A^2 + 5A \\ &= 4(4A + 5I) + 5A \quad \text{using (3)} \\ &= 21A + 20I \\ &= \begin{pmatrix} 41 & 42 \\ 84 & 83 \end{pmatrix} \end{aligned}$$

To obtain A^5 multiply (3) by A^3

$$\begin{aligned} \Rightarrow A^5 &= 4A^4 + 5A^3 \\ &= 521A + 520I \quad (\text{on using (3) repeatedly}) \\ &= \begin{pmatrix} 1041 & 1042 \\ 2084 & 2083 \end{pmatrix} \end{aligned}$$

Another way of obtaining this expression for A^5 as a linear polynomial in A is to divide x^5 by $x^2 - 4x - 5$

$$\begin{aligned} \text{Thus } x^5 &= (x^3 + 4x^2 + 21x + 104)(x^2 - 4x - 5) + 521x + 520 \\ \text{so that } A^5 &= (A^3 + 4A^2 + 21A + 104I)(A^2 - 4A - 5I) + 521A + 520I \\ &= 521A + 520I, \quad \text{using (2)} \\ &= \begin{pmatrix} 1041 & 1042 \\ 2084 & 2083 \end{pmatrix} \end{aligned}$$

Let A be a $n \times n$ matrix with characteristic equation

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 = 0 \quad \dots (1)$$

By Cayley Hamilton theorem, A satisfies the matrix equation

$$a_n A^n + a_{n-1} A^{n-1} + \dots + a_0 I = 0$$

so that

$$a_n A^n + a_{n-1} A^{n-1} + \dots + a_0 I = 0 \quad \dots (2)$$

We know that $|A| = a_0$, so that A^{-1} exists if $a_0 \neq 0$.

Pre-multiplying (2) by A^{-1} , we get,

$$\begin{aligned} a_n A^{n-1} + a_{n-1} A^{n-2} + \dots + a_0 A^{-1} &= 0 \\ \Rightarrow A^{-1} &= \frac{-1}{a_0} (a_n A^{n-1} + a_{n-1} A^{n-2} + \dots + a_1 I) \end{aligned}$$

Thus A^{-1} has been expressed as a polynomial in A of degree $n - 1$, and can be evaluated easily.

Suppose we want to evaluate a polynomial in A , namely

$$g(A) = b_0 A^m + b_1 A^{m-1} + \dots + b_m I$$

If

$$f(x) = 0 \quad \dots (1)$$

is the characteristic equation of A , then by Cayley Hamilton theorem,

$$f(A) = 0 \quad \dots (2)$$

Applying division algorithm to $f(x)$ and $g(x)$, there exist polynomials $q(x)$ and $r(x)$ such that

$$g(x) = f(x)q(x) + r(x), \text{ where } r(x) = 0 \text{ or } \deg(r(x)) < n. \text{ Replacing } x \text{ by } A,$$

$$\text{we get } g(A) = f(A)q(A) + r(A)$$

$$\Rightarrow g(A) = r(A), \text{ using (2)}$$

Thus $g(A)$ is expressible as a polynomial in A of degree at most $n-1$. This helps

us to evaluate $g(A)$, when $g(x)$ is a polynomial of any degree.

Example 11.24. If $A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$, find

(1) A^{-1} , if it exists.

(2) $A^7 - 7A^6 + 15A^5 - 12A^4 - A^3 + 8A^2 - 15A - 4I$

Solution: Characteristic equation A is

$$|A - xI| = 0$$

$$\Rightarrow \begin{vmatrix} 2-x & -1 & 1 \\ -1 & 2-x & -1 \\ 1 & -1 & 2-x \end{vmatrix} = 0$$

$$\Rightarrow x^3 - 6x^2 + 9x - 4 = 0$$

$$\text{let } f(x) = x^3 - 6x^2 + 9x - 4$$

By Cayley Hamilton Theorem,

$$A^3 - 6A^2 + 9A - 4I = 0$$

$$\Rightarrow f(A) = 0 \quad \dots (2)$$

In (1) constant term is non-zero, so that A^{-1} exists. Pre-multiplying (2) by A^{-1} and transposing, we get

$$4A^{-1} = A^2 - 6A + 9I$$

$$= \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$$

$$A^{-1} = \frac{1}{4} \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{pmatrix}$$

$$(2) \text{ Let } g(x) = x^7 - 7x^6 + 15x^5 - 12x^4 - x^3 + 8x^2 + 15x - 4.$$

Applying division algorithm to $f(x)$ and $g(x)$ there exists $q(x)$ and $r(x)$ such that

$$g(x) = f(x)q(x) + r(x) \quad (3) \text{ where } q(x) = x^4 - x^3 + x + 1$$

$$r(x) = 5x^2 + 10x$$

In (3), replacing x by A the relation still holds.

$$\therefore g(A) = f(A)q(A) + r(A)$$

$$= r(A), \text{ using (2)}$$

$$= 5A^2 + 10A$$

$$= 5 \begin{pmatrix} 6 & -5 & -5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{pmatrix} + 10 \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 50 & -35 & 35 \\ -35 & 50 & -35 \\ 35 & -35 & 50 \end{pmatrix}$$

11.17 Solved Problems

Problem 11.12. Find the eigenvalues and the corresponding eigenvectors of

$$\text{the matrix } \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 5 \end{pmatrix}$$

$$\textbf{Solution:} \text{ Let } A = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 5 \end{pmatrix}$$

Characteristic equation of A is $|A - xI| = 0$

$$\Rightarrow \begin{vmatrix} 2-x & 1 & 2 \\ 0 & 2-x & 3 \\ 0 & 0 & 5-x \end{vmatrix} = 0$$

$$\Rightarrow (2-x)^2(5-x) = 0$$

$$\Rightarrow x = 2, 2, 5.$$

Thus the eigenvalues are $\lambda_1 = 5, \lambda_2 = 2, \lambda_3 = 2$.

We now obtain the eigenvectors corresponding to each eigenvalue, noting that two of the eigenvalues are the same. If $X = (x_1, x_2, x_3)^t$ is an eigenvector corresponding to the eigenvalue λ , then $(A - \lambda I)X = 0$

$$\Rightarrow \begin{pmatrix} 2 - \lambda & 1 & 2 \\ 0 & 2 - \lambda & 3 \\ 0 & 0 & 5 - \lambda \end{pmatrix} X = 0$$

$$\text{Substituting } \lambda = 5 \text{ in (1), we get } \begin{pmatrix} -3 & 1 & 2 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{pmatrix} X = 0$$

$$\Rightarrow -3x_1 + x_2 + 2x_3 = 0$$

$$\quad -3x_2 + 3x_3 = 0$$

$$\Rightarrow \quad x_1 = k$$

$$\quad x_2 = k$$

$$\quad x_3 = k \text{ where } k \text{ is any real number.}$$

Thus $X_1 = (1, 1, 1)^t$ is an eigenvector corresponding to $\lambda_1 = 5$.

Now, $\lambda_2 = 2 = \lambda_3$

So, substituting $\lambda = 2$ in (1), we get

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 3 \end{pmatrix} X = 0$$

$$\Rightarrow \quad x_2 + 2x_3 = 0$$

$$\quad 3x_3 = 0$$

$$\therefore x_1 = k, x_2 = 0, x_3 = 0$$

Taking $k = 1$.

$\therefore X_2 = (1, 0, 0)^t$ is an eigenvector corresponding to the repeated eigenvalues 2.

Note that in above example, though 2 is a repeated eigenvalue, there is only one eigenvector corresponding to it. There are only 2 linearly independent eigenvectors, X_1 and X_2 .

Problem 11.13. Find the eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Solution: Let $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$

The characteristic equation of A is $|A - xI| = 0$

$$\Rightarrow \begin{vmatrix} 2-x & 0 & 1 \\ 0 & 3-x & 0 \\ 1 & 0 & 2-x \end{vmatrix} = 0$$

$$\Rightarrow (3-x)[(2-x)^2 - 1] = 0$$

$$\Rightarrow (3-x)(3-x)(x-1) = 0$$

$$\Rightarrow x = 1, 3, 3.$$

Thus, the eigenvalues are $\lambda_1 = 1, \lambda_2 = \lambda_3 = 3$.

If $X = (x_1, x_2, x_3)^t$ is an eigenvector corresponding to the eigenvalue λ , then $(A - \lambda I)X = 0$

$$\Rightarrow \begin{pmatrix} 2-\lambda & 0 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{pmatrix} X = 0 \quad \dots (1)$$

For $\lambda_1 = 1$, Equation (1) gives $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} X = 0$

$$\Rightarrow x_1 + x_3 = 0$$

$$2x_2 = 0$$

$$x_1 + x_3 = 0$$

$$\therefore x_1 = -k, x_2 = 0, x_3 = k$$

where k is any real number. $\therefore X = k(-1, 0, 1)^t$ is an eigenvector corresponding to $\lambda_1 = 1$.

For $\lambda_2 = \lambda_3 = 3$, Equation (1) gives

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

$$-x_1 + x_3 = 0$$

$$0 = 0$$

$$x_1 - x_3 = 0$$

$\therefore x_1 = x_3 = k_1$ (say). Also x_2 is arbitrary so that $X = (k_1, k_2, k_1)^t$ is an eigenvector. Taking $k_1 = 1, k_2 = 0$, we get $X_1 = (1, 0, 1)^t$ as an eigenvector. Taking $k_1 = 0, k_2 = 1$, we get $X_2 = (0, 1, 0)$. Also X_1, X_2 are linearly independent, thus corresponding to $\lambda_2 = 3$ we get two linearly independent eigenvectors X_1, X_2 .

Problem 11.14. Show that the square matrices A and $P^{-1}AP$ have the same eigenvalues, where P is an arbitrary non-singular matrix.

Solution: Let λ be any complex number.

$$\text{Then } (P^{-1}AP - \lambda I) = (P^{-1}AP - \lambda P^{-1}P) = P^{-1}(A - \lambda I)P$$

$$\begin{aligned} \therefore |P^{-1}AP - \lambda I| &= |P^{-1}(A - \lambda I)P| \\ &= |P^{-1}| |A - \lambda I| |P| \\ &= |P|^{-1} |A - \lambda I| |P| \\ &= |A - \lambda I|. \text{ Thus } |A - \lambda I| = 0 \end{aligned}$$

$$\Leftrightarrow |P^{-1}AP - \lambda I| = 0$$

so that A and $P^{-1}AP$ have the same characteristic equation, and \therefore the same characteristic roots.

Problem 11.15. If λ is a characteristic root of a non-singular matrix A , then $\frac{|A|}{\lambda}$ is a characteristic root of $\text{adj}A$.

Solution: Since A is non-singular, $\therefore 0$ is not a characteristic root of A . Let λ be a characteristic root of A . Then $\lambda \neq 0$. Also, there exists a non-zero vector X such that

$$AX = \lambda X$$

$$\text{Then } (\text{adj}A)AX = (\text{adj}A)(\lambda X)$$

$$\Rightarrow |A|IX = \lambda(\text{adj}A)X \therefore (\text{adj}A)A = I$$

$$\Rightarrow \frac{|A|}{\lambda}X = (\text{adj}A)X (\because \lambda \neq 0)$$

$$\Rightarrow (\text{adj}A)X = \frac{|A|}{\lambda}X$$

$$\Rightarrow \frac{|A|}{\lambda} \text{ is a characteristic root of } (\text{adj}A)$$

Problem 11.16. If A is a square matrix of order n , prove that the trace of A is the sum of eigenvalues of A .

Proof: Let $A = (a_{ij})_{n \times n}$. The characteristic equation of A is $|A - xI| = 0$

$$\Rightarrow \begin{vmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{vmatrix} = 0 \quad \dots (1)$$

Expanding (1) along the I^{st} row, the co factors of a_{12}, \dots, a_{1n} are of degree at most $n - 2$ in x and the cofactor of $a_{11} - x$ is of degree $n - 1$ in x . Thus, in the expansion of (1) the term of x^{n-1} is obtained only from $(a_{11} - x)(a_{22} - x) \dots (a_{nn} - x)$.

$$\begin{aligned} \text{Thus coefficient of } x^{n-1} &= (-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn}) \\ &= (-1)^{n-1} \text{trace} A \end{aligned}$$

$$\begin{aligned} \text{coefficient of } x^n &= (-1)^n. \text{ Sum of the eigenvalues} = \text{Sum of the roots of the} \\ \text{characteristic equation of } A &= -\frac{\text{Coefficient of } x^{n-1}}{\text{Coefficient of } x^n} \\ &= -\frac{(-1)^{n-1} \text{trace} A}{(-1)^n} \\ &= \text{trace} A \end{aligned}$$

Problem 11.17. Let A and B be n -rowed matrices. If $I - AB$ is invertible, then $I - BA$ is invertible and $(I - BA)^{-1} = I + B(I - AB)^{-1}A$

Solution: Since $I - AB$ is invertible, therefore there exists a matrix C such that $(I - AB)C = C(I - AB) = I \quad \dots (1)$

$$\text{Now } (I - AB)C = I$$

$$\Rightarrow C - (AB)C = I$$

$$\Rightarrow I + (AB)C = C$$

$$\Rightarrow BIA + B(AB)CA = BCA$$

$$\Rightarrow BA = BCA - BABCA$$

$$\Rightarrow -BA + BCA - BABCA = 0$$

$$\Rightarrow -BA(I + BCA) + I + BCA = I \quad \dots (2)$$

$$\Rightarrow (I - BA)(I + BCA) = I$$

$$\text{Similarly } (2) \Rightarrow (I + BCA)(I - BA) = I$$

$$\text{Thus } (I + BCA)(I - BA) = (I - BA)(I + BCA) = I$$

$$\text{Hence } I - BA \text{ is invertible and } (I - BA)^{-1} = I + BCA$$

$$= I + B(I - AB)^{-1}A$$

Remark 11.4. From this we can say that $|I - AB| = 0 \Leftrightarrow |I - BA| = 0$. Consequently we have the following problem

Problem 11.18. If A and B are n -rowed matrices then AB and BA have the same characteristic roots.

Solution: Let λ be a characteristic root of AB . Two cases arise.

Case 1. $\lambda = 0$

$$\text{Then } |AB| = 0$$

$$\Leftrightarrow |A||B| = 0$$

$$\Leftrightarrow |B||A| = 0$$

$$\Leftrightarrow |BA| = 0$$

Thus 0 is a characteristic root of $AB \Leftrightarrow 0$ is a characteristic root of BA .

$$\begin{aligned}
& \text{Case 2. } \lambda \neq 0 \\
& |AB - \lambda I| = 0 \\
\Leftrightarrow & \left| \frac{1}{\lambda} AB - I \right| = 0 \\
\Leftrightarrow & \left| \left(\frac{1}{\lambda} A \right) B - I \right| = 0 \\
\Leftrightarrow & \left| B \left(\frac{1}{\lambda} A \right) - I \right| = 0 \\
\Leftrightarrow & \left| \frac{1}{\lambda} BA - I \right| = 0 \\
\Leftrightarrow & |BA - \lambda I| = 0
\end{aligned}$$

Thus λ is a characteristic root of $AB \Leftrightarrow \lambda$ is a characteristic root of BA .

Cases 1 and 2 prove that AB and BA have the same characteristic roots.

Problem 11.19. Let A be a 2×2 real symmetric matrix. Prove that A is similar to a diagonal matrix.

Solution: Since A is symmetric, let

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

characteristic equation of A is $|A - xI| = 0$

$$\begin{vmatrix} a-x & b \\ b & d-x \end{vmatrix} = 0$$

$$\Rightarrow (a-x)(d-x) - b^2 = 0$$

$$x^2 - (a+d)x + ad - b^2 = 0$$

$$\begin{aligned}
\text{Discriminant} &= (a+d)^2 - 4(ad - b^2) \\
&= (a-d)^2 + 4b^2 \\
&\geq 0
\end{aligned}$$

Hence roots are real. Roots are equal when $a - d = 0$ and $b = 0$

$$\therefore A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

So that A itself is a diagonal matrix. Let the roots be distinct, say λ_1, λ_2 . Let X_1, X_2 be corresponding eigenvectors. Then X_1, X_2 are linearly independent so that $P = [X_1 \ X_2]$ is an invertible matrix.

$$AX_1 = \lambda_1 X_1, AX_2 = \lambda_2 X_2$$

$$\text{Now } AP = A[X_1 \ X_2] = [AX_1 \ AX_2]$$

$$= [\lambda_1 X_1 \ \lambda_2 X_2]$$

$$= [X_1 \ X_2] \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$= PD, \text{ where } D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$AP = PD$$

$$\Rightarrow P^{-1}AP = D$$

$\Rightarrow A$ is similar to a diagonal matrix.

Problem 11.20. Show that if the two characteristic roots of a Hermitian matrix of order 2 are equal, then the matrix must be a scalar multiple of the unit matrix.

Solution: Let $A = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}$ be a given 2×2 Hermitian matrix.

The characteristic equation of A is

$$\begin{vmatrix} a-x & b \\ \bar{b} & c-x \end{vmatrix} = 0$$

$$\Rightarrow (a-x)(c-x) - b\bar{b} = 0$$

$$\Rightarrow x^2 - (a+c)x + ac - b\bar{b} = 0 \quad \dots (1)$$

Since the characteristic roots are equal

$$\begin{aligned} \therefore \quad & \text{Disc. of (1)} = 0 \\ \Rightarrow & (a+c)^2 - 4(ac - b\bar{b}) = 0 \\ \Rightarrow & (a-c)^2 + 4b\bar{b} = 0 \end{aligned}$$

Since the sum of two non-negative terms is zero, \therefore each term = 0 so that $(a-c)^2 = 0$ and $b\bar{b} = 0$

$$\begin{aligned} \Rightarrow & a = c, b = 0 \\ \therefore A &= \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Hence proved.

Problem 11.21. Prove that the coefficient in the characteristic equation of a real skew symmetric matrix are non-negative.

Solution: Let A be a real skew symmetric matrix whose characteristic equation is

$$a_0x^n + a_1x^{n-1} + \dots + a_n = 0 \dots (1)$$

Since the characteristic roots of a real symmetric matrix are either zero or pure imaginary, therefore 0 is the only real root of (1).

If 0 is a root of (1) of multiplicity k , then x^k is a factor of (1). Then (1) can be rewritten as

$$x^k(a_0x^{n-k} + a_1x^{n-k-1} + \dots + a_{n-k}) = 0.$$

Therefore $a_{n-k+1} = a_{n-k+2} = \dots = a_n = 0$

$a_0x^{n-k} + a_1x^{n-k-1} + \dots + a_{n-k} = 0$, has no real roots

Therefore by Descartes Rule of Sign, there is no change of sign, so that a_0, a_1, \dots, a_{n-k} all have the same sign. Hence they all can be taken to be positive.

Thus the coefficients of (1) are positive or zero i.e. they are non-negative.

Problem 11.22. For a skew symmetric matrix of order n , show that

- (i) if λ is a characteristic root of A , so is $-\lambda$.
- (ii) every non-zero characteristic root of A^2 , occurs with even multiplicity.
- (iii) $|A^2 - xI|$ is a perfect square, if n is even, and $|A^2 - xI| = x[f(x)]^2$ if n is odd.

Solution: Since A is a skew symmetric

$$\therefore A^t = -A.$$

$$\begin{aligned} |A - xI| &= |-A^t - xI| \\ &= |-(A^t + xI)| \\ &= (-1)^n |A^t + xI| \\ &= (-1)^n |(A^t + xI)^t| \because |B^t| = |B| \\ &= (-1)^n |A + xI| \end{aligned}$$

$$\therefore |A - xI| = (-1)^n |A + xI| \dots (1)$$

- (i) If λ is a characteristic root of A , then $|A - \lambda I| = 0$
 $\Rightarrow |A + \lambda I| = 0$, using (1)
 $\Rightarrow -\lambda$ is a characteristic root of A .
- (ii) Let λ be a non-zero characteristic root of A . By (i), $-\lambda$ is also a characteristic root of A . Since $\lambda, -\lambda$ are both characteristic roots of A , therefore

$\lambda^2, (-\lambda)^2$ are characteristic roots of A^2 occurring twice. Hence multiplicity of λ^2 as a characteristic root of $A^2 = 2$ (multiplicity of λ as a characteristic root of A)

Thus every non-zero root of A^2 occurs with even multiplicity.

- (iii) Suppose n is even. Let $n = 2k$, for some k . Since the non-zero roots occur in pairs, by (i), if zero is a root it must have even multiplicity. Let the roots be $\lambda_1, \lambda_2, \dots, \lambda_k, -\lambda_1, -\lambda_2, \dots, -\lambda_k$ then the roots of A^2 are

$$\begin{aligned} \lambda_1^2, \lambda_1^2, \lambda_2^2, \lambda_2^2, \dots, \lambda_k^2, \lambda_k^2 \\ |A^2 - xI| = (x - \lambda_1^2)^2 (x - \lambda_2^2)^2 \dots (x - \lambda_k^2)^2 \\ = [(x - \lambda_1^2)(x - \lambda_2^2) \dots (x - \lambda_k^2)]^2 \end{aligned}$$

Suppose n is odd. Then $n = 2k + 1$. By (i), the non-zero roots occur in pairs and the number of roots is even, $\therefore 0$ must be a root with odd multiplicity $2m + 1$ (say). Hence the roots are

$$\begin{aligned} 0, \lambda_1, \dots, \lambda_l, -\lambda_1, \dots, -\lambda_l \text{ so that } |A - xI| = x(x - \lambda_1^2)^2 \dots (x - \lambda_l^2)^2 \\ = x[(x - \lambda_1^2) \dots (x - \lambda_l^2)]^2 \\ x(f(x))^2 \end{aligned}$$

Problem 11.23. Let A be a 2×2 matrix with complex entries such that $A^2 = 0$.

Prove that either $A = 0$ or A is similar to $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

Solution: Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $A^2 = 0$. Then $|A^2| = 0 \Rightarrow |A| = 0 \Rightarrow$

$ad - bc = 0$. If λ_1, λ_2 are the characteristic roots of A , the characteristic roots of A^2 are λ_1^2, λ_2^2 .

Since $A^2 = 0$

$$\therefore \lambda_1^2 = \lambda_2^2 = 0$$

$$\Rightarrow \lambda_1 = \lambda_2 = 0$$

$\Rightarrow 0$ is the only characteristic root of A . The characteristic equation of A is

$$|A - xI| = 0$$

$$\Rightarrow x^2 - (a + d)x + ad - bc = 0$$

$$\Rightarrow x^2 - (a + d)x = 0 \dots (1)$$

By Cayley Hamilton theorem, A satisfies (1) so that $A^2 - (a + d)A = 0$

$$\Rightarrow (a + d)A = 0 \because A^2 = 0$$

$$\Rightarrow (a + d) = 0 \text{ or } A = 0$$

If $A = 0$ result is proved. Let us now take $a + d = 0$

$\therefore d = -a$, so that

$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$. We now find a characteristic vector X associated with the

characteristic value 0. If $X = (x_1 \ x_2)$ is such a vector then $AX = 0X$

$$\Rightarrow \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$x_1 = a, x_2 = c \text{ satisfies this (as } a^2 + bc = -ad + bc = 0)$$

$\therefore X = \begin{pmatrix} a \\ c \end{pmatrix} \neq 0$ is a characteristic vector. If $c = 0$ then $a^2 + bc = 0 \Rightarrow a = 0$

so that $X = 0$, a contradiction. So $c \neq 0$. If $Y = (1, 0)^t$, then X, Y are linearly

independent vectors. If $P = \begin{pmatrix} 1 & a \\ 0 & c \end{pmatrix}$, then $P^{-1}AP = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ so that A is

similar to $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

11.18 Exercise

1. If A is a non-singular matrix, prove that the characteristic roots of A^{-1} are the reciprocals of the characteristic roots of A .
2. Prove that 0 is a characteristic root of A if and only if A is a singular matrix.
3. If λ is a characteristic roots of A , prove that $\lambda - k$ is a characteristic root of $A - kI$.
4. Give an independent proof of the following:
 - (i) The characteristic roots of a skew Hermitian matrix are zero or pure imaginary.
 - (ii) The characteristic roots of a real symmetric matrix are real.
 - (iii) The characteristic roots of a real skew symmetric matrix are zero or pure imaginary.
5. Let A be a 2×2 symmetric matrix with real entries. Prove that A is either a scalar matrix or similar to a diagonal matrix.

6. Find the characteristic equation of the matrix $A = \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -1 \\ 10 & -5 & -3 \end{pmatrix}$.

Also find A^{-1} , if it exists.

7. Find the eigenvalues and the corresponding eigenvectors of the matrix

$$\begin{pmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{pmatrix}$$

8. If the characteristic equation of a non-singular matrix A of order 3 is $x^3 - px^2 + qx - r = 0$, then prove that the characteristic equation of $\text{adj} A$ is $x^3 - qx^2 + prx - r^2 = 0$

9. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ 4 & -4 & 4 \end{pmatrix}$

10. Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 2 & 0 & 1 \\ 2 & 3 & 1 \\ 1 & 0 & 2 \end{pmatrix}$

11. For any non-singular matrix P , prove that A and $P^{-1}AP$ have the same trace.

12. In Q.9, is the matrix diagonalizable? If yes, find a matrix P such that $P^{-1}AP$ is diagonal.

11.19 Supplementary Exercises

1. State whether the following the statements are true or false.

- (i) $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ is a 2×2 matrix.
- (ii) If $A = \begin{pmatrix} 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix}$, then $\det A = 0$
- (iii) The null matrix of order 3 is a diagonal matrix.
- (iv) Every null matrix is a scalar matrix .
- (v) The product of two triangular matrices is a triangular matrix.
- (vi) If $A = (a_{ij})$ and $B = (b_{ij})$, then $A + B = (a_{ij} + b_{ij})$.
- (vii) If A and B are matrices such that AB and BA are both defined, then $AB = BA$
- (viii) If α is any real number and A is a 3×10 matrix, then $\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix} A = \alpha A$
- (ix) Ae_i is the i^{th} row of A , where e_i is the i^{th} column of the unit matrix.
- (x) There exists a non-zero matrix A such that $AX = 0$, for all column vectors X .
- (xi) If A is a non-zero diagonal matrix, then $\text{trace } A = |A|$.
- (xii) Every matrix can expressed as the sum of a symmetric and skew symmetric matrix.
- (xiv) Every square matrix can be expressed as $P + iQ$, where P and Q are Hermitian matrices.
- (xv) A square matrix A is skew Hermitian matrix if iA is Hermitian.
- (xvi) If a matrix has a right inverse B and a left inverse C , then $B = C$
- (xvii) If A is any square matrix, then A^n is defined for every integer n .
- (xviii) If A is an orthogonal matrix, then the vector X and AX are orthogonal to each other.
- (xix) If P is a unitary matrix, then APA^θ is also a unitary matrix.
- (xx) The column of a normal matrix are of unit length.
- (xxi) A scalar λ is called an eigenvalue of a square matrix A if there exists a vector X such that $AX = \lambda X$.
- (xxii) For any square matrix A , two distinct eigenvalues can have the same eigenvector.
- (xxiii) For any square matrix A , X is an eigenvector of A if AX and X are collinear.
- (xxiv) An n -rowed square matrix has exactly n eigenvalue.
- (xxv) An n -rowed square matrix always has n linearly independent eigenvectors.
- (xxvi) A square matrix may not have any eigenvector.
- (xxvii) Every matrix satisfies its characteristic equation.
- (xxviii) The number of eigenvectors of a square matrix over complex numbers is infinite.
- (xxix) For any square matrix A , the matrices A and A^θ have same characteristic roots.
- (xxx) Every square matrix is similar to a diagonal matrix.
- (xxx) If the constant term in the characteristic equation of a matrix A is zero, then A is a singular matrix.

2. Give example of a 2×2 matrix over \mathbb{R} which has no eigen vector.
3. If A and B are matrices of suitable sizes, prove that $\text{tr}(AB - BA) = 0$.
4. If A is a skew symmetric matrix, prove that A^n is symmetric or skew symmetric according n is even or odd.
5. If A is a square matrix and $A = P + iQ$ where P and Q are Hermitian matrices, prove that A is normal if and only if P and Q commute.
6. Prove that the characteristic roots of a real skew symmetric matrix are either zero or pure imaginary.
7. Prove that the characteristic roots of a real symmetric matrix are real.
8. Prove that the characteristic vectors corresponding to the two characteristic roots of a skew Hermitian matrix are orthogonal.
9. If λ is a characteristic root of a unitary matrix, then prove that $\frac{1}{\lambda}$ is also a characteristic root.
10. If A is Hermitian (skew Hermitian), show that for any matrix B , $B^\theta AB$, is a Hermitian (skew hermitian).

11. A square matrix A is said to be an idempotent i.e. $A^2 = A$. Show that the matrices

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \\ -2 & 4 & -2 \end{pmatrix}$$
 are idempotents.

12. Find all 2-rowed idempotent matrices.

13. A square matrix A is said to be nilpotent if $A^n = 0$ for some positive integer n . Show that the matrices

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ are nilpotent.}$$

14. Prove that a nilpotent matrix is singular.
15. Prove that a non-zero idempotent matrix cannot be nilpotent.
16. Prove that the only matrix which is both idempotent and nilpotent is the zero matrix.

17. If $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ -1 & 1 & 0 \end{pmatrix}$, show that

$$(i) P^t A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$P^{-1} = P^t$$

18. Find the flaw in the following argument

Let A and B be two n -rowed square matrices. We know that
 $(AB)^{-1} = B^{-1}A^{-1}$
 $\Rightarrow \frac{\text{adj}(AB)}{|AB|} = \frac{\text{adj}B}{|B|} \frac{\text{adj}A}{|A|}, \therefore A^{-1} = \frac{\text{adj}A}{|A|}$
 $\Rightarrow \text{adj}(AB) = (\text{adj}B)(\text{adj}A) \therefore |AB| = |A||B|$

19. Prove that every square matrix can be expressed uniquely as the sum of a Hermitian and skew hermitian matrix.
20. Find the inverse of the matrix using partitioning of matrices.

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

21. Find the inverse of the matrix

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

using partitioning of matrices. Partition its in two different ways. Do you get different inverse?

22. Using partitioning of matrices, prove that the inverse of the matrix

$$X = \begin{pmatrix} A & B & C \\ 0 & D & 0 \\ 0 & E & F \end{pmatrix} \text{ is } \begin{pmatrix} A^{-1} - A^{-1}(B - CF^{-1}E)D^{-1} & -A^{-1}CF^{-1} \\ 0 & D^{-1} & 0 \\ 0 & -F^{-1}ED^{-1} & F^{-1} \end{pmatrix}$$

where A, D, F are invertible matrices and all the matrices A, B, C, D, E, F are of suitable sizes for the products to be defined.

23. If A is a singular matrix, prove that the columns of $\text{adj}A$ are the solutions of $AX = 0$.
24. If A is a singular matrix, prove that the columns of A are the solutions of $(\text{adj}A)X = 0$.
25. If A and B commute, prove that PAP^t and PBP^t also commute for any orthogonal matrix P .
26. If A and B commute and U is a unitary matrix, then prove that UAU^θ and UBU^θ also commute.
27. If A and B are skew symmetric matrices, prove that AB is skew symmetric if and only if $AB = -BA$.
28. Prove that the product of two Hermitian matrices is Hermitian if and only if they commute.

11.20 Answers to Exercises

Exercise - 11.5

1. (i) $\begin{pmatrix} 1+3i & 2-5i \\ -2 & -4+2i \end{pmatrix}$
 (ii) $\begin{pmatrix} 3+5i & 7-11i \\ -8-2i & -5+3i \end{pmatrix}$
 (iii) $\begin{pmatrix} -2+2i & 3+9i \\ -4i & 0 \end{pmatrix}$
 (iv) $\begin{pmatrix} 4+5i \\ -1-3i \end{pmatrix}$
 (v) $\begin{pmatrix} 1-i & 1-3i \\ -2 & -1-i \end{pmatrix}$
 (vi) $\begin{pmatrix} 1-i & -2 \\ 1+3i & -1-i \end{pmatrix}$
 (vii) $\begin{pmatrix} -2-4i \\ 5 \end{pmatrix}$
 (viii) $\begin{pmatrix} -6-5i \\ -2-2i \end{pmatrix}$
 (ix) $\begin{pmatrix} 3-i & -2-8i \\ -2-2i & 4i \end{pmatrix}$
 (x) $\begin{pmatrix} -2-2i & 4i \\ 3-9i & 0 \end{pmatrix}$
2. $\pm \begin{pmatrix} i & -i \\ -i & i \end{pmatrix}$
3. (i) $\begin{pmatrix} 19 & 0 \\ 0 & 19 \end{pmatrix}$
 (ii) $\begin{pmatrix} 4 & 0 \\ 0 & 7 \end{pmatrix}$
 (iii) $\begin{pmatrix} 3 & -3i \\ -3i & 3 \end{pmatrix}$
 (iv) $\begin{pmatrix} 4 & -2 \\ 0 & 4 \end{pmatrix}$
5. $\pm iI_2$
6. (i) symmetric
 (ii) skew Hermitian
 (iii) symmetric
 (iv) all
 (v) symmetric, Hermitian
 (vi) Hermitian
 (vii) skew symmetric
 (viii) skew symmetric, skew Hermitian.
9. (i) $P = \frac{1}{2} \begin{pmatrix} 4+2i & 8+i & -4+3i \\ 8+i & 4 & 5+7i \\ -4+3i & 5+7i & 12i \end{pmatrix},$
 $Q = \frac{1}{2} \begin{pmatrix} 0 & -2-i & -4-3i \\ 2+i & 0 & 7-i \\ 4+3i & -7+i & 0 \end{pmatrix}$

$$(ii) A = \frac{1}{2} \begin{pmatrix} 4 & 8-i & -4-3i \\ 8+i & 4 & 5-i \\ -4+3i & 5+i & 0 \end{pmatrix},$$

$$Q = \frac{1}{2} \begin{pmatrix} 2i & -2+i & -4+3i \\ 2+i & 0 & 7+7i \\ 4+3i & -7+7i & 12i \end{pmatrix}$$

$$(iii) P = \frac{1}{2} \begin{pmatrix} 4 & 8-i & -4-3i \\ 8+i & 4 & 5-i \\ -4+3i & 5+i & 0 \end{pmatrix},$$

$$Q = \frac{1}{2} \begin{pmatrix} 2 & 1+2i & 3+4i \\ 1-2i & 0 & 7-7i \\ 3-4i & 7+7i & 12 \end{pmatrix}$$

$$10. (i) P = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 4 \\ 3 & 4 & 2 \end{pmatrix},$$

$$Q = \begin{pmatrix} 0 & -3 & 4 \\ 3 & 0 & -5 \\ -4 & 5 & 0 \end{pmatrix}$$

(ii) Not possible, as A is not Hermitian.

$$11. (i) mn$$

$$(ii) n^2$$

$$(iii) n^2 - 1$$

$$(iv) n$$

$$(v) 1$$

$$(vi) \frac{1}{2}n(n+1)$$

$$(vii) \frac{1}{2}n(n-1)$$

$$(viii) \frac{1}{2}n(n+1), \text{ with diagonal entries real}$$

$$(ix) \frac{1}{2}n(n+1), \text{ with diagonal entries 0 or pure imaginary.}$$

$$13. \text{ Hint } |A^t| = |-A| = (-1)^n |A|/$$

$$19. (i) 4 \times 7, 5 \times 5, 5 \times 7$$

$$(ii) 7 \times 6, 2 \times 6, 2 \times 6$$

$$(iii) 3 \times 8, 6 \times 4, 6 \times 8$$

$$(iv) 6 \times 3, 3 \times 9, 3 \times 3$$

Exercise - 11.11

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

$$2. \begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}, \text{ No, } \begin{pmatrix} k_1 & 1-2k_1 & k_1 \\ k_2-1 & 2-2k_2 & k_2 \end{pmatrix}, k_1, k_2 \in \mathbb{R}$$

14. (i) $\frac{1}{15} \begin{pmatrix} -11 & 9 & -1 \\ 7 & -9 & 2 \\ -2 & 3 & -1 \end{pmatrix}$
- (ii) $\begin{pmatrix} 1 & 2 & -1 \\ -4 & -7 & 4 \\ -4 & -9 & 5 \end{pmatrix}$
- (iii) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 3 & -1 \end{pmatrix}$
- (iv) $\begin{pmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{pmatrix}$
- (v) $\begin{pmatrix} 1 & -x & x^2 \\ 0 & 1 & -x \\ 0 & 0 & 1 \end{pmatrix}$
- (vi) $\begin{pmatrix} 1 & 0 & 0 \\ -y & 1 & 0 \\ y^2 & -y & 1 \end{pmatrix}$
18. (i) $\frac{1}{12} \begin{pmatrix} 24 & -18 & 3 \\ -12 & 12 & 4 \\ 0 & 0 & 2 \end{pmatrix}$
- (ii) $\begin{pmatrix} 1 & -\alpha & 0 \\ 0 & 1 & 0 \\ 0 & -\beta & 1 \end{pmatrix}$
22. (i) $\frac{1}{3} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 2 & 1 \end{pmatrix}$ (iv) $\frac{1}{4} \begin{pmatrix} 2 & -2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ -1 & 1 & 2 & 0 \\ -1 & 1 & -2 & 4 \end{pmatrix}$
- (ii) $\frac{1}{3} \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & -1 & -2 \\ -1 & 2 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{pmatrix}$ (v) $\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 2 \\ 1 & 1 & 0 & 0 \end{pmatrix}$
- (iii) $\begin{pmatrix} 2 & -1 & -2 & 1 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}$ (vi) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$
23. $\begin{pmatrix} \frac{1}{2} & \frac{-1}{2} & \frac{-1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{-1}{2} & \frac{1}{2} \end{pmatrix}$

Exercise - 11.14

3. 5

4. $x = \pm \frac{1}{2}, y = \mp \frac{1}{2}$

5. $\frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 1 & -2 \end{pmatrix}^t$
 6. $\frac{1}{\sqrt{7}} \begin{pmatrix} 1 & -1 & -2 & 1 \end{pmatrix}^t$
 12. $q = 0 \Rightarrow A = I$

Exercise - 11.18

5. *Hint:* $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ Characteristic roots are real if $(a - d)^2 + 4b^2 \geq 0$,
 roots are equal if $a = d$ and $b = 0$. $\therefore A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$
 if roots are α, β then $A \sim \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$

6. $x^3 - 2x^2 + x - 2 = 0; 2, \pm i$
 $\begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ 1 & 1 & -1 \\ -5 & 0 & 3 \end{pmatrix}$

7. $-1, -1, 3$. $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}$ are the corresponding eigenvectors.

8. *Hint:* If λ is a characteristic root of A then $\frac{|A|}{\lambda}$ is a characteristic root of $\text{adj}A$. Also $|A| = \lambda_1 \lambda_2 \lambda_3 = -r$.

9. $\lambda = 3, 3, 12$ when $\lambda = 3, X = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} k + \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} k_1; k, k_1 \in \mathbb{R}, X \neq 0$, when

$$\lambda = 12, X = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} k, k \in \mathbb{R}^*$$

10. $\lambda = 1, 3, 3$

$$\text{For } \lambda = 1, X = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} k; k \in \mathbb{R}^*, \text{ and for } \lambda = 3, X = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} k; k \in \mathbb{R}^*.$$

Supplementary Exercises

1. (i) F, it is determinant of a 2×2 matrix.
- (ii) F, determinant is defined only for square matrices.
- (iii) T
- (iv) F, if it is a square matrix.
- (v) F, $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$
- (vi) F, A, B should be of same order.
- (vii) F
- (viii) T
- (ix) F, i^{th} column.
- (x) F, using (ix)

(xi) F, $A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, $|A| = 0$, $trA = 2$

(xii) F, matrix should be square.

(xiii) F, zero or pure imaginary.

(xiv) T

(xv) T

(xvi) F

(xvii) F

(xviii) F

(xix) T

(xx) F, unitary matrix.

(xxi) F

(xxii) F

(xxiii) T

(xxiv) F

(xxv) F

(xxvi) T

(xxvii) F

(xxviii) T

(xxix) F

(xxx) F

(xxxi) T

12. $\begin{pmatrix} a & b \\ c & 1-a \end{pmatrix}$

18. *Flaw:* The I_{st} step is incorrect. A and B should be non-singular matrices.

20. $\frac{1}{2} \begin{pmatrix} 0 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

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

24. *Hint:* $|A| = 0 \Rightarrow AadjA = 0$

If $AdjA = [c_1 \ c_2 \ \dots \ c_n]$ then $Ac_i = 0, i = 1, 2, \dots, n$.

Chapter 12

Matrices and Linear Transformations

In this chapter we are going to study a special kind of function that arises very naturally in the study of linear algebra and has many applications in other fields such as physics and engineering.

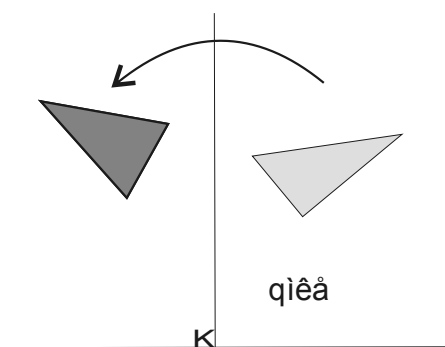
These functions are called linear transformations. We shall show that there is a close relationship between matrices and linear transformations. In fact every $m \times n$ matrix gives rise to a linear transformation and vice versa.

12.1 Introduction to Linear Transformations

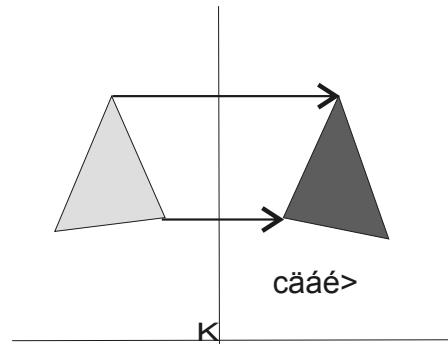
This section is devoted mostly to the basic definitions and facts associated with this special kind of function. In this section we are going to look at functions from \mathbb{R}^n into \mathbb{R}^m .

In other words, we are going to look at functions that take elements/points/vectors from \mathbb{R}^n and associate them with elements/points/vectors from \mathbb{R}^m . These kinds of functions are called transformations.

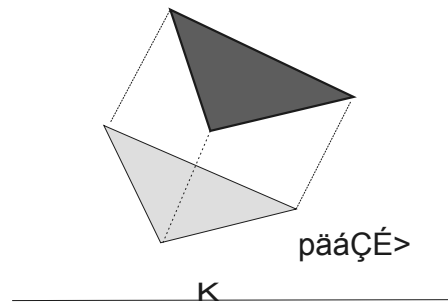
The three examples of transformations of \mathbb{R}^2 are:



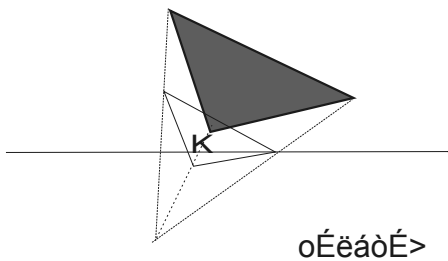
Rotation



Reflection



Translation



Resizing

The other important transformation is resizing (also called dilation, contraction, compression, enlargement or even expansion). The shape becomes bigger or smaller.

Definition 12.1. (Transformation)

A transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a rule that assigns to each vector X in \mathbb{R}^n a unique vector, denoted by $T(X)$, in \mathbb{R}^m .

\mathbb{R}^n is called the domain and \mathbb{R}^m the co-domain of T .

$T(X)$ is called image of X under T .

The subset $\{T(X) : X \in \mathbb{R}^n\}$ of \mathbb{R}^m is called the range of T .

If for a $Y \in \mathbb{R}^m$, there is some $X \in \mathbb{R}^n$ such that $T(X) = Y$ then X is called the pre-range of Y under T .

Transformations are described either by giving the rule explicitly or in terms of matrix multiplication, for example, if T is a transformation from \mathbb{R}^2 to \mathbb{R}^3 then we can write

$T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, defined by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \\ y \end{pmatrix} \quad (\text{Rule method})$$

$$\text{or } T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{Matrix multiplication method})$$

Domain of T is \mathbb{R}^2 , co-domain is \mathbb{R}^3 . The image of any vector $\begin{pmatrix} x \\ y \end{pmatrix}$ is vector

$$\begin{pmatrix} x + y \\ x - y \\ y \end{pmatrix}. \text{ A pre-image of } v = \begin{pmatrix} -1 \\ 3 \\ -2 \end{pmatrix} \text{ is } u = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \text{ as } T(u) = v$$

Examples of Transformations

Some transformation of \mathbb{R}^2 to \mathbb{R}^2 can be defined as follows:

$$\text{Reflection through } y \text{ axis: } T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$$

$$\text{Reflection through } x \text{ axis: } T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$

$$\text{Scaling by a factor } k: T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix}$$

$$\text{Projection on } x \text{ axis: } T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$$

It is easily seen that reflections and scaling are injective transformations which are also onto \mathbb{R}^2 , whereas, projection is neither injective nor onto.

$T : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by $T \begin{pmatrix} x \\ y \end{pmatrix} = x$ is a projection of \mathbb{R}^2 onto \mathbb{R} which is not injective.

$T : \mathbb{R} \rightarrow \mathbb{R}^2$ defined by $T(x) = \begin{pmatrix} x \\ 1 \end{pmatrix}$ is injective but not onto.

Example 12.1. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be defined by

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} |x_1| \\ x_2 \\ x_1 + x_2 \end{pmatrix}$$

(i) If $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Is $T(u+v) = T(u) + T(v)$?

(ii) What are the pre-images of $\begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -4 \\ -5 \end{pmatrix}$?

(iii) Is T onto?

(iv) Is T injective?

Solution:

$$\begin{aligned} \text{(i) } u + v &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ T(u) &= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad T(v) = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \\ T(u + v) &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad T(u) + T(v) = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \\ \therefore T(u + v) &\neq T(u) + T(v). \end{aligned}$$

$$\begin{aligned} \text{(ii) Let } X &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \text{ such that} \\ T(X) &= \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \quad \therefore \begin{pmatrix} |x_1| \\ x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \end{aligned}$$

Since $|x_1| \geq 0$, \therefore we can't have $|x_1| = -2$

Hence no such X exists. Thus $\begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$ doesn't have a pre-image.

Let $X \in \mathbb{R}^3$ such that

$$T(X) = \begin{pmatrix} 1 \\ -4 \\ -3 \end{pmatrix}.$$

$$\text{Hence } \begin{pmatrix} |x_1| \\ x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ -3 \end{pmatrix}$$

This gives $|x_1| = 1$

$\Rightarrow x_1 = \pm 1$. Last two equations give $x_1 = -1$. Hence

$$x_1 = -1, \quad x_2 = -4 \quad \therefore T \begin{pmatrix} -1 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ -5 \end{pmatrix}$$

(iii) T is not a onto because $\begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$ does not have a pre-image.

(iv) Let $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ such that $T(X) = T(Y)$,

$$\text{then } \begin{pmatrix} |x_1| \\ x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} |y_1| \\ y_2 \\ y_1 + y_2 \end{pmatrix}$$

thus $|x_1| = |y_1|, x_2 = y_2, x_1 + x_2 = y_1 + y_2$.

Last two equations give $x_1 = y_1$ and $x_2 = y_2$.
Thus $X = Y$, so that T is injective.

12.2 Exercise

- Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a mapping
 - What is the domain of T ?
 - What is the co-domain of T ?
 - Range T is a subset of \mathbb{R}^3 or \mathbb{R}^2 ?
- Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a transformation defined by $T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$
Find the image under T of u , v and $u + v$, for $u = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$
Is $T(u + v) = T(u) + T(v)$?
- T is a transformation from \mathbb{R}^2 to \mathbb{R}^2 defined by $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 + 2 \end{pmatrix}$
Find the images under T of u , v and $u + v$, where $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$
. Is $T(u) + T(v) = T(u + v)$?
Also find the pre-image of 0.
- T is a transformation from \mathbb{R}^2 to \mathbb{R}^2 defined by $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix}$
 - Find $T(u)$ for $u = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.
 - Is $T(u) = T(-u)$, for u in (i).
 - Is there an element in the co-domain which is not in the range of T .
 - Is T onto?
 - Is T injective?

12.3 Matrix Transformations

In this section we discuss transformations which arise by matrix multiplication. Let A be any $m \times n$ matrix and X be a vector in \mathbb{R}^n , then AX is a $m \times 1$ matrix which belongs to \mathbb{R}^m .

Thus for every $X \in \mathbb{R}^n$ there is defined a vector AX in \mathbb{R}^m which gives rise to a transformation

$X \rightarrow AX$ from \mathbb{R}^n into \mathbb{R}^m .

This leads to the following definition.

Definition 12.2. (Matrix transformation)

If A is a $m \times n$ matrix, then for the transformation $X \rightarrow AX$ to be defined, X must be a vector in \mathbb{R}^n and then AX is a vector in \mathbb{R}^m .

Therefore $m \times n$ matrix always determines a transformation from \mathbb{R}^n into \mathbb{R}^m .

Definition 12.3. (Zero transformation)

If A is the $m \times n$ null matrix, then $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ determined by A is called the zero transformation. It maps every element of \mathbb{R}^n to the zero element of \mathbb{R}^m .

Definition 12.4. (Identity transformation)

If A is the $n \times n$ identity matrix, then the transformation from $\mathbb{R}^n \rightarrow \mathbb{R}^n$ determined by A is called identity transformation.

It maps every element of \mathbb{R}^n to itself.

Example 12.2. Let $A = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ 4 & -1 \end{pmatrix}$. The matrix transformation determined

by A is $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$

$$\text{where } T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - 2y \\ 3x \\ 4x - y \end{pmatrix}$$

Similarly a 2×3 matrix will define a transformation from \mathbb{R}^3 into \mathbb{R}^2 .

Properties of matrix multiplication give that

- (i) $A(u + v) = Au + Av$ and $A(ku) = kAu$, where k is a real number.
- (ii) $Ae_1 = c_1, Ae_2 = c_2, \dots, Ae_n = c_n$ where e_1, e_2, \dots, e_n are the columns of the identity matrix I_n and c_1, c_2, \dots, c_n are the columns of A .

12.4 Surjective and Injective Matrix Transformations

Now we study under what conditions a given matrix transformation is surjective and injective. Consider the matrix transformation T determined by $m \times n$ matrix A .

The following questions arise:

- (i) Given b in R^m , does b lie in the range of T ?
- (ii) Is T onto?
- (iii) Is T injective?
- (iv) Given b in range T , does there exist a unique x such that $T(x) = b$.

We answer these questions one by one.

- (i) Given b in R^m

b lies in the Range T

\Leftrightarrow There exists some x in R^n such that $T(x) = b$

$\Leftrightarrow Ax = b$ has a solution.

\Leftrightarrow Augmented column of $[A : b]$ does not have a pivot position.

- (ii) T is onto

\Leftrightarrow Range $T = R^m$

\Leftrightarrow For all b in R^m there exists some x in R^n such that $T(x) = b$

$\Leftrightarrow Ax = b$ has a solution for all b in R^m

\Leftrightarrow every row of A has a pivot position.

- (iii) T is injective
- \Leftrightarrow Given x_1, x_2 in R^n such that $T(x_1) = T(x_2)$, then $x_1 = x_2$.
 - $\Leftrightarrow Ax_1 = Ax_2$ implies that $x_1 = x_2$.
 - $\Leftrightarrow A(x_1 - x_2) = 0$ implies that $x_1 - x_2 = 0$
 - $\Leftrightarrow Ax = 0$ implies that $x = 0$
 - $\Leftrightarrow 0$ is the only solution of $Ax = 0$
 - \Leftrightarrow Every column of A has a pivot position.
- (iv) There exists a unique x in R^n such that $T(x) = b$
- $\Leftrightarrow Ax = b$ has a unique solution.
 - \Leftrightarrow In the augmented matrix $[A : b]$ every column of A has pivot position and augmented column does not have pivot position.

The above discussion can be summarized as follows:

Property of T	Meaning in terms of T	Meaning in terms of A	Test on A
b lies in range T	There exists x in R^n such that $T(x) = b$	There exists x in R^n such that $Ax = b$	Augmented column of $[A : b]$ doesn't have a pivot.
T is onto (surjective)	For every b in R^m there exists x in R^n such that $Tx = b$	$Ax = b$ has a solution for every b in R^m	Every row of A has a pivot position
T is injective	Given x_1, x_2 in R^n such that $T(x_1) = T(x_2)$ then $x_1 = x_2$	0 is the only solution of $Ax = 0$	Every column of A has pivot position
b has unique pre-image under T	There exists a unique x in R^n such that $T(x) = b$	$Ax = b$ has a unique solution	In $[A : b]$, every column of A has a pivot position but augmented column doesn't have a pivot position.

Example 12.3. Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \end{pmatrix}$, $u = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$,

$$c = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Let T be the matrix transformation determined by A

- (i) Find $T(u)$.
- (ii) Find a vector x such that $T(x) = b$.
- (iii) Can you find more than one x in part (ii)?
- (iv) Does c lie in the range of T ?

Solution: The transformation determined by a 2×3 matrix A is from R^3 to R^2 .

For any $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ in R^3

$$T(x) = T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = Ax = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 - x_3 \\ 3x_1 + x_3 \end{pmatrix}$$

$$(i) T(u) = T \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 + 0 - 1 \\ -3 + 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

(ii) Finding x in \mathbb{R}^3 such that $T(x) = b$

Solving $Ax = b$,

$$\text{Augmented matrix } [A : b] = \left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 3 & 0 & 1 & -1 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & -\frac{2}{3} & -\frac{1}{3} \end{array} \right)$$

Solution is $x_1 = -\frac{1}{3}k - \frac{1}{3}$

$$x_2 = \frac{2}{3}k + \frac{2}{3}$$

$x_3 = k$, where k is any real number.

Taking a particular value of $k = 0$ (say) $x_1 = -\frac{1}{3}$, $x_2 = \frac{2}{3}$, $x_3 = 0$

Therefore $x = \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \end{pmatrix}$ is a required vector.

(iii) The general solution in (ii) is

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\frac{k}{3} - \frac{1}{3} \\ \frac{2k}{3} + \frac{2}{3} \\ k \end{pmatrix} = \frac{k}{3} \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \end{pmatrix}. \text{ For every}$$

value of k we get a solution.

Thus there are infinitely many solutions in (ii).

(iv) c will lie in Range T

\Leftrightarrow there exists x in \mathbb{R}^3 such that $T(x) = c$,

$\Leftrightarrow Ax = c$ has a solution.

\Leftrightarrow Augmented matrix $[A : c]$ doesn't have a pivot in the augmented column.

$$[A : c] \sim \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 3 & 0 & 1 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & -\frac{2}{3} & \frac{4}{3} \end{array} \right)$$

Pivots lie in 1st and 2nd columns and not in the augmented column. Therefore $Ax = c$ has a solution and so c lies in the range of T .

Problem 12.1. Define $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by $TX = AX$, where

$$A = \begin{pmatrix} 1 & 0 & -2 \\ -2 & 1 & 6 \\ 3 & -2 & -5 \end{pmatrix}$$

If $b = \begin{pmatrix} -1 \\ 7 \\ -3 \end{pmatrix}$, find $X \in \mathbb{R}^3$ whose image under T is b . Is X unique?

Solution: *Step 1* Suppose $X \in \mathbb{R}^3$ be such that $TX = b$. Hence $AX = b$. Thus finding X amounts to solving the linear system $AX = b$.

To do this we shall reduce the augmented matrix to echelon form.

$$\text{Step 2 } [A : b] = \left(\begin{array}{ccc|c} 1 & 0 & -2 & -1 \\ -2 & 1 & 6 & 7 \\ 3 & -2 & -5 & -3 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 5 \\ 0 & -2 & 1 & 0 \end{array} \right)$$

Applying $R_2 \rightarrow R_2 + 2R_1$

$$R_3 \rightarrow R_3 - 3R_1$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 5 & 10 \end{array} \right) \text{Applying } R_3 \rightarrow R_3 + 2R_2$$

To obtain the solution we reduced it to reduced echelon form.

$$[A : b] \sim \left(\begin{array}{ccc|c} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 1 & 2 \end{array} \right) \text{Applying } R_3 \rightarrow \frac{1}{5}R_3$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right) R_2 \rightarrow R_2 - 2R_3, R_1 \rightarrow R_1 + 2R_3$$

Hence solution is

$$x_1 = 3$$

$$x_2 = 1$$

$$x_3 = 2$$

Step 3 Thus $X = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$ is such that

$$T(X) = AX = b$$

Since there is only one value of X , therefore X is unique.

Problem 12.2. Define T by $T(X) = A(X)$, where $A = \begin{pmatrix} 1 & -5 & -7 \\ -3 & 7 & 5 \end{pmatrix}$

(i) Is T onto?

(ii) Find X such that $T(X) = b$, where $b = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$.

(iii) Is X obtained in (ii) unique? If not, find all X such that $T(X) = b$.

Solution: Clearly $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$

(i) T will be onto if given any $b \in \mathbb{R}^2$, $\exists X \in \mathbb{R}^3$ such that $T(X) = b$
i.e. if $AX = b$ has a solution for every $b \in \mathbb{R}^2$

$$\text{Let } b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\text{Then } [A : b] = \left(\begin{array}{ccc|c} 1 & -5 & -7 & b_1 \\ -3 & 7 & 5 & b_2 \end{array} \right)$$

the row reduced echelon form is

$$[A : b] \sim \left(\begin{array}{ccc|c} 1 & 0 & 3 & -(7b_1 + 5b_2)/8 \\ 0 & 1 & 2 & -(3b_1 + b_2)/8 \end{array} \right)$$

Since every row of A has a pivot,
therefore there is a solution for all $b \in \mathbb{R}^2$.

Hence T is onto.

(ii) For the given b , we get from (i)

$$[A : b] \sim \left(\begin{array}{ccc|c} 1 & 0 & 3 & 3 \\ 0 & 1 & 2 & 1 \end{array} \right)$$

The equations are

$$x_1 + 3x_3 = 3$$

$$x_2 + 2x_3 = 1$$

The solution is

$$x_1 = 3 - 3k$$

$$x_2 = 1 - 2k$$

$x_3 = k$,
where k is any real number.

Taking $k = 0$,

$X = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ is such that

$$T(X) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

(iii) From (ii) we get

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 - 3k \\ 1 - 2k \\ k \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} - k \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = u - kv \text{ (say)}$$

Thus the X obtained in (ii) is not unique.

$$\text{In fact } T(u - kv) = b = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

for all real numbers k .

Problem 12.3. Let T be a linear operator on \mathbb{R}^3 defined by $TX = AX$, where

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix}$$

Does b lie in range of T , where $b = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$?

Solution: b will lie in the range of T if there exists some $X \in \mathbb{R}^3$ such that $T(X) = b$

i.e. if $AX = b$ has a solution.

The row reduced echelon form of $[A:b]$ is $\begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Since there is a pivot in the augmented column, therefore the system $AX = b$ is inconsistent, that is, it does not have a solution.

Hence b does not lie in the range of T .

Problem 12.4. Find all $X \in \mathbb{R}^4$ which are mapped to zero by the transformation $X \rightarrow AX$, for

$$A = \begin{pmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{pmatrix}$$

Solution: If X is mapped to 0, then we have to find the solution of $AX = 0$.

The row reduced echelon form of $[A : 0]$ is $\begin{pmatrix} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Thus the equivalent system of equations is

$$x_1 + 2x_4 = 0$$

$$x_2 + x_4 = 0$$

$$x_3 + x_4 = 0$$

$\therefore x_1 = -2x_4, x_2 = -x_4, x_3 = -x_4, x_4$ is free.

$$\text{Hence } X = \begin{pmatrix} -2x_4 \\ -x_4 \\ -x_4 \\ x_4 \end{pmatrix} = -x_4 \begin{pmatrix} 2 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

Thus, all multiples of $\begin{pmatrix} 2 \\ 1 \\ 1 \\ -1 \end{pmatrix}$ are mapped to 0.

Problem 12.5. Let T be the transformation defined on \mathbb{R}^3 by $T(X) = AX$

where $A = \begin{pmatrix} 1 & -3 & -8 \\ 3 & 1 & -4 \\ 2 & 5 & 6 \end{pmatrix}$. If $b = \begin{pmatrix} -10 \\ 0 \\ 13 \end{pmatrix}$, find

- (i) some vector X such that $T(X) = b$.
 (ii) all vectors X such that $T(X) = b$.

Solution:

- (i) Finding a vector X such that $T(X) = b$ is equivalent to solving $AX = b$.

Step 1 We shall obtain the reduced echelon form of the augmented matrix $[A : b]$.

$$\text{Then } [A : b] \sim \begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Step 2 The corresponding equations are

$$\begin{aligned} x_1 - 2x_3 &= -1 \\ x_2 + 2x_3 &= 3 \\ 0 &= 0 \end{aligned}$$

Step 3 Taking $x_3 = 0$, we get

$$x_1 = -1, x_2 = 3$$

Thus $X_0 = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}$ is a vector such that

$$T(X_0) = b$$

- (ii) To find all vectors X such that $T(X) = b$

We proceed up to Step 2 of (i).

Taking $x_3 = k$ any real number, we get from Step 2,

$$x_1 = -1 + 2k$$

$$x_2 = 3 - 2k$$

$$x_3 = k$$

$$\text{Thus } X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} + k \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

$$= X_0 + kX_1 \text{ where } X_1 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

and X_0 is the same as in Step 3.

Thus $X = X_0 + kX_1, k \in \mathbb{R}$

such that $T(X) = b$.

Note: It is important to note that X_1 is a solution of $AX = 0$ and X_0 is one solution of $AX = b$

$X = X_0 + kX_1$ is the general solution of $AX = b$.

Problem 12.6. For the transformation in the above problem find all vectors X such that $T(X) = 0$

Solution: $T(X) = 0$

$\Leftrightarrow AX = 0$

As in the previous problem,

$$[A : 0] \sim \left(\begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

The corresponding equations are

$$x_1 - 2x_3 = 0$$

$$x_2 + 2x_3 = 0$$

$$0x_3 = 0$$

$$x_1 = 2k, \quad x_2 = -2k, \quad x_3 = k, \quad \text{where } k \in \mathbb{R}$$

$$\therefore X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} k = X_1 k \text{ (say)}$$

General solution of $T(X) = 0$ is $X = kX_1, \quad k \in \mathbb{R}$

Problem 12.7. Let A be an $m \times n$ matrix.

Define $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by $T(X) = AX$.

Then

(i) The columns of A are images of the columns of I_n .

(ii) If the i th row of A is zero, then the i th coordinate of $T(X)$ is zero.

Solution: Let $A = [c_1 \ c_2 \dots \ c_n]$

(i) Let $e_i, \ 1 \leq i \leq n$, be the columns of the $n \times n$ unit matrix.

Then for $1 \leq i \leq n, \ T(e_i) = Ae_i$

$$= [c_1 \ c_2 \dots \ c_n]e_i$$

$$= [c_1 \ c_2 \dots \ c_n] \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1(\text{ith position}) \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= [c_1 0 + c_2 0 + \dots + c_i 1 + \dots + c_n 0]$$

$$= c_i$$

$$\text{Hence } T(e_i) = c_i, \quad 1 \leq i \leq n$$

(ii) Let $A = (a_{ij})_{m \times n}$

i th coordinate of $TX = i$ th coordinate of AX

$$AX = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix} X = \begin{pmatrix} R_1 X \\ R_2 X \\ \vdots \\ R_m X \end{pmatrix}$$

$$= a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$$

Thus if i th row of A is zero,
 then i th coordinate of $TX = i$ th coordinate of AX
 $= a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0$.
 Thus i th row of A is zero,
 then i th coordinate of $TX = 0$.

12.5 Exercise

1. Let $T : R^5 \rightarrow R^3$ be defined by $T(x) = Ax$. The matrix A , u and v are

$$\text{defined as } A = \begin{pmatrix} 1 & -1 & 2 & 3 & 0 \\ 4 & 0 & 2 & 1 & -2 \\ 2 & 3 & 1 & 3 & 2 \end{pmatrix}, u = \begin{pmatrix} 2 \\ 4 \\ 3 \\ 1 \\ 3 \end{pmatrix}, v = \begin{pmatrix} 3 \\ 1 \\ 4 \\ 0 \\ 1 \end{pmatrix}$$

- (a) Compute

- (i) $T(u)$, $T(v)$ and $T(u) + T(v)$
 (ii) $u + v$ and $T(u + v)$

What do you observe about the relationship between $T(u) + T(v)$ and $T(u + v)$? Give and prove your statement (for any $m \times n$ matrix A and any vectors u, v in R^n) using matrix algebra.

- (b) Compute $3T(u)$ and $T(3u)$. What do you deduce? Give and prove your statement (for any $m \times n$ matrix A any vector u in R^n , any scalar c) using matrix algebra.

2. Suppose that $T : R^2 \rightarrow R^4$ is defined via multiplication by a matrix A . What is the size of the matrix A ?

3. Let A be a 3×4 matrix. What must be p and q so that $T : R^p \rightarrow R^q$ is defined by $T(X) = AX$?

4. Let A be a 4×3 matrix and T be defined by $T : R^3 \rightarrow R^4$ by $T(X) = AX$. Without actually verifying, can you tell whether T is a linear transformation or not.

5. Let T be a matrix transformation defined from R^2 to R^4 . If $T((1, 0)) = (2, 1, -2, 0)$ and $T((0, 1)) = (-1, -5, 2, 1)$, Use matrix algebra to find
 (i) $T((2, 0))$
 (ii) $T((1, 1))$
 (iii) $T((3, -1))$

6. Let T be the matrix transformation determined by 3×2 matrix A such

$$\text{that } T(e_1) = \begin{pmatrix} -3 \\ 6 \\ 0 \end{pmatrix}, T(e_2) = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

- (i) Find the matrix A .
 (ii) Find $T(u)$, where $u = \begin{pmatrix} -6 \\ 5 \end{pmatrix}$

7. Let $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ and $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the transformation determined by A .
- (i) What is the image of the line $x = a + tb$, $t \in \mathbb{R}$, under T , where $a = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$
- (ii) Verify that the image of the point $a + b$ on the line lies on the image of the line.

8. Plot the line and the image of the line in Q(7).

9. Let T be a matrix transformation defined by $T(x) = Ax$. Find a vector x whose image under T is b and determine whether x is unique or not.

(i) $A = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$

(ii) $A = \begin{pmatrix} -3 & 7 & 5 \\ -5 & 9 & 3 \end{pmatrix}$, $b = \begin{pmatrix} -2 \\ -6 \end{pmatrix}$

(iii) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & -2 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$

(iv) $A = \begin{pmatrix} 1 & 3 & 9 & 2 \\ 1 & 0 & 3 & -4 \\ 0 & 1 & 2 & 3 \\ -2 & 3 & 0 & 5 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ 3 \\ -1 \\ 3 \end{pmatrix}$

10. Let u and v be vectors in \mathbb{R}^n . Show that the set P of all points in the parallelogram determined by u and v has the form $au + bv$, for $0 \leq a \leq 1$, $0 \leq b \leq 1$. Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Show that the image of point in P under the transformation T lies in the parallelogram determined by $T(u)$ and $T(v)$.

11. Let T be a transformation defined by $T(X) = AX$. Find a vector X whose image under T is b , and determine whether X is unique or not.

(i) $A = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$

(ii) $A = \begin{pmatrix} -3 & 7 & 5 \\ -5 & 9 & 3 \end{pmatrix}$, $b = \begin{pmatrix} -2 \\ -6 \end{pmatrix}$

(iii) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & -2 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$

(iv) $A = \begin{pmatrix} 1 & 3 & 9 & 2 \\ 1 & 0 & 3 & -4 \\ 0 & 1 & 2 & 3 \\ -2 & 3 & 0 & 5 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ 3 \\ -1 \\ 3 \end{pmatrix}$

12. Let T be a transformation defined by $T(X) = AX$. Does c lie in the range of T ? If yes, find all vectors X whose image is c .

$$\begin{aligned}
 \text{(i)} \quad A \text{ as in Q11} \quad c &= \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \\
 \text{(ii)} \quad A \text{ as in Q13} \quad c &= \begin{pmatrix} 4 \\ 4 \\ 3 \end{pmatrix} \\
 \text{(iii)} \quad A \text{ as in Q14} \quad c &= \begin{pmatrix} 3 \\ 1 \\ 1 \\ 3 \end{pmatrix} \\
 \text{(iv)} \quad A \text{ as in Q15} \quad c &= \begin{pmatrix} 6 \\ -2 \\ -6 \end{pmatrix}
 \end{aligned}$$

13. Find all vectors X that are mapped in to the zero vector by the transformation $X \mapsto AX$, for the given matrix A .

$$\begin{aligned}
 \text{(i)} \quad A &= \begin{pmatrix} 2 & -3 & 4 \\ 1 & 0 & -3 \end{pmatrix} \\
 \text{(ii)} \quad A &= \begin{pmatrix} 3 & -4 \\ 5 & 2 \\ -1 & 3 \end{pmatrix} \\
 \text{(iii)} \quad A &= \begin{pmatrix} 1 & -2 & 1 & -1 \\ 1 & 1 & -2 & 3 \\ 4 & 1 & -5 & 8 \\ 5 & -7 & 2 & -1 \end{pmatrix} \\
 \text{(iv)} \quad A &= \begin{pmatrix} 1 & 1 & -3 & 2 \\ 2 & -1 & 2 & -3 \\ 3 & -2 & 1 & -4 \\ -4 & 1 & -3 & 1 \end{pmatrix}
 \end{aligned}$$

14. Let T be the transformation defined by matrix A . Determine whether

(a) T is onto. (b) T is injective.

$$\begin{aligned}
 \text{(i)} \quad A &= \begin{pmatrix} 2 & -3 & 1 \\ 1 & 2 & -3 \\ 4 & -1 & -2 \end{pmatrix} && \text{onto, injective.} \\
 \text{(ii)} \quad A &= \begin{pmatrix} 4 & 5 & 6 \\ 5 & 6 & -7 \\ 7 & 8 & 9 \end{pmatrix} && \text{not onto, not injective.} \\
 \text{(iii)} \quad A &= \begin{pmatrix} 2 & 1 & 3 & -5 \\ 0 & 7 & 3 & -7 \\ -3 & -4 & 2 & 0 \end{pmatrix} && \text{onto, not injective.} \\
 \text{(iv)} \quad A &= \begin{pmatrix} 1 & 6 & 3 & 8 \\ 2 & 4 & 6 & -1 \\ 3 & 10 & 9 & 7 \\ 4 & 16 & 12 & 15 \end{pmatrix} && \text{not onto, not injective.} \\
 \text{(v)} \quad A &= \begin{pmatrix} 1 & -1 & 1 \\ -3 & 1 & -4 \\ 7 & -3 & -9 \\ 4 & -2 & -5 \end{pmatrix} && \text{not onto, injective.}
 \end{aligned}$$

15. A transformation T is defined by $X \mapsto AX$

If the echelon form of A is given, determine whether the mapping is onto and/or injective.

$$\begin{aligned}
 \text{(i)} & \begin{pmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \\ \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 \text{(ii)} & \begin{pmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 \end{pmatrix} \\
 \text{(iii)} & \begin{pmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & \blacksquare \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
 \text{(iv)} & \begin{pmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 \text{(v)} & \begin{pmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

16. Find all $X \in \mathbb{R}^4$ which are mapped to zero, by the transformation $X \mapsto AX$ for the given matrix A ,

$$\begin{aligned}
 \text{(i)} & \begin{pmatrix} 0 & 1 & -4 & 3 \\ 1 & -2 & -1 & 1 \\ 1 & -4 & 7 & -5 \end{pmatrix} \\
 \text{(ii)} & \begin{pmatrix} -1 & 6 & 9 & 7 \\ 1 & -3 & -3 & -1 \\ -1 & 2 & 1 & 4 \\ 0 & 7 & 14 & 6 \end{pmatrix} \\
 \text{(iii)} & \begin{pmatrix} 1 & -1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 13 & -1 \\ -1 & 3 & 2 & -10 \\ 3 & -3 & 13 & 3 \end{pmatrix} \\
 \text{(iv)} & \begin{pmatrix} 1 & 2 & -1 & 0 \\ -3 & 0 & 1 & 2 \end{pmatrix}
 \end{aligned}$$

12.6 Linear Transformation

If T is a matrix transformation determined by a matrix A , then $T(u+v) = A(u+v) = Au + Av = T(u) + T(v)$ and $T(cu) = A(cu) = cA(u) = cT(u)$ for all u, v in \mathbb{R}^n and c in \mathbb{R} .

These properties of T lead us to a special class of transformations, called linear transformation, which we shall study in this section.

Definition 12.5. (Linear transformation):

A transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called a linear transformation if

- (i) $T(u+v) = T(u) + T(v)$ for all $u, v \in \mathbb{R}^n$. (T preserves vector addition)
- (ii) $T(cu) = cT(u)$ for all $u \in \mathbb{R}^n$, $c \in \mathbb{R}$. (T preserves scalar multiplication)

Definition 12.6. (Linear transformation):

A linear transformation of \mathbb{R}^n into itself is called a linear operator.

Linear transformations form a very important class of transformations in linear algebra. These transformations are called linear as they preserve linearity, that is to say, that under a non-zero linear transformation, lines are mapped onto lines and planes are mapped onto planes. This will be proved later. As an important consequence of the definition of a linear transformation we have the following result.

Theorem 12.1. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Then

- (i) $T(0) = 0$
- (ii) $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$
- (iii) $T(u - v) = T(u) - T(v)$
- (iv) $T(-v) = -T(v)$
- (v) $T(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha_1 T(u_1) + \dots + \alpha_n T(u_n)$.

Proof:

- (i) Let $u \in \mathbb{R}^n$. Then

$$T(0) = T(0u) = 0T(u) \quad \because T \text{ is a linear transformation.}$$

$$= 0$$
 Hence $T(0) = 0$.
- (ii) Let $u, v \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$. Then

$$T(\alpha u + \beta v) = T(\alpha u) + T(\beta v)$$

$$= \alpha T(u) + \beta T(v)$$
 Hence $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$.
- (iii) In particular, for $\alpha = 1$, $\beta = -1$ we get $T(u - v) = T(u) - T(v)$. Thus T preserves subtraction.
- (iv) For $\alpha = 0$, $\beta = -1$ in (ii) we get $T(-1v) = -1 T(v)$ i.e. $T(-v) = -T(v)$.
- (v) Repeated application of (ii) of the above theorem gives $T(\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k) = \alpha_1 T(u_1) + \dots + \alpha_k T(u_k)$ for $u_1, \dots, u_k \in \mathbb{R}^n$, $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. \square

Remark 12.1. A direct consequence of the theorem is:

If $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a mapping such that $T(0_{\mathbb{R}^n}) \neq 0_{\mathbb{R}^m}$, then T is not a linear transformation.

How to prove non-linearity?

When we want to disprove linearity of a given transformation that is, to prove that a transformation is not linear, we need to find only one counter-example. That is, if we can find just one case in which the transformation does not preserve addition, scalar multiplication, or the zero vector, we can conclude that the transformation is not linear.

Mathematically a transformation T is not linear if

- (i) There exist some vectors u and v such that $T(u + v) \neq T(u) + T(v)$ or
- (ii) There exists a vector u and a scalar c such that $T(cu) \neq cT(u)$ or
- (iii) $T(0) \neq 0$.

Example 12.4. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be defined by $T(x, y, z) = (x + y, x - z)$

If $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$, $\alpha \in \mathbb{R}$ then $T((x_1, y_1, z_1) + (x_2, y_2, z_2)) =$

$$\begin{aligned}
T((x_1 + x_2, y_1 + y_2, z_1 + z_2)) &= (x_1 + x_2 + y_1 + y_2, x_1 + x_2 - z_1 - z_2) \\
&= (x_1 + y_1, x_1 - z_1) + (x_2 + y_2, x_2 - z_2) \\
&= T(x_1, y_1, z_1) + T(x_2, y_2, z_2) \\
T(\alpha(x_1, y_1, z_1)) &= T((\alpha x_1, \alpha y_1, \alpha z_1)) \\
&= (\alpha x_1 + \alpha y_1, \alpha x_1 - \alpha z_1) \\
&= \alpha(x_1 + y_1, x_1 - z_1) \\
&= \alpha T((x_1, y_1, z_1))
\end{aligned}$$

Hence T is a linear transformation.

Example 12.5. Define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(x, y) = (x^2, 2y)$

Here $T(0, 0) = (0, 0)$.

Thus $0 \in \mathbb{R}^2$ is mapped to 0.

But $(1, 2), (2, 1) \in \mathbb{R}^2$

$$T((1, 2) + (2, 1)) = T((3, 3)) = (9, 6)$$

$$T((1, 2)) + T((2, 1)) = (1, 4) + (4, 2) = (5, 6)$$

Hence $T((1, 2) + (2, 1)) \neq T((1, 2)) + T((2, 1))$, so that T is not a linear transformation.

Example 12.6. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be defined by $T(x, y, z) = (x + 1, y - x)$

In this case $T(0, 0, 0) = (1, 0) \neq (0, 0)$

Hence we can conclude that T is not a linear transformation.

Example 12.7. Consider the transformation

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^3 \\ y^2 \end{pmatrix}$$

We suspect it is not linear (components in the image vector $T(X)$ must homogeneous of degree 1, i.e. all terms must be only first degree in components of vector X). To prove it is not linear, take the vector

$$v = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

then

$$T(2v) = \begin{pmatrix} 64 \\ 16 \end{pmatrix}, \quad 2T(v) = \begin{pmatrix} 16 \\ 8 \end{pmatrix}$$

Thus $T(2v) \neq 2T(v)$, so T is not linear.

Note that in above example $T(0) = 0$, but T is not linear. Thus $T(0) = 0$ is only necessary condition for linearity but not sufficient.

If a linear transformation T is defined on a set S of vectors, then T can be extended to the linear span of S . In particular e_1, e_2, \dots, e_n denote the column of the identity matrix and we are given $T(e_1), T(e_2), \dots, T(e_n)$, then $T(X)$ can be found for any vector X in R^n . This is because X can be expressed as a linear combination of the columns e_1, e_2, \dots, e_n .

Example 12.8. Given $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ such that $T(e_i) = f_i$, $i = 1, 2, 3$,

where $f_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 4 \end{pmatrix}$, $f_2 = \begin{pmatrix} 2 \\ -3 \\ 3 \\ 0 \end{pmatrix}$, $f_3 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ -3 \end{pmatrix}$. Find $T \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$ and

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

$$\begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} = 3e_1 + (-2)e_2 + 4e_3$$

$$T \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} = T(3e_1 + (-2)e_2 + 4e_3) = 3T(e_1) + (-2)T(e_2) + 4T(e_3) \dots (T \text{ is linear})$$

$$= 3f_1 + (-2)f_2 + 4f_3 = 3 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 4 \end{pmatrix} + (-2) \begin{pmatrix} 2 \\ -3 \\ 3 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 2 \\ 1 \\ -3 \end{pmatrix}$$

$$= \begin{pmatrix} -7 \\ 17 \\ -2 \\ 0 \end{pmatrix}$$

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = T(xe_1 + ye_2 + ze_3) = xT(e_1) + yT(e_2) + zT(e_3) = xf_1 + yf_2 + zf_3$$

$$= x \begin{pmatrix} -1 \\ 1 \\ 0 \\ 4 \end{pmatrix} + y \begin{pmatrix} 2 \\ -3 \\ 3 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 2 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -x + 2y \\ x - 3y + 2z \\ 3y + z \\ 4x - 3z \end{pmatrix}$$

This shows that a linear transformation on \mathbb{R}^n is completely determined by the images of the columns of the identity matrix I_n .

Geometrical Properties of Linear Transformation

Remark 12.2. In view of the fact that a linear transformation preserves addition and scalar multiplication, we see the effect of a linear transformation on some geometrical figures in \mathbb{R}^2

- (i) A linear transformation maps a line onto a line.
- (ii) A linear transformation maps a line segment to a line segment.
- (iii) A linear transformation maps a parallelogram to a parallelogram.

Proof:

- (i) Let L be any line in \mathbb{R}^n and $T(L)$ be its image under T .

Equation of L is of the form $x = a + tb$, t is a parameter for some vectors a and b in \mathbb{R}^n . Let P be any point on $T(L)$. Then P is the image of some point Q on L . If points P and Q are tips of vectors x' and x respectively then $x' = T(x)$. Now $x = a + tb$, $\therefore x' = T(x) = T(a + tb) = T(a) + tT(b) = a' + tb'$ where $a' = T(a)$ and $b' = T(b)$. Hence equation of $T(L)$ is of the form $x = a' + tb'$ where t is a parameter. Which is the equation of a line.

- (ii) Let AB be a segment of a line whose equation is $x = a + tb$, t is a parameter. Let vector $OA = a + t_1b$ and $OB = a + t_2b$. Then segment AB is the set of points $a + tb$ where $t_1 \leq t \leq t_2$. Image of AB under T is the set of points $T(a + tb) = T(a) + tT(b)$ where $t_1 \leq t \leq t_2$ which is the line segment joining points $T(a) + t_1T(b)$ and $T(a) + t_2T(b)$.
- (iii) Let $ABCD$ be a parallelogram with a and b as adjacent sides. Let P be any point inside the parallelogram then AP is the diagonal of the parallelogram $ALPM$ for some points L and M respectively. Then $\overrightarrow{AP} = \overrightarrow{AL} + \overrightarrow{AM}$. Since L lies on $AB (= a)$, $\overrightarrow{AL} = ka$ for some $0 \leq k \leq 1$. Similarly $\overrightarrow{AM} = mb$ for some $0 \leq m \leq 1$. This gives that $\overrightarrow{AP} = ka + mb$. \square

Problem 12.8. Check whether the following mappings are linear transformations

- (i) $T_1 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, where $T_1(x, y, z) = (0, x + y, y - z)$
(ii) $T_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, where $T_2(x, y) = (x^2, x + y)$
(iii) $T_3 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, where $T_3(x, y) = (x + 1, y)$

Solution:

- (i) Let $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbb{R}^3$, $\alpha \in \mathbb{R}$
Then $T_1((x_1, y_1, z_1) + (x_2, y_2, z_2)) = T_1((x_1 + x_2, y_1 + y_2, z_1 + z_2))$
 $= (0, x_1 + x_2 + y_1 + y_2, y_1 + y_2 - z_1 - z_2)$
 $= (0, x_1 + y_1, y_1 - z_1) + (0, x_2 + y_2, y_2 - z_2)$
 $= T_1((x_1, y_1, z_1)) + T_1((x_2, y_2, z_2))$
 $T_1(\alpha(x_1, y_1, z_1)) = T_1((\alpha x_1, \alpha y_1, \alpha z_1))$
 $= (0, \alpha x_1 + \alpha y_1, \alpha y_1 - \alpha z_1) = \alpha(0, x_1 + y_1, y_1 - z_1)$
 $= \alpha T_1((x_1, y_1, z_1))$
Hence T_1 is a linear transformation.
- (ii) $(1, 2), (3, 4) \in \mathbb{R}^2$ (Choose $(x_1, y_1), (x_2, y_2)$ such that $(x_1 + x_2)^2 \neq x_1^2 + x_2^2$).
 $T_2((1, 2) + (3, 4)) = T_2((4, 6)) = (16, 10)$
 $T_2((1, 2)) + T_2((3, 4)) = (1, 3) + (9, 7) = (10, 10)$
Thus $T_2((1, 2) + (3, 4)) \neq T_2((1, 2)) + T_2((3, 4))$
Hence T_2 is not a linear transformation.
- (iii) $T_3((0, 0)) = (1, 0) \neq (0, 0)$. $\therefore T(0, 0) \neq 0$, so that T_3 is not a linear transformation.

12.7 Exercise

- Examine whether the following mappings $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ are linear transformations.
 - $T((x \ y \ z)^t) = (x \ y)^t$
 - $T((x \ y \ z)^t) = (x + 1 \ y - x)^t$
 - $T((x \ y \ z)^t) = (x + y \ 0)^t$
 - $T((x \ y \ z)^t) = (x + y \ 2x - z)^t$
 - $T((x \ y \ z)^t) = (xy^2 \ z - x)^t$
 - $T((x \ y \ z)^t) = (x + z \ y + z)^t$
- Examine whether the following mappings $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ are linear transformations. Also give the geometrical interpretation of those which are

linear transformations.

- (i) $T((x \ y \ z)^t) = k(x \ y \ z)^t$, k is a fixed constant > 1
- (ii) $T((x \ y \ z)^t) = k(x \ y \ z)^t$, k is a fixed constant < 1 .
- (iii) $T((x \ y \ z)^t) = (x \ 0 \ 0)$
- (iv) $T((x \ y \ z)^t) = (0 \ y \ z)$
- (v) $T((x \ y \ z)^t) = (-x \ -y \ -z)$
- (vi) $T((x \ y \ z)^t) = (x \ y \ -z)$
- (vii) $T((x \ y \ z)^t) = (-x \ -y \ z)$
- (viii) $T((x \ y \ z)^t) = (x \ y^2 \ z)$

3. Verify whether the mappings defined in Q3, 4, 5 of exercise are linear transformations or not.

4. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be a linear transformation. Let $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$, $T(u) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $T(v) = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$. Find $T(3u + 4v)$, $T(u - 2v)$, $T(0)$.

5. Does there exist a linear transformation T such that $T \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $T \begin{pmatrix} -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$? Justify your answer.

6. Does there exist a linear transformation T such that $T \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $T \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $T \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$? Justify your answer.

7. A linear transformation $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ is such that $T(e_1) = u_1$, $T(e_2) = u_2$, $T(e_3) = u_3$, where $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$; $u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $u_3 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$. Obtain (i) $T(0)$ (ii) $T(v)$, where $v = \begin{pmatrix} 3 \\ -4 \\ 5 \end{pmatrix}$, (iii) $T(X)$ for any $X \in \mathbb{R}^3$

8. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a linear transformation such that $T((1, 0, 0)^t) = (2, -1)^t$, $T((0, 1, 0)^t) = (3, 1)^t$, $T((0, 0, 1)^t) = (-1, 2)^t$, find $T((-3, 4, 2)^t)$ and $T((x, y, z)^t)$.

9. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation such that $T((1, 0)^t) = (2, 1)^t$, $T((1, 1)^t) = (0, 1)^t$ find $T((0, 1)^t)$ and $T((x, y)^t)$.

10. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a linear transformation such that $T((1, 1, 0)^t) = (2, -1)$, $T((0, 1, 1)^t) = (3, 2)$, $T((0, 0, 1)^t) = (1, -1)$, find

- (i) $T((1, 3, -2)^t)$
- (ii) $T((x, y, z)^t)$.

12.8 The Matrix of a Linear Transformation

We shall now see that any linear transformation is indeed a matrix transformation. Describing a linear transformation T means having a formula for $T(X)$. In fact, if $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation and we know the effect of T on the columns e_1, e_2, \dots, e_n of I_n , then we will prove that T is completely determined. Moreover, we will prove that with every T , we can associate a matrix A such that T is actually the matrix transformation $X \rightarrow AX$. Let us illustrate this by an example, before proceeding to the proof.

Example 12.9. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a linear transformation such that $T(e_1) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $T(e_2) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, $T(e_3) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, where e_1, e_2, e_3 are the columns of I_3 and $a_1, a_2, b_1, b_2, c_1, c_2$ are any real numbers. Can we find $T(X)$ for any $X \in \mathbb{R}^3$? The answer is yes. We shall express any $X \in \mathbb{R}^3$ in terms of e_1, e_2, e_3 .

$$\text{If } X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} = x_1 e_1 + x_2 e_2 + x_3 e_3 \quad \dots (1)$$

$$\begin{aligned} \text{then } T(X) &= T(x_1 e_1 + x_2 e_2 + x_3 e_3) \\ &= x_1 T(e_1) + x_2 T(e_2) + x_3 T(e_3) \quad \dots (2) \\ &= x_1 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + x_2 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + x_3 \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 a_1 + x_2 b_1 + x_3 c_1 \\ x_1 a_2 + x_2 b_2 + x_3 c_2 \end{pmatrix} \quad \dots (3) \end{aligned}$$

This shows that a knowledge of $T(e_1)$, $T(e_2)$, and $T(e_3)$ is sufficient to obtain $T(X)$ for any $X \in \mathbb{R}^3$.

In fact, (3) can be written as

$$T(X) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = [T(e_1) \ T(e_2) \ T(e_3)]X$$

$T(X) = AX$ (say).

Thus A is a matrix associated with the linear transformation T . This shows that T is the matrix transformation $X \mapsto AX$. Note that the columns of A are $T(e_1)$, $T(e_2)$, $T(e_3)$.

Two linear transformations T_1 and T_2 defined from \mathbb{R}^n to \mathbb{R}^m are same (or equal) if $T_1(X) = T_2(X)$ for all $X \in \mathbb{R}^n$. Thus T_1 and T_2 will be equal when their corresponding matrices are the same.

We shall now consider a general linear transformation T from \mathbb{R}^n to \mathbb{R}^m .

Theorem 12.2. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a given linear transformation. Then there exists a unique matrix A such that $T(X) = AX$ for all $X \in \mathbb{R}^n$.

Proof: Existence

Since T is given, therefore $T(e_1), \dots, T(e_n)$ are known, where e_1, e_2, \dots, e_n are the columns of I_n , the $n \times n$ unit matrix. Since $T(e_j) \in \mathbb{R}^m$, let

$$T(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \text{ for } j = 1, 2, \dots, n$$

Let $X \in \mathbb{R}^n$. Then

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

$$T(X) = T(x_1 e_1 + x_2 e_2 + \dots + x_n e_n)$$

$$= x_1 T(e_1) + x_2 T(e_2) + \dots + x_n T(e_n) \quad \because T \text{ is linear.}$$

$$= [T(e_1) \dots T(e_n)] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= AX \text{ where } A = (a_{ij})_{m \times n}$$

Thus $T(X) = AX$ for a $m \times n$ matrix A , whose columns are $T(e_1), T(e_2), \dots, T(e_n)$.

Uniqueness

If A, B are two $m \times n$ matrices such that $T(X) = A(X)$ also $AX = BX$ for all $X \in \mathbb{R}^n$. Then comparing the images of e_1, e_2, \dots, e_n we get $Ae_i = Be_i \forall i, 1 \leq i \leq n$.

\Rightarrow i th column of $A = i$ th column of $B \forall i, 1 \leq i \leq n$.

$\Rightarrow A = B$. □

The matrix A obtained above is called the *standard matrix for the linear transformation T* . Thus we also get that every linear transformation is a matrix transformation and vice versa.

It is important to note that if T is not a linear transformation then $A = [T(e_1) \ T(e_2) \ \dots \ T(e_n)]$ will not implement T , that is, $X \mapsto AX$ is not T .

Working rules: To obtain the standard matrix for a given linear transformation T from \mathbb{R}^n to \mathbb{R}^m .

Step 1 Obtain $T(e_1) \ T(e_2) \ \dots \ T(e_n)$ the images of the columns of I_n , as column vectors of \mathbb{R}^m .

Step 2 Let $A = [T(e_1) \ T(e_2) \ \dots \ T(e_n)]$ be the matrix whose columns are $T(e_1) \ T(e_2) \ \dots \ T(e_n)$ respectively. Then A is the required $m \times n$ matrix.

Example 12.10. Find the standard matrix for the linear transformation T :

$$\mathbb{R}^3 \rightarrow \mathbb{R}^4 \text{ defined by } T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x + y \\ x - y + z \\ -3x + 2z \\ y - 13z \end{pmatrix}$$

$$\text{Step 1 Here } T(e_1) = T\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 1 \\ -3 \\ 0 \end{pmatrix}$$

$$T(e_2) = T\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

$$T(e_3) = T\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \\ 2 \\ -13 \end{pmatrix}$$

Step 2 $A = [T(e)_1 \ T(e)_2 \ T(e)_3]$

$$= \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ -3 & 0 & 2 \\ 0 & 1 & -13 \end{pmatrix}$$

is the required matrix. It is 4×3 matrix.

Example 12.11. $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 x_2 \\ x_2 \end{pmatrix}$

$$T(e_1) = T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad T(e_2) = T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

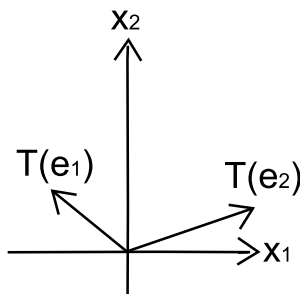
$$\text{If } A = [T(e_1) \ T(e_2)] = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$\text{Then } AX = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \end{pmatrix} \neq T(X).$$

Thus $[T(e_1) \ T(e_2)]$ does not implement T .

12.9 Exercises

- Find a matrix that implement the following transformations and hence prove that they are linear. Note that x_1, x_2, \dots are not vectors but are entries in vectors.
 - $T(x_1, x_2, x_3) = (x_1 + x_2 - x_3, -x_1 + 2x_2 + 3x_3)$
 - $T(x_1, x_2) = (2x_1, 3x_2, 2x_1 - 3x_2)$
 - $T(x_1, x_2, x_3, x_4) = (x_1 - 2x_2 + x_3, x_2 - 3x_3 + x_4)$
- Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation such that $T(e_1)$ and $T(e_2)$ are the vectors shown in the figure. Using the figure sketch the vector $T(-1, 2)$.



3. Find the standard matrix representing the following linear transformations.

(i) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x - y \\ x + y \end{pmatrix}$
 (ii) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(u) = 2u$, $u \in \mathbb{R}^2$
 (iii) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by $T((x \ y \ z)^t) = (2x - 3y \ -4y)^t$
 (iv) $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ defined by $T((x \ y \ z \ w)^t) = (x - y \ x - w \ x - z \ x + y + z)^t$

4. If the standard matrix of a linear transformation is $\begin{pmatrix} 1 & -2 & 3 & -1 \\ 0 & 1 & 2 & 4 \\ -1 & -3 & 2 & 0 \end{pmatrix}$,

find $T\left(\begin{pmatrix} 1 \\ -1 \\ 2 \\ -3 \end{pmatrix}\right)$

5. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by $T((x \ y \ z)^t) = (x + y \ y - z \ z + x)^t$. Find the standard matrix of T .

6. Find the matrix of the linear transformation of \mathbb{R}^2 defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix}$$

Plot $(x \ y)^t$ and its image under T for

- (i) $x = 2$, $y = 3$
 (ii) $x = -5$, $y = 6$
 (iii) $x = 8$, $y = -4$
 (iv) $x = -3$, $y = -5$

Give the geometrical interpretation of T .

7. If the standard matrix of a linear transformation is $\begin{pmatrix} -1 & 2 & 3 \\ 5 & 0 & 5 \end{pmatrix}$, find the linear transformation T .

8. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined such that $T((1 \ 2)^t) = (1 \ 0 \ 0)^t$

$$T((2 \ 3)^t) = (0 \ 1 \ 0)^t$$

Find

- (i) $T((1 \ 0)^t)$
 (ii) $T((0 \ 1)^t)$
 (iii) Standard matrix for T
 (iv) $T((x \ y)^t)$

9. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a transformation defined by $T(x_1, x_2) = (x_1 + 1, x_2)$. If $A = [T(e_1), T(e_2)]$, show that A does not implement T . Give reasons.

10. Fill in the entries in the matrix assuming that the equation holds for all values of the variables.

$$\begin{aligned}
 \text{(i)} \quad & \begin{pmatrix} - & - & - & - \\ - & - & - & - \\ - & - & - & - \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_3 + x_4 \\ -x_1 + x_2 - 3x_4 \\ x_2 - 3x_3 + 4x_4 \end{pmatrix} \\
 \text{(ii)} \quad & \begin{pmatrix} - & - \\ - & - \\ - & - \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - 3x_2 \\ -x_1 + 2x_2 \\ x_1 + x_2 \end{pmatrix} \\
 \text{(iii)} \quad & \begin{pmatrix} - & - & - \\ - & - & - \\ - & - & - \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 - x_3 \\ x_3 - x_1 \\ x_1 - x_2 \end{pmatrix}
 \end{aligned}$$

12.10 Geometric Transformations of \mathbb{R}^2 and \mathbb{R}^3

We now study some basic transformations of \mathbb{R}^2 namely scaling, projection, reflection, shearing and rotation.

Translation is a transformation which is not linear so it cannot be described by a matrix.

Scaling

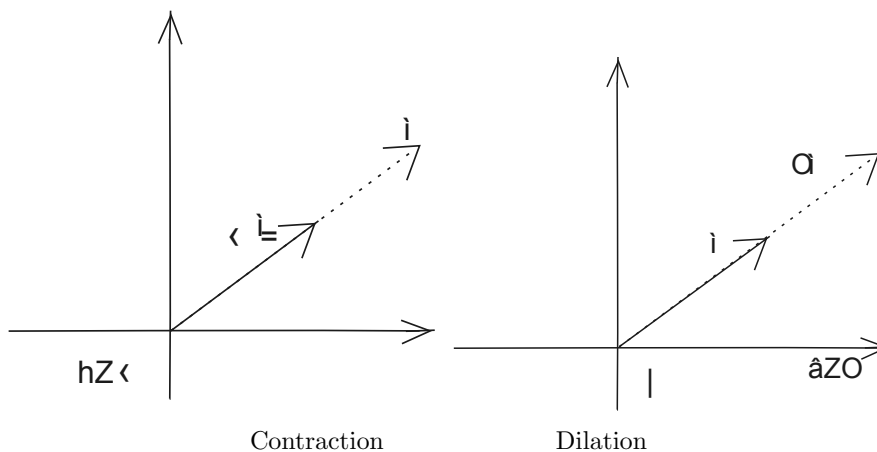
Scaling is a transformation which increases or decreases the length of a vector.

Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ or $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by

$$T(X) = kX, \text{ where } k \in \mathbb{R}.$$

For every value of k , T is a linear operator (verify!). k is called the scaling factor.

(1) If $0 < k < 1$, then T is called contraction and if $k > 1$, T is called a dilation. Contraction shrinks the vector whereas dilation expands (or dilates) the vector.



Similarly if we define $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by

$T(X) = kX$, for all $X \in \mathbb{R}^3$, where k is fixed real number, then T is linear operator, and k is called scaling factor.

(2) If $0 < k < 1$, then T is called contraction and if $k > 1$, T is called a dilation.

For $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$

Matrix of scaling (Dilation/Contraction)

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T(e_1) = \begin{pmatrix} k \\ 0 \end{pmatrix}$$

$$e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 0 \\ k \end{pmatrix}$$

Matrix of $T = [T(e_1) \ T(e_2)]$

$$= \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

For $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3: X \rightarrow kX$

$$T \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} k \\ 0 \\ 0 \end{pmatrix}, \quad T \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ k \\ 0 \end{pmatrix}, \quad T \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ k \end{pmatrix}$$

$$\text{The matrix of } T = \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix}$$

Example 12.12. Given points, $A(3, 6)$, $B(0, 2)$, $C(6, 1)$ of $\triangle ABC$ find its image under

(i) contraction by a factor of $\frac{1}{3}$.

(ii) dilation by a factor of 3.

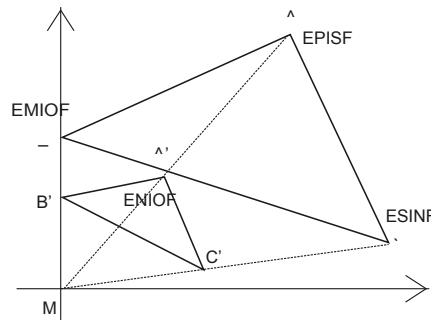
Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(X) = \frac{1}{3}X$

$$\text{Point } A = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad T(A) = T \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = A'(\text{say})$$

$$B = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad T(B) = T \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2/3 \end{pmatrix} = B'(\text{say})$$

$$C = \begin{pmatrix} 6 \\ 1 \end{pmatrix}, \quad T(C) = \begin{pmatrix} 2 \\ 1/3 \end{pmatrix} = C'(\text{say})$$

Then $A'B'C'$ is the image of ABC under contraction.



Point A is $(3, 6)^t$.

$$T(A) = 3(3, 6)^t = \begin{pmatrix} 9 \\ 18 \end{pmatrix} = A'(\text{say})$$

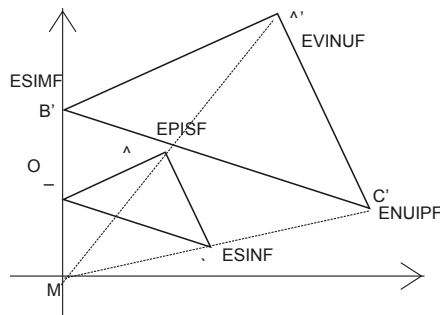
Point B is $(0, 2)^t$.

$$T(B) = (0, 6)^t = B'(\text{say})$$

Point C is $(6, 1)^t$.

$$T(C) = (18, 3)^t = C'(\text{say}).$$

Thus $A'B'C'$ is the image of ABC under dilation.



Projection

Let us take the projection of vectors in \mathbb{R}^2 to vectors on the x -axis. Let's call this transformation T .

Then $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

Clearly this is linear.

Scalar multiplication is preserved

We wish to show that for all vectors X and all scalars λ , $T(\lambda X) = \lambda T(X)$.

$$\text{Let } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{Then } \lambda X = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}$$

$$\text{Now } T(\lambda X) = T \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \lambda T(X)$$

This is the same vector as above, so under the transformation T , scalar multiplication is preserved.

Addition is preserved

We wish to show for all vectors X and Y , $T(X + Y) = T(X) + T(Y)$.

$$\text{Let } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{and } Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\text{Now } T(X + Y) = T \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) = T \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ 0 \end{pmatrix}$$

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} y_1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ 0 \end{pmatrix}$$

So we have that the transformation T preserves addition.

We have shown T preserve addition and scalar multiplication. So T must be linear.

This mapping is called projection of \mathbb{R}^2 on to x -axis.

$$\text{Since } Te_1 = T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad Te_2 = T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

\therefore the matrix of T is $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. This is the matrix of projection on to X -axis.

Similarly $T_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$T_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ y_1 \end{pmatrix} \text{ is the projection of } \mathbb{R}^2 \text{ on the } Y\text{-axis.}$$

The matrix of T_1 is $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. This is the matrix of projection on to Y -axis.

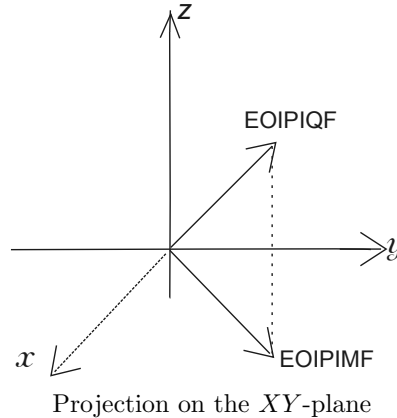
In \mathbb{R}^3 the mappings $T_1 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ defined by

$$(i) \quad T_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \text{ is the projection on the } XY\text{-plane.}$$

$$(ii) \quad T_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} \text{ is the projection on the } XZ\text{-plane.}$$

$$(iii) \quad T_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} \text{ is the projection on the } YZ\text{-plane.}$$

Similarly we can define projection on the X -axis, Y -axis and Z -axis.



Reflection

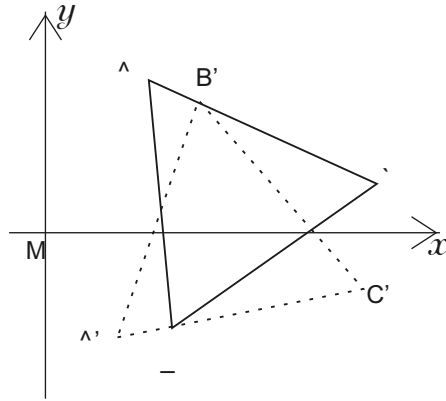
$$\text{Let } T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \text{ be defined by } T \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x \\ -y \end{pmatrix}$$

Then T is a linear operator called the reflection of \mathbb{R}^2 with respect to the X -axis.

$$\text{Since } T(e_1) = T \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$T(e_2) = T \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

\therefore Matrix of T is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This is the matrix of reflection in the X -axis.



ABC gives triangle.
 $A \rightarrow A', B \rightarrow B', C \rightarrow C'$
 $A'B'C'$ reflection in X -axis.

Similarly $T_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$T_2\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -x \\ y \end{pmatrix}$ is the reflection of \mathbb{R}^2 with respect to the Y -axis. The

matrix of reflection in y -axis is $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

$T_3 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T_3\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -x \\ -y \end{pmatrix}$

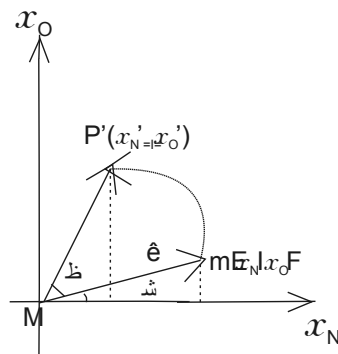
is the reflection of \mathbb{R}^2 w.r.t the origin. The matrix of reflection through the origin is $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

Rotation

Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the transformation that rotates each vector in \mathbb{R}^2 about origin through an angle φ counterclockwise. Then T is a linear transformation.

Let $u = \vec{OP} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be any vector in \mathbb{R}^2 , under rotation about φ , suppose

\vec{OP} maps to $\vec{OP}' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$



N=====O

Suppose $OP = r$ and \vec{OP} makes an angle θ with x_1 -axis, we see from figure that

$$x_1 = r \cos \theta \quad x_2 = r \sin \theta \quad \dots (1)$$

$$x'_1 = r \cos(\theta + \varphi) \quad x'_2 = r \sin(\theta + \varphi) \quad \dots (2)$$

Using the formula for sine and cosine of sum of angles, equations (2) become

$$x'_1 = r \cos \theta \cos \varphi - r \sin \theta \sin \varphi$$

$$x'_2 = r \sin \theta \cos \varphi + r \cos \theta \sin \varphi \quad \dots (3)$$

Using (1), equations (3) can be written as

$$x'_1 = x_1 \cos \varphi - x_2 \sin \varphi$$

$$x'_2 = x_2 \cos \varphi + x_1 \sin \varphi$$

$$\text{Thus } T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} x_1 \cos \varphi - x_2 \sin \varphi \\ x_1 \sin \varphi + x_2 \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Au, \text{ where } A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

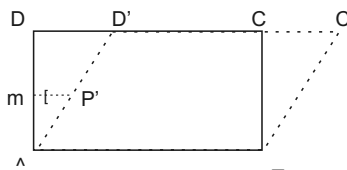
This proves that T is a matrix transformation. Since every matrix transformation is linear therefore T is linear.

We see from above that the matrix of rotation about origin through an angle φ is

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Shear Transformation

Consider a rectangle $ABCD$. Keeping AB fixed, if it is pushed parallel to AB , then it takes the shape $ABC'D'$ (see figure).

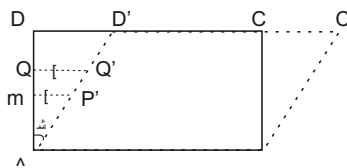


Such a transformation is called a shear. Note that every point is not moved by the same distance. The point P is moved to P' , D is moved to D' . The distance moved by a point is proportional to its distance from the fixed line. How far a point is pushed is determined by its shearing factor. Thus we have the following definition.

Definition 12.7. (Shear along a line):

A transformation in which all points along a given line L remain fixed, while other points are shifted parallel to L by a distance proportional to their perpendicular distance from L .

The distance a point P moves due to shear divided by the perpendicular distance of P from the invariant line is constant, and is called the shearing factor.



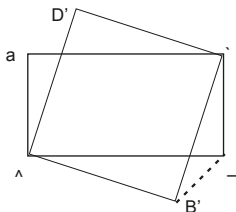
$ABCD$ is a square.

$$\frac{PP'}{AP} = \frac{QQ'}{AQ} = \frac{DD'}{AD} = \text{shearing factor.}$$

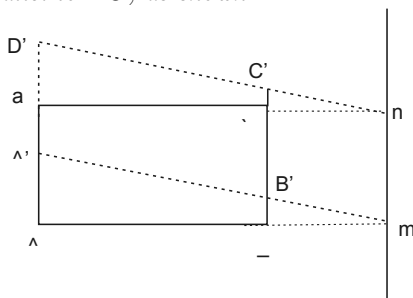
Shear transformations can also be generalized to three dimensional, in which planes are translated instead of lines. We shall restrict ourselves to two dimensional only.

It is important to note that a shear transformation is invertible. Moreover on applying a shear transformation to a plane figure, the area of the figure remains unchanged.

Example 12.13. Consider a rectangle $ABCD$. We would like to shear it along AC . Keeping AC fixed, every point is moved parallel to AC . Thus B is moved to B' and D to D' . A and C remain fixed. $AB'CD'$ is sheared figure.



Example 12.14. Consider a rectangle $ABCD$. We would like to shear it along a line L which is parallel to BC , as shown.



$A'B'C'D'$ is the sheared figure. How do we obtain it? Produce AB and CD to meet L in P and Q respectively. Then shear $PADQ$ along PQ . The corresponding figure is $PA'D'Q$. If $A'P$ intersects BC in B and $D'Q$ intersects BC produced in C' , then B has moved to B' , C to C' , D to D' and A to A' . Thus $ABCD$ has been sheared to $A'B'C'D'$.

Mathematically, shear can be defined as follows:

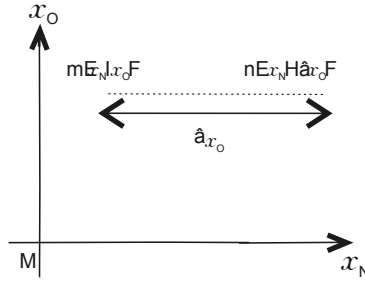
Definition 12.8. A shear in the x_1 - direction (or horizontal shear) is the linear operator T defined by

$$T(X) = T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 + kx_2 \\ x_2 \end{pmatrix}$$

where k is scalar. k is called the shear factor. The point $P(x_1, x_2)$ is moved parallel to the x_1 - axis to $Q(x_1 + kx_2, x_2)$.

The x_1 - coordinate is increased by an amount kx_2 , where as the x_2 - coordinate remains unchanged. See figure

Thus every point on the x_1 - axis is fixed.



$$T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} k \\ 1 \end{pmatrix}$$

$$\therefore \text{Matrix of } T = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

$$\text{Since } \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + kx_2 \\ x_2 \end{pmatrix} = T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Therefore T is also a matrix transformation and is consequently a linear transformation.

Definition 12.9. A shear in the x_2 -direction (or a vertical shear) in the linear operator T defined by

$$T(X) = T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ x_2 + kx_1 \end{pmatrix}$$

where k is scalar.

As above, the x_1 -coordinate is fixed whereas the x_2 -coordinate is increased by kx_1 . Every point on the x_2 -axis is fixed.

The standard matrix of $T = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$ and T is linear transformation follows as for horizontal shear.

Summarizing we have,

Type of shear	Transformation	Standard matrix	Figure	Remarks
Horizontal	$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 + kx_2 \\ x_2 \end{pmatrix}$	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$		x_1 coordinate changes.

Type of shear	Transformation	Standard matrix	Figure	Remarks
				x_2 coordinate is unchanged $e_1 \rightarrow e_1$, $e_2 \rightarrow e_2 + ke_1$
Vertical	$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rightarrow \begin{pmatrix} x_1 \\ x_2 + kx_1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$		x_1 coordinate unchanged. x_2 coordinate changes. $e_1 \rightarrow e_1 + ke_2$, $e_2 \rightarrow e_2$

Example 12.15. A shear in the x_1 -direction is defined by

$$T(X) = T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 + kx_2 \\ x_2 \end{pmatrix}, \text{ where } k \text{ is scalar.}$$

(i) Determine the standard matrix of T .

(ii) If this shear is applied on a rectangle with vertices $(0, 0)$, $(2, 0)$, $(2, 1)$ and $(0, 1)$, sketch the image of the rectangle for $k = 3, -3$.

$$T(e_1) = T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e_1$$

$$T(e_2) = T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} k \\ 1 \end{pmatrix} = ke_1 + e_2$$

\therefore Standard matrix of $T = [T(e_1) \ T(e_2)]$

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

We find the images of the points under T .

$$T\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$T\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

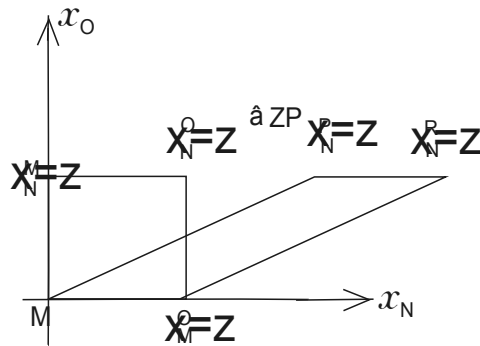
$$T\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2+k \\ 1 \end{pmatrix} = \begin{pmatrix} 2+k \\ 1 \end{pmatrix}$$

$$T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} k \\ 1 \end{pmatrix} = \begin{pmatrix} k \\ 1 \end{pmatrix}$$

when $k = 3$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

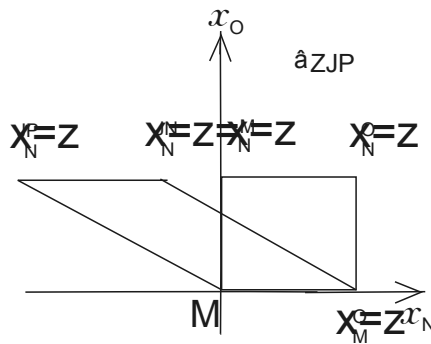
points on x_1 -axis remain unchanged.



when $k = -3$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

points on x_1 -axis are unchanged.



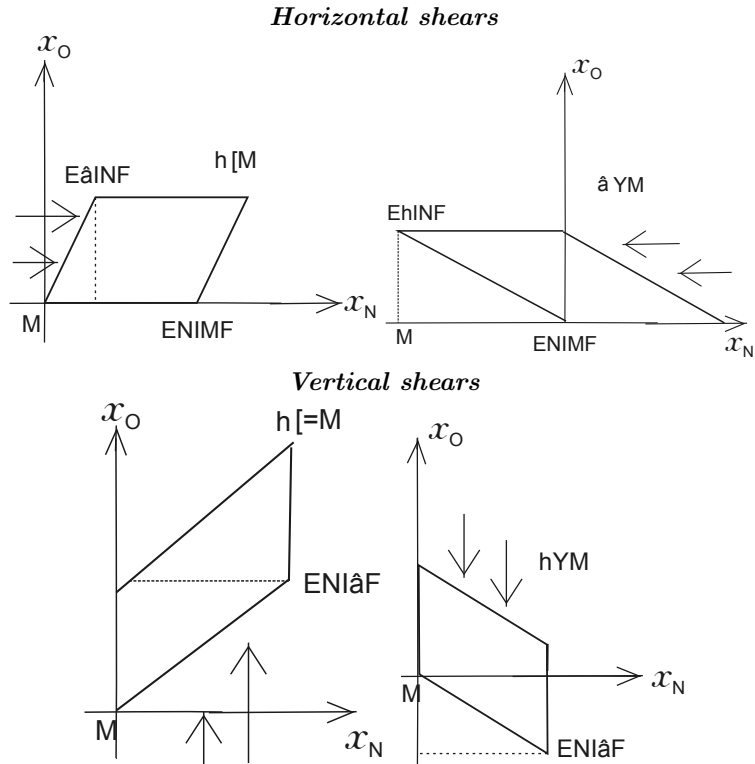
Example 12.16. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation, $T(e_1) = e_1 - 2e_2$, $T(e_2) = e_2$. What is the standard matrix of T ? What is this transformation called?

$$T(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The standard matrix of T is $[T(e_1) \ T(e_2)]$, i.e., $\begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix}$.

This transformation is called a vertical shear.

Example 12.17. We shall now sketch the image of the unit square under horizontal and vertical shears.

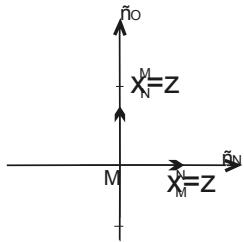
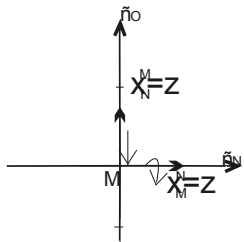
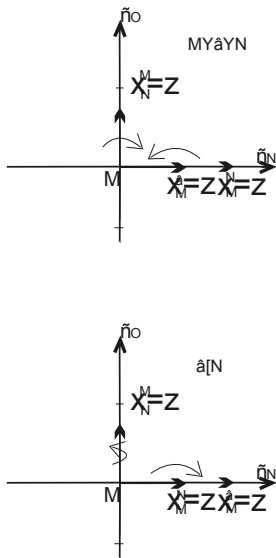


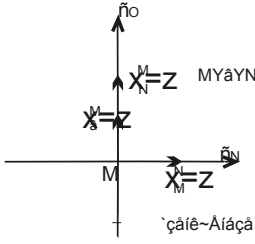
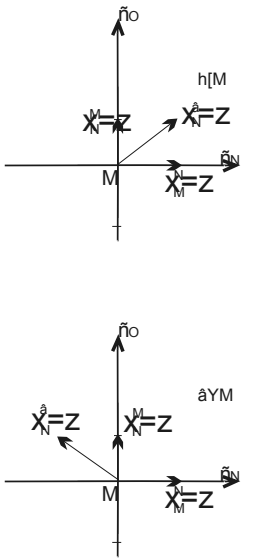
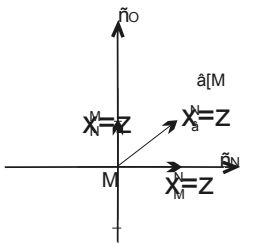
Matrices of Geometric Linear Transformation in \mathbb{R}^2

Given a geometric linear transformation on \mathbb{R}^2 , we see how to write the standard matrix of this transformation. For this, it is required to find the effect of the transformation on the columns e_1 and e_2 of I_2 .

Transformation	Geometrical effect on e_1 and e_2	Algebraic effect on e_1 and e_2	Standard matrix
Reflection through x_1 - axis		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Transformation	Geometrical effect on e_1 and e_2	Algebraic effect on e_1 and e_2	Standard matrix
Reflection through x_2 axis		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Reflection through $x_2 = x_1$		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Reflection through $x_2 = -x_1$		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} -1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
Reflection through origin		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

Transformation	Geometrical effect on e_1 and e_2	Algebraic effect on e_1 and e_2	Standard matrix
Projection onto x_1 axis		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
Projection onto x_2 axis		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
Horizontal scaling		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} k \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$

Transformation	Geometrical effect on e_1 and e_2	Algebraic effect on e_1 and e_2	Standard matrix
Vertical scaling		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ k \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$
Horizontal shear		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ k \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} k \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$
Vertical shear		$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ k \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$

Transformation	Geometrical effect on e_1 and e_2	Algebraic effect on e_1 and e_2	Standard matrix

Problem 12.9. A linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ first reflects points through the vertical axis (x_2 axis) and then reflects points through the line $x_2 = x_1$. Find the standard matrix of T . Also prove that T is a rotation about the origin. What is the angle of rotation? Plot the image of $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ under T .

Solution: Let T_1 be the reflection through x_2 -axis and T_2 reflection through the line $x_2 = x_1$.

Then $T = T_2 T_1$

The standard matrix of T is $[T(e_1) \ T(e_2)] = A$ (say).

First we see the effect of T_1, T_2 on e_1 and e_2 .

$$T_1(e_1) = -e_1, \quad T_1(e_2) = e_2$$

$$T_2(e_1) = e_2, \quad T_2(e_2) = e_1$$

$$\therefore T(e_1) = T_2 T_1(e_1) = T_2(T_1(e_1)) = T_2(-e_1) = -e_2$$

$$T(e_2) = T_2 T_1(e_2) = T_2(e_2) = e_1$$

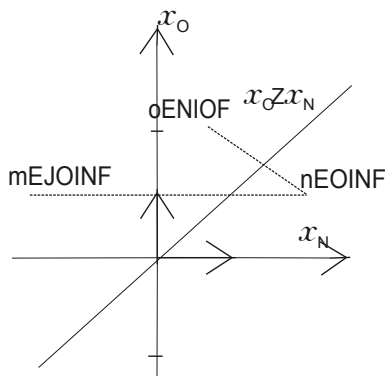
$$\therefore A = [-e_2 \ e_1] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

The matrix of rotation through an angle θ in the anticlockwise direction is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$\text{Now } A = \begin{pmatrix} \cos(-\pi/2) & -\sin(-\pi/2) \\ \sin(-\pi/2) & \cos(-\pi/2) \end{pmatrix}$$

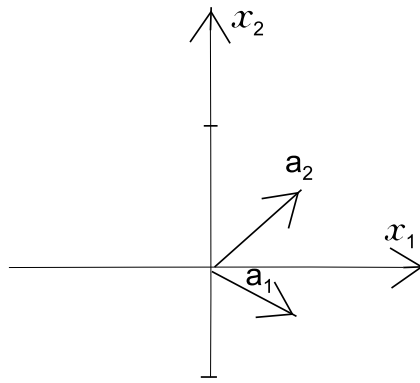
Thus T represents a rotation through $-\pi/2$ in the anticlockwise direction that is through an angle $\pi/2$ clockwise.



The point $P \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ lies in the second quadrant. The image of P under T_1 is $Q(2, 1)$ and under T_2 is $R(1, 2)$. Thus image of $P(-2, 1)$ under T_2T_1 i.e. T is $R(1, 2)$. This can also be obtained as

$$\begin{pmatrix} -2 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Problem 12.10. Let $A = [a_1 a_2]$ be the standard matrix of the linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, where a_1, a_2 are shown in the figure. Using the figure, draw the image of $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ under T .

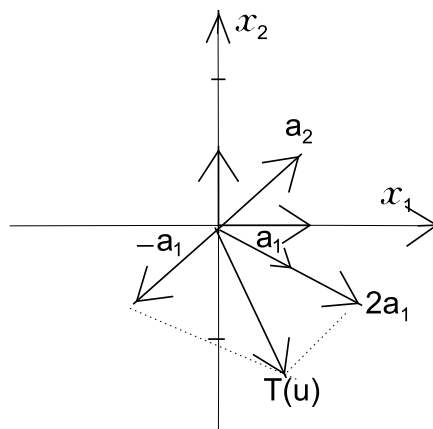


Solution: Since $A = [T(e_1) \ T(e_2)]$, therefore $a_1 = T(e_1)$ and $a_2 = T(e_2)$.

If $u = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, then $u = 2e_1 - e_2$.

$T(u) = T(2e_1 - e_2) = 2T(e_1) - T(e_2)$, T is linear.
 $= 2a_1 - a_2$.

How $T(u)$ is obtained is shown in the figure.



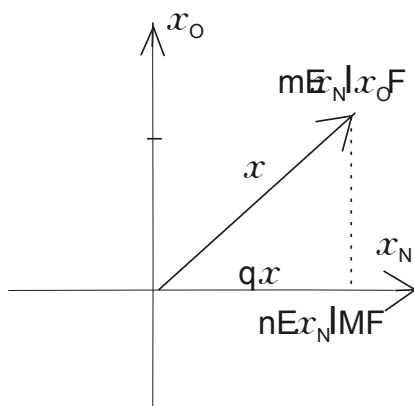
Geometrical Interpretation of Some Transformation

We now show that some matrix transformations can be interpreted geometrically.

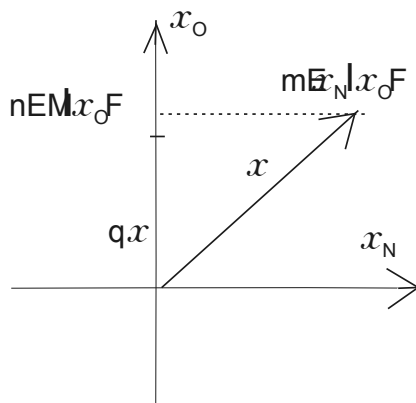
1. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, A determines a linear transformation of \mathbb{R}^2 which maps

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \text{ to } AX = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

This is the projection of X on the x_1 -axis.



2. The matrix transformation determined by $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ defines the projection of \mathbb{R}^2 on the x_2 -axis.



3. Let $A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$, where k is any fixed real number. Let T be the transformation determined by A .

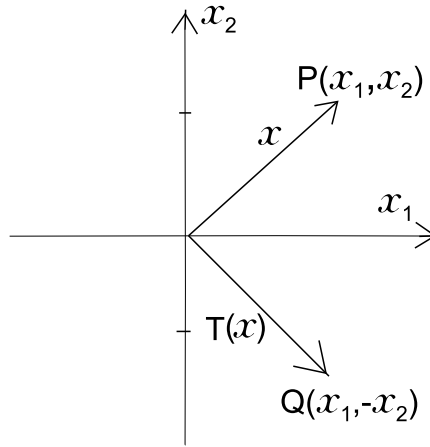
$$\text{If } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \text{ then } T(X) = AX = \begin{pmatrix} kx_1 \\ kx_2 \end{pmatrix} = k \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = kX.$$

Thus each vector X is mapped to kX . This is called scaling. If $k > 1$, T is a dilation.

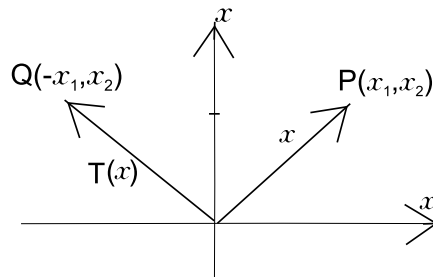
If $0 < k < 1$, T is a contraction.

4. The matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ defines a transformation T of \mathbb{R}^2 . If

$X = [x_1 \ x_2]^t$, $T(X) = AX = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$. T defines a reflection through the x_1 -axis.



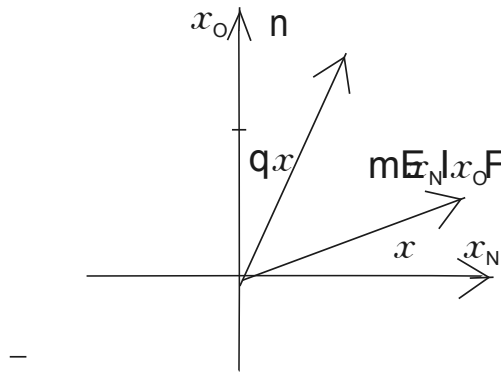
5. The matrix transformation determined by $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ defines a reflection in the x_2 -axis.



6. The matrix transformation T is determined by $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ maps the point $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to $T(X) = AX = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$. This gives a reflection in the line $y = x$.

7. The matrix transformation T is determined by $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ defines a reflection about the origin. The image of point $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is $\begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$.

8. The matrix transformation T is determined by $A = \begin{pmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{pmatrix}$ maps the point $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to $T(X) = AX = \begin{pmatrix} \frac{1}{\sqrt{2}}(x_1 - x_2) \\ \frac{1}{\sqrt{2}}(x_1 + x_2) \end{pmatrix}$. $T(X)$ can be obtained from X by rotating X about the origin through an angle of 45° anticlockwise.



In general, the matrix transformation T determined by $A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ defines a rotation about the origin through an angle θ in the anticlockwise direction.

12.11 Exercises

1. For the following linear transformations, find the standard matrix of T .

(i) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $T(e_1) = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$, $T(e_2) = \begin{pmatrix} -1 \\ 4 \\ 0 \end{pmatrix}$

(ii) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T(e_1) = e_1$, $T(e_2) = -3e_1 + e_2$. What is T called?

(iii) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ rotates points about the origin through $3\pi/4$ radians counterclockwise.

(iv) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ first rotate points through $-3\pi/4$ radians clockwise and then reflects points through the horizontal axis.

(v) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ first performs a horizontal shear that transforms e_2 to $2e_1 + e_2$ and then reflects points through the line $x_2 = -x_1$.

If T_1 is the transformation which performs the above two transformations in the reverse order, that is, first reflection then shearing. Is $T_1 = T$?

(vi) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ first reflects points through the vertical axis (x_2 -axis) and then rotates points through $\pi/2$ radians in anticlockwise direction.

2. If T_1, T_2 are two linear transformations, find the standard matrix of T_2T_1 (i.e. T_1 followed by T_2). Moreover, verify that standard matrix of $(T_2T_1) = (\text{standard matrix of } T_2)(\text{standard matrix of } T_1)$. Is $T_2T_1 = T_1T_2$?
- T_1 is reflection in the X_1 -axis. T_2 is rotation through an angle θ about 0.
 - T_i is rotation about the origin through an angle $\phi_i, i = 1, 2$
3. Let T be a linear transformation from \mathbb{R}^2 to \mathbb{R}^2 . Find the standard matrix of T when
- T rotates points about the origin through $5\pi/6$ radians counter-clockwise.
 - T rotates points about the origin through $\pi/3$ radians clockwise.
4. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation defined by $T(u) = Au$, where
- $$A = \begin{pmatrix} \cos\phi^\circ & -\sin\phi^\circ \\ \sin\phi^\circ & \cos\phi^\circ \end{pmatrix}, \phi = 30^\circ$$
- If $T_1(u) = A^2u$, describe the matrix action of T_1 on u as a matrix of rotation.
 - $T_2(u) = A^{-1}u$, describe the matrix action of T_2 on u as a matrix of rotation.
 - What is the smallest positive value of k for which $T_3(u) = A^k u = u$?
5. A shear in the x_2 direction is defined by $T(X) = T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ x_2 + kx_1 \end{pmatrix}$, where k is a scalar. If this shear is applied on a rectangle with vertices $(1, 1), (1, 4), (3, 1)$ and $(3, 4)$. Sketch the image of the rectangle for $k = 3$ and $k = -2$.
6. Let $O : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the zero linear transformation defined by $O(u) = 0$ for every $u \in \mathbb{R}^n$. Find the standard matrix of O .
7. Let $I : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the identity linear transformation defined by $I(u) = u$ for every $u \in \mathbb{R}^n$. Find the standard matrix of I .
8. Find the standard matrix representing a clockwise rotation of \mathbb{R}^2 about origin through $\pi/6$ radians.

12.12 Supplementary Problems

Problem 12.11. Define $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by

$$T(X) = AX + b$$

where A is a $m \times n$ matrix, $b \in \mathbb{R}^m$. Show that T is linear if and only if $b = 0$.

Solution: Let T be a linear transformation.

Then $T(0) = 0$ so that $A0 + b = 0$

$$\Rightarrow b = 0$$

Conversely, let $b = 0$ so that $T(X) = AX$

This is the matrix transformation which is linear.

Note: The transformation defined above is called an affine transformation. These transformations are important in computer graphics.

Problem 12.12. Show that the line through the vectors a and b in \mathbb{R}^n may be written in the parametric form $x = (1-t)a + tb$.

Solution: Let $A(a)$ and $B(b)$ be the two points and $P(x)$ any point on AB . Then
 $\rightarrow_{AP} = t \rightarrow_{AB}$ for some $t \in \mathbb{R}$
 $\therefore x - a = t(b - a)$, so that $x = (1-t)a + tb$

Problem 12.13. The line segment from a to b is the set of the form $(1-t)a + tb$ for $0 \leq t \leq 1$. Show that a linear transformation T maps this line segment onto a line segment or onto a single point.

Solution: The line segment $\rightarrow_{AP} = \{(1-t)a + tb / 0 \leq t \leq 1\}$. Since T is a linear transformation, therefore
 $T(\rightarrow_{AB}) = \{(1-t)T(a) + tT(b) / 0 \leq t \leq 1\}$ line segment joining $T(a)$ and $T(b)$.
 If $T = 0$ then it is a single point namely 0 .

Problem 12.14. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. Let $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ be such that $T(v_1), \dots, T(v_k)$ are known. Then T is completely determined on $\text{Span}\{v_1, v_2, \dots, v_k\}$. In particular if $\{v_1, v_2, \dots, v_k\}$ spans \mathbb{R}^n then T is completely determined.

Solution: Let $T(v_1) = u_1, T(v_2) = u_2, \dots, T(v_k) = u_k$. Then $u_1, u_2, \dots, u_k \in \mathbb{R}^m$.

If $X \in \text{span}\{v_1, v_2, \dots, v_k\}$ then there exist weights c_1, c_2, \dots, c_k such that $X = c_1v_1 + \dots + c_kv_k$.

$T(X) = T(c_1v_1 + \dots + c_kv_k) = c_1T(v_1) + \dots + c_kT(v_k)$, as T is a linear transformation.

$$= c_1u_1 + \dots + c_ku_k.$$

Hence $T(X)$ is known for all $X \in \text{Span}\{v_1, \dots, v_k\}$. If $\text{Span}\{v_1, \dots, v_k\} = \mathbb{R}^n$ then T is determined on \mathbb{R}^n .

Problem 12.15. Let P be the quadrilateral with vertices $A(1, 3), B(-2, 1), C(-1, -3)$ and $D(4, -1)$.

Find the coordinates of the vertices of the image of the quadrilateral P under the transformation T . Also sketch the quadrilateral P and its image under T .

(i) Horizontal shear T defined by

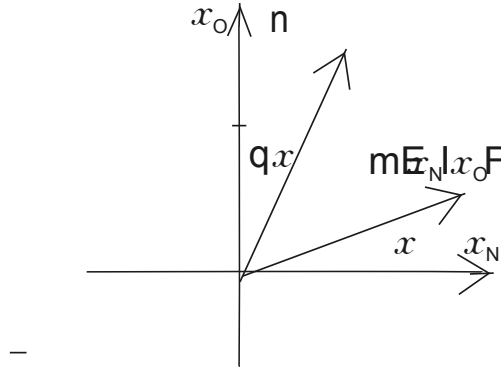
$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 - 3x_2 \\ x_2 \end{pmatrix}$$

(ii) T is clockwise rotation through $\pi/6$ radians.

Solution:

$$(i) T = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -8 & \\ & 3 \end{pmatrix}$$

Thus A is mapped to $A'(-8, 3)$ under T . Similarly B is mapped to $B'(-5, 1)$, C is mapped to $C'(8, -3)$ and D is mapped to $D'(7, -1)$ under T .



(ii) The matrix of T is $\begin{pmatrix} \cos(-\pi/6) & -\sin(-\pi/6) \\ \sin(-\pi/6) & \cos(-\pi/6) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$

$$\text{Therefore, } T(X) = \begin{pmatrix} \cos(\pi/6) & -\sin(-\pi/6) \\ \sin(-\pi/6) & \cos(-\pi/6) \end{pmatrix} X$$

$$= \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} X = \begin{pmatrix} 0.87 & 0.5 \\ -0.5 & 0.87 \end{pmatrix} X$$

$$T\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 0.87 & 0.5 \\ -0.5 & 0.87 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2.37 \\ 1.91 \end{pmatrix} \approx \begin{pmatrix} 2.4 \\ 1.9 \end{pmatrix}$$

Thus $A(1, 3)$ is mapped to $A'(2.4, 1.9)$ under T . Similarly $B(-2, -1)$ is mapped to $B'(-1.23, 1.87)$, $C(-1, -3)$ is mapped to $C'(-2.37, -1.91)$ and $D(4, -1)$ is mapped to $D'(2.96, -2.87)$ under T .

12.13 Supplementary Exercise

1. Indicate whether the following statements are true or false. Justify the false ones.

(i) If $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ is a mapping then \mathbb{R}^3 is the domain and \mathbb{R}^2 is the range of T .

(ii) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} |x_1| \\ x_2 \end{pmatrix} \text{ is not a linear transformation.}$$

(iii) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \text{ is onto but not injective.}$$

(iv) If A is a 5×3 matrix then $X \mapsto AX$ defines a mapping from \mathbb{R}^5 to \mathbb{R}^3 .

(v) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the matrix of reflection through x_2 -axis.

(vi) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is the matrix of rotation about origin through an angle $\pi/2$ radians clockwise.

- (vii) $\begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}$ is the matrix of dilation.
- (viii) A reflection through x_2 -axis followed by a rotation about origin through $\pi/2$ radians counter-clockwise is rotation about origin through $3\pi/2$ radians counter-clockwise.
- (ix) A reflection through x_1 -axis followed by a reflection through x_2 -axis is rotation about origin through π radian clockwise.
- (x) The composition of two linear transformations is a linear transformation.
- (xi) A linear transformation T from \mathbb{R}^2 to \mathbb{R}^3 can be onto.
- (xii) A linear transformation $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ can be injective.
- (xiii) Every linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is always onto and injective.
- (xiv) The standard matrix of a linear transformation from \mathbb{R}^2 to \mathbb{R}^2 that reflects points through x_1 -axis, or x_2 -axis or the origin has the form $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, where a, b can take values ± 1 .
- (xv) Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. For T to be injective, $T(X) = b$ must have a unique solution, for every $b \in \mathbb{R}^m$.
2. If T is a linear transformation from \mathbb{R}^5 to \mathbb{R}^6 defined by $T(X) = AX$ then what is the order of A ?
3. Consider the following transformation
 $T(x_1, x_2, x_3) = (x_1 + x_2 - x_3, x_2 + x_3 - x_1, x_3 + x_1 - x_2, x_1 + x_2 + x_3)$
 (i) Prove that T is linear.
 (ii) Find a matrix that implements T and hence prove that it is linear.
4. Let T be the matrix transformation defined by $T(X) = AX$. Find a vector X , if it exists, such that $T(X) = b$.

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 3 & -4 & 5 \\ 0 & 1 & 1 \\ -3 & 5 & 4 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 9 \\ 3 \\ -6 \end{pmatrix}$$
5. Is the vector $(1, 2, 3)$ a linear combination of the vectors $(1, -1, -3)$, $(1, 5, 1)$ and $(1, 2, -1)$?
6. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear transformation such that
 $T(1, -1, -3) = (2, 3, -4)$
 $T(1, 5, 1) = (4, 5, -6)$
 $T(1, 2, -1) = (6, 2, -3)$
 Find $T(1, 2, 3)$.
7. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ such that

$$T(v_1) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, T(v_2) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, T(v_3) = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\text{where } v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Determine T , that is, find $T(X)$ for any $X \in \mathbb{R}^3$.

8. If $X \in \mathbb{R}^3$, $X = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, express X in terms of v_1, v_2, v_3 .

In fact $X = cv_1 + (b-c)v_2 + (a-b)v_3$

$$T(X) = \begin{pmatrix} a \\ a \\ b \\ c \end{pmatrix}.$$

9. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. If $v_1, \dots, v_k \in \mathbb{R}^n$ such that $T(v_i) = 0$ for all $i = 1, \dots, k$ then $T(X) = 0$ for all $X \in \text{span}\{v_1, \dots, v_k\}$.
10. Let $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ be a linear transformation defined by $T(X) = AX$, where
- $$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$
- (i) Show that T is onto.
- (ii) Find a vector X such that $T(X) = b$, where $b = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}$
- (iii) Is the vector X obtained in (ii) unique? If not, obtain all X such that $T(X) = b$.
- (iv) What is the range of T ?
11. Given an echelon form of the standard matrix for a linear transformation T .
- (i) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, T is injective.
- (ii) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, T is onto and injective.
- (iii) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, T is not onto.
- (iv) $T : \mathbb{R}^4 \rightarrow \mathbb{R}^5$, T is injective.
- (v) $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3$, T is onto.
- (vi) $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$, T is neither injective nor onto.
12. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. What is the relationship which must hold between m and n for T to be
- (i) onto.
- (ii) injective.
- (iii) onto and injective.
13. Given an example of a linear transformation T , for the following:
- (i) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ such that T is not onto.
- (ii) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ such that T is not injective.
- (iii) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that T is neither onto nor injective.
14. Let T be a matrix transformation defined by $T(X) = AX$. Find all vectors X , if they exist, such that $T(X) = b$.

15. Let T be a matrix transformation defined by $T(X) = AX$. Find all vectors X which are mapped to zero.

16. Check whether the matrix transformation determined by the given matrix A is (a) onto, (b) injective.

$$(i) \quad A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & -1 & 3 \\ 3 & -5 & 4 \\ 1 & 17 & 4 \end{pmatrix} \quad \text{not injective, not onto.}$$

$$(ii) \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 7 & 13 & 9 \end{pmatrix} \quad 1-1, \text{ onto.}$$

$$(iii) \quad A = \begin{pmatrix} 1 & 3 & 4 & 7 \\ 2 & 4 & 5 & 8 \\ 3 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} \quad \text{not 1-1, onto.}$$

$$(iv) \quad A = \begin{pmatrix} 8 & 5 & 1 & 4 \\ 5 & 6 & 8 & 1 \\ 8 & 3 & 7 & 2 \end{pmatrix} \quad 1-1, \text{ onto.}$$

17. Let T be a linear operator on \mathbb{R}^n . Then T is onto \mathbb{R}^n if and only if T is injective.

Interpret the linear transformation determined by the matrix A geometrically.

$$(i) \quad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$(ii) \quad A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

18. Let P be the triangle with vertices $A(1, 1)$, $B(-2, -3)$ and $C(2, -1)$. Find the coordinates of the vertices of the image of the triangle P under the transformation T . Also sketch the triangle P and its image under T .

(i) Vertical shear T with shear factor 2. Also verify that the area of the triangle remains unchanged under T .

(ii) T is the counterclockwise rotation through $\pi/3$ radians.

19. Plot the image of the parabola $x_2 = x_1^2$ when it is rotated counterclockwise through $\pi/3$ radians.

20. A linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ first reflects points through the x_2 -axis and then reflects points through x_1 -axis. Show that T can also be described as a linear transformation that rotate points about the origin.

What is the angle of rotation? Plot the image of $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ under T .

21. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = ax + b.$$

Prove that f is a linear transformation if and only if $b = 0$.

12.14 Answers to Exercises

Exercise - 12.11

3.

$$(i) \begin{pmatrix} -\frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{pmatrix}$$

$$(ii) \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

4.

$$(i) A^2 = \begin{pmatrix} \cos 2\phi & -\sin 2\phi \\ \sin 2\phi & \cos 2\phi \end{pmatrix}. T_1(u) = A^2u \text{ is a rotation through } 2\phi.$$

(ii) T_2 is rotation through ϕ in the clockwise direction.

(iii) A^k is rotation through $k\phi$. Thus $k\phi = 360^\circ$ gives $k = 12$.

Supplementary Exercises

1.

(i) F. \mathbb{R}^3 is domain and \mathbb{R}^2 is co-domain.

(ii) T

(iii) F. Neither onto nor injective.

(iv) F. Defines a mapping from \mathbb{R}^3 to \mathbb{R}^5 .

(v) F. Reflection through x_1 -axis.

(vi) T

(vii) F. $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ is the matrix of dilation. or $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ is the matrix of dilation followed by reflection through $y = x$.

(viii) F. It is reflection about $y = -x$

(ix) T

(x) T

(xi) F. The matrix of T is a 3×2 which can have at most 2 pivots (one in each column). So each row can not have a pivot.

(xii) F. Matrix of T is 2×3 which cannot have a pivot.

(xiii) F. Zero transformation is neither onto nor injective.

(xiv) T

(xv) F. Either no solution or a unique solution.

$$4. \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}$$

10.

$$(ii) \quad X = \begin{pmatrix} -5 \\ 4 \\ 2 \\ 0 \end{pmatrix}$$

$$(iii) \quad \text{No. } X = \begin{pmatrix} -5+k \\ 4-k \\ 2-k \\ k \end{pmatrix}$$

$$(iv) \quad \text{Range } T = \mathbb{R}^3$$

11. One of the possible answers is given:

$$(i) \quad \begin{pmatrix} \blacksquare & \star \\ 0 & \blacksquare \\ 0 & 0 \end{pmatrix}$$

$$(ii) \quad \begin{pmatrix} \blacksquare & \star & \star \\ 0 & \blacksquare & \star \\ 0 & 0 & \blacksquare \end{pmatrix}$$

$$(iii) \quad \begin{pmatrix} 0 & \blacksquare & \star \\ 0 & 0 & 0 \end{pmatrix}$$

$$(iv) \quad \left. \begin{pmatrix} \blacksquare & \star & \star & \star \\ 0 & \blacksquare & \star & \star \\ 0 & 0 & \blacksquare & \star \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}$$

$$(v) \quad \left. \begin{pmatrix} \blacksquare & \star & \star & \star \\ 0 & \blacksquare & \star & \star \\ 0 & 0 & \blacksquare & \star \end{pmatrix} \right\}$$

$$(vi) \quad \begin{pmatrix} \blacksquare & \star & \star & \star \\ 0 & \blacksquare & \star & \star \\ 0 & 0 & \blacksquare & \star \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

12.

$$(i) \quad m \leq n$$

$$(ii) \quad m \geq n$$

$$(iii) \quad m = n$$

13. $T : X \rightarrow AX$, where A is

$$(i) \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(ii) \quad \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$(iii) \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

16.

$$(i) \quad X = \begin{pmatrix} -11 \\ -1 \\ 7 \end{pmatrix} k : TX = 0$$

$$(iii) \quad X = \begin{pmatrix} 1 \\ 3 \\ -6 \\ 2 \end{pmatrix} k, \quad k \in \mathbb{R}$$

17. One possible answer is:

- (i) Projection on x_2 -axis followed by reflection in the line $x_2 = x_1$.
- (ii) Projection on x_1 -axis followed by reflection in the line $x_2 = x_1$.

19. *Hint:* Take a few points (say 4 or 5) on the parabola, find their images. Join them.

This page is intentionally left blank.

UNIT - 5

Chapter 13

Vector Space

13.1 Definition and Examples

There are objects in mathematics which can be added together as well as multiplied by numbers also, like polynomials, matrices, real valued functions etc. In these systems, the operations of addition and multiplication by numbers have properties which are the same as those of \mathbb{R}^n as given in chapter. The elements of \mathbb{R}^n are called vectors and the real numbers are called scalars. For this reason, an algebraic structure which has \mathbb{R}^n , like properties is called a vector space. In this chapter, we define and study vector spaces. A good intuitive model for a vector space is provided by \mathbb{R}^2 and \mathbb{R}^3 .

Definition 13.1. (Vector space): Let V be a non-empty set and F a field. Let a binary operation $+$ be defined on V , and let scalar multiplication αv be defined for every $\alpha \in F$ and $v \in V$. Then, V is said to be a vector space over F if the following axioms hold:

- A1 $u + v \in V$ for all $u, v \in V$
- A2 $(u + v) + w = u + (v + w)$, for all $u, v, w \in V$
- A3 $u + v = v + u$, for all $u, v \in V$
- A4 There is a zero vector 0_v in V such that $u + 0_v = u$, for all $u \in V$
- A5 For each $u \in V$, there exists $v \in V$ such that $u + v = 0_v$
- M1 $\alpha u \in V$, for all $\alpha \in F, u \in V$
- M2 $\alpha(u + v) = \alpha u + \alpha v$, for all $\alpha \in F, u, v \in V$
- M3 $(\alpha + \beta)u = \alpha u + \beta u$, for all $\alpha, \beta \in F, u \in V$
- M4 $(\alpha\beta)(u) = \alpha(\beta u)$, for all $\alpha, \beta \in F, u \in V$
- M5 $1u = u$, for all $u \in V$

The elements of V are called vectors and the elements of F are called scalars. A vector space is also called a linear space.

Remark 13.1. Properties A1 to A5 are equivalent to saying that $(V, +)$ is an Abelian group. Consequently, identity element and additive inverse of an every element is unique. The additive identity element e of V is called the null vector (or zero vector) and is denoted by 0_v . The additive inverse of an element $v \in V$ is denoted by $-v$.

Remark 13.2. The zero of the field F is also denoted by the same symbol “0”. It will clear from the context when “0” denotes the zero vector or the zero (scalar) of the field F .

Notation

The vector space V over a field F , is denoted by $V(F)$ or V_F . If F is understood from the context, then we simply say that V is a vector space.

Elementary Properties

The following theorem gives us some very simple, yet useful properties of a vector space.

Theorem 13.1. Let V be a vector space over a field F . Then the following properties hold.

- (i) Scalar multiplication with the zero vector gives the zero vector, i.e. $\alpha 0 = 0$, for all $\alpha \in F$
- (ii) Multiplication by the zero scalar yields the zero vector, i.e. $0v = 0$, for all $v \in V$
- (iii) If $\alpha \in F$ and $v \in V$ such that $\alpha v = 0$, then either $\alpha = 0$ or $v = 0$
- (iv) $(-\alpha)v = \alpha(-v) = -\alpha v$, for all $\alpha \in F$, $v \in V$
- (v) $(-1)v = -v$, for all $v \in V$
- (vi) For $0 \neq v \in V$, $\alpha, \beta \in F$ such that $\alpha \neq \beta$, then $\alpha v \neq \beta v$

Proof: (i) Let $\alpha \in F$

$$\begin{aligned} \text{Then } \alpha 0 &= \alpha(0 + 0) \\ \Rightarrow \alpha 0 &= \alpha 0 + \alpha 0 \quad \text{by M2} \\ \Rightarrow \alpha 0 + 0 &= \alpha 0 + \alpha 0 \quad \text{by A4} \\ \Rightarrow 0 &= \alpha 0 \quad \text{by cancellation property for addition in } V \\ \therefore \alpha 0 &= 0 \end{aligned}$$

(ii) Let $v \in V$. Then

$$\begin{aligned} 0v &= (0 + 0)v \\ \Rightarrow 0v &= 0v + 0v \quad \text{using M3} \\ \Rightarrow 0 + 0v &= 0v + 0v \quad \text{using A4} \\ \Rightarrow 0 &= 0v \quad \text{by cancellation property of addition in } V \end{aligned}$$

Hence $0v = 0$ for all $v \in V$

(iii) Let $\alpha \in F$, $v \in V$ such that

$$\alpha v = 0$$

If $\alpha = 0$ then proof is complete.

If $\alpha \neq 0$ then $\alpha^{-1} \in F$, so that

$$\begin{aligned} \alpha^{-1}(\alpha v) &= \alpha^{-1}0 \\ \Rightarrow (\alpha^{-1}\alpha)v &= 0 \quad \text{using M4 and (i)} \\ \Rightarrow 1v &= 0 \\ \Rightarrow v &= 0 \quad \text{using M5} \end{aligned}$$

Hence proved.

(iv) Let $\alpha \in F$, $v \in V$

$$\begin{aligned} (-\alpha)v + \alpha v &= (-\alpha + \alpha)v, \quad \text{using M3} \\ &= 0v = 0 \quad \text{using (ii)} \end{aligned}$$

$\therefore (-\alpha)v = -\alpha v$, similarly $\alpha(-v) = -\alpha v$

(v) In (iv) take $\alpha = 1$

(vi) Let $v \in V, \alpha, \beta \in F$ and $\alpha \neq \beta$. Suppose that $\alpha v = \beta v$. Then
 $\alpha v - \beta v = 0 \Rightarrow (\alpha - \beta)v = 0$ since $v \neq 0 \Rightarrow \alpha - \beta = 0$ using (iii)
 $\Rightarrow \alpha = \beta$, which contradicts our hypothesis that $\alpha \neq \beta$.
Hence our assumption is wrong, so that $\alpha v \neq \beta v$. \square

Remark 13.3. In this chapter we shall study vector spaces over \mathbb{R} and \mathbb{C} only. V will always be taken as a vector spaces over \mathbb{R} unless specified otherwise, in which case it will be over \mathbb{C} .

Notation

An element of \mathbb{R}^n has been written as $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $x_i \in \mathbb{R}$. Interchangeably, we

write it as (x_1, x_2, \dots, x_n) .

If $V = \{0\}$ and we define

$$0 + 0 = 0$$

$$\alpha 0 = 0, \forall \alpha \in \mathbb{R}$$

Then V is called a vector space over \mathbb{R} . It is called the zero vector space.

If a vector space is such that $V \neq \{0\}$, then it is called a non-zero vector space.

If V is a non-zero vector space over the field \mathbb{R} , then can V have only a finite number of elements?

The answer is in the negative, \because if $0 \neq v \in V$ and $\alpha, \beta \in \mathbb{R}, \alpha \neq \beta$ then $\alpha v \neq \beta v$.

Thus if $S = \{\alpha v \mid \alpha \in \mathbb{R}\}$, then S is infinite and $S \subseteq V$, so that S has infinitely many elements.

In fact, if $V \neq \{0\}$ is a vector space over a field F , then V has infinite number of elements if F is infinite.

Example 13.1. The set of all real numbers \mathbb{R} is a vector space over itself. The addition and scalar multiplication is defined as follows: (1) $x_1 + x_2$ (2) αx for every $x_1, x_2, x \in \mathbb{R}$ treated as vectors and for every $\alpha \in \mathbb{R}$ treated as a scalar.

Example 13.2. \mathbb{C} with respect to the usual addition and multiplication is a vector space over the field \mathbb{C} . The zero vector is the complex number 0 and the negative of the vector x is the complex number $-x$.

Example 13.3. \mathbb{C} is a vector space over \mathbb{R} , with respect to the operations defined as follows:

If $v_1, v_2 \in \mathbb{C}, \alpha \in \mathbb{R}$, and $v_1 = a_1 + ib_1, v_2 = a_2 + ib_2$, then

$$v_1 + v_2 = (a_1 + a_2) + i(b_1 + b_2), \alpha v_1 = (\alpha a_1) + i(\alpha b_1)$$

Example 13.4. \mathbb{R} with usual addition and multiplication forms a vector space over the field \mathbb{Q} of rational numbers.

Example 13.5. \mathbb{R}^2 is a vector space over \mathbb{R} with addition and scalar multiplication defined as follows:

For $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2, \alpha \in \mathbb{R}$,

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \alpha(x_1, y_1) = (\alpha x_1, \alpha y_1)$$

Example 13.6. Let $n \geq 1$ be a fixed integer. \mathbb{R}^n , the set of all n tuples of real numbers, is a vector space over \mathbb{R} with component-wise addition and scalar multiplication.

Example 13.7. Let $n \geq 1$ be a fixed integer. Then \mathbb{C}^n is a vector space over \mathbb{R} , with component-wise addition and scalar multiplication.

Example 13.8. Let $V = \{(x_1, x_2, 0) | x_1, x_2 \in \mathbb{R}\}$. Then V is a vector space over \mathbb{R} with respect to component-wise addition and scalar multiplication in V . In fact, V represents x_1x_2 plane.

Example 13.9. Let $V = \mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} | a, b \in \mathbb{Q}\}$. Then V is a vector space over \mathbb{Q} with usual addition of real numbers and scalar multiplication.

Example 13.10. $M_2(\mathbb{R})$ is a vector space over \mathbb{R} with respect to matrix addition and scalar multiplication of matrices.

Example 13.11. $M_3(\mathbb{C})$ is a vector space over \mathbb{C} with respect to matrix addition and scalar multiplication of matrices.

Example 13.12. Let P_2 be the set of all polynomials of degree ≤ 2 over \mathbb{R} . For $f = a_0 + a_1x + a_2x^2$, $g = b_0 + b_1x + b_2x^2 \in P_2$, where $a_i, b_i \in \mathbb{R}$, $i = 0, 1, 2$: Define $f+g = (a_0+b_0) + (a_1+b_1)x + (a_2+b_2)x^2$ and $\alpha f = (\alpha a_0) + \alpha a_1x + \alpha a_2x^2$. It is easy to check that P_2 is a vector space over \mathbb{R} .

V ceases to be a vector space if (i) even one of the axioms fails to hold, or (ii) any one of the properties of a vector space fails to hold. Note that whether a given set V is a vector space over a field or not depends very much on the operations defined on V , i.e. vector addition and the scalar multiplication. This is shown by the following examples.

Example 13.13. Consider $V = \mathbb{R}^3$, define
 $(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$
 $\alpha(x_1, x_2, x_3) = (\alpha x_1, x_2, x_3)$ for all $\alpha \in \mathbb{R}$
 Then for $x, y \in V$, $\alpha \in \mathbb{R}$, $x + y \in V$, $\alpha x \in V$
 So the composition are well defined. Let $(1, 1, 1) \in V$.
 Then $0(1, 1, 1) = (0, 1, 1)$, so that $0v = 0$ is not satisfied for all $v \in V$
 Hence V is not a vector space over \mathbb{R} .

Example 13.14. Let $V = \mathbb{R}^2$. Define addition and scalar multiplication in V as follows:

For $x, y \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$, $x = (x_1, x_2)$, $y = (y_1, y_2)$
 $x + y = (0, x_2 + y_2)$
 $\alpha x = (\alpha x_1, \alpha x_2)$

Suppose $(e_1, e_2) \in V$ is the zero vector of V . Then for any $(x_1, x_2) \in V$

$$(x_1, x_2) + (e_1, e_2) = (x_1, x_2)$$

$$\Rightarrow (0, x_2 + e_2) = (x_1, x_2)$$

which is not possible when $x_1 \neq 0$. Hence V does not have a zero element, so that V is not a vector space over \mathbb{R} under the given operations.

Example 13.15. Let $V = \mathbb{R}^2$. Define

$$(x_1, x_2) + (y_1, y_2) = (x_1 + x_2, y_1 + y_2)$$

$$\alpha(x_1, x_2) = (\alpha x_1, 0)$$

Observe that $v = (2, 3) \in V$ is such that $1v \neq v$ Therefore V is not a vector space over \mathbb{R} .

Problem 13.1. Verify that \mathbb{R}^n is a vector space over \mathbb{R} , with component-wise addition and scalar multiplication.

Solution:

1. Let $u, v \in \mathbb{R}^n$, so that

$$u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n) \text{ for } u_i, v_i \in \mathbb{R}, 1 \leq i \leq n.$$

$$\text{Then } u + v = (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \in \mathbb{R}^n$$

$$\because u_i + v_i \in \mathbb{R}, 1 \leq i \leq n. \text{ Thus } u + v \in \mathbb{R}^n$$

2. Let $u, v, w \in \mathbb{R}^n$. Then $u = (u_1, u_2, \dots, u_n)$, $v = (v_1, v_2, \dots, v_n)$, $w = (w_1, w_2, \dots, w_n)$, where $u_i, v_i, w_i \in \mathbb{R}$, $1 \leq i \leq n$.

Then

$$\begin{aligned} (u + v) + w &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) + (w_1, w_2, \dots, w_n) \\ &= ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \dots, (u_n + v_n) + w_n) \\ &= (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \dots, u_n + (v_n + w_n)) \\ &\quad \text{(using Associative law for addition in } \mathbb{R} \text{)} \\ &= (u_1, u_2, \dots, u_n) + (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n) \\ &= u + ((v_1, v_2, \dots, v_n) + (w_1, w_2, \dots, w_n)) \\ &= u + (v + w) \end{aligned}$$

$$\text{Hence } (u + v) + w = u + (v + w), \forall u, v, w \in \mathbb{R}^n.$$

3. Let $u, v \in \mathbb{R}^n$. Then

$$u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n) \text{ where } u_i, v_i \in \mathbb{R}, 1 \leq i \leq n$$

$$\begin{aligned} \therefore u + v &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \\ &= (v_1 + u_1, v_2 + u_2, \dots, v_n + u_n) \text{ (using Commutative law} \\ &\quad \text{for addition in } \mathbb{R} \text{)} \\ &= (v_1, v_2, \dots, v_n) + (u_1, u_2, \dots, u_n) \\ &= v + u \end{aligned}$$

$$\text{Hence } u + v = v + u, \forall u, v \in \mathbb{R}^n$$

4. $0 = (0, 0, \dots, 0) \in \mathbb{R}^n$ such that

$$u + 0 = u, \forall u \in \mathbb{R}^n$$

5. Let $u \in \mathbb{R}^n$. Then $u = (u_1, u_2, \dots, u_n)$, $u_i \in \mathbb{R}$.

$$\text{Let } v = (-u_1, -u_2, \dots, -u_n). \text{ Then } v \in \mathbb{R}^n \text{ and } u + v = 0.$$

Thus v is the negative of u .

6. Let $u = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n, \alpha \in \mathbb{R}$. Then $\alpha u = (\alpha u_1, \alpha u_2, \dots, \alpha u_n) \in \mathbb{R}^n$.

7. Let $u, v \in \mathbb{R}^n, \alpha \in \mathbb{R}$

$$\text{Then } u = (u_1, u_2, \dots, u_n), v = (v_1, v_2, \dots, v_n)$$

$$\begin{aligned}
\alpha(u+v) &= \alpha(u_1+v_1, u_2+v_2, \dots, u_n+v_n) \\
&= (\alpha(u_1+v_1), \alpha(u_2+v_2), \dots, \alpha(u_n+v_n)) \\
&= (\alpha u_1 + \alpha v_1, \alpha u_2 + \alpha v_2, \dots, \alpha u_n + \alpha v_n) \\
&= (\alpha u_1, \alpha u_2, \dots, \alpha u_n) + (\alpha v_1, \alpha v_2, \dots, \alpha v_n) \\
&= \alpha u + \alpha v
\end{aligned}$$

Hence $\alpha(u+v) = \alpha u + \alpha v$, $\forall u, v \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$.

8. Let $u = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$

$$\begin{aligned}
\text{Then } (\alpha + \beta)u &= ((\alpha + \beta)u_1, (\alpha + \beta)u_2, \dots, (\alpha + \beta)u_n) \\
&= (\alpha u_1 + \beta u_1, \alpha u_2 + \beta u_2, \dots, \alpha u_n + \beta u_n) \\
&= (\alpha u_1, \alpha u_2, \dots, \alpha u_n) + (\beta u_1, \beta u_2, \dots, \beta u_n) \\
&= \alpha(u_1, u_2, \dots, u_n) + \beta(u_1, u_2, \dots, u_n) \\
&= \alpha u + \beta u
\end{aligned}$$

$\therefore (\alpha + \beta)u = \alpha u + \beta u$, $\forall u \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$.

9. Let $u = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$. Then

$$\begin{aligned}
(\alpha\beta)u &= ((\alpha\beta)u_1, (\alpha\beta)u_2, \dots, (\alpha\beta)u_n) \\
&= (\alpha(\beta u_1), \alpha(\beta u_2), \dots, \alpha(\beta u_n)) \quad (\text{using Associative law for} \\
&\hspace{15em} \text{multiplication in } \mathbb{R}) \\
&= \alpha(\beta u_1, \beta u_2, \dots, \beta u_n) \\
&= \alpha(\beta(u_1, u_2, \dots, u_n)) \\
&= \alpha(\beta u)
\end{aligned}$$

$\therefore (\alpha\beta)u = \alpha(\beta u)$, $\forall u \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$.

10. Let $u = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$. Then $1 \in \mathbb{R}$ and

$$\begin{aligned}
1u &= (1u_1, 1u_2, \dots, 1u_n) \\
&= (u_1, u_2, \dots, u_n) \\
&= u
\end{aligned}$$

Thus, it follows from the above that \mathbb{R}^n is a vector space over \mathbb{R} .

Problem 13.2. Let S be a non-empty set and V be the set of all functions from S to \mathbb{R} . Define the sum of two functions f and g to be the function $(f+g)$ by the rule

$$(f+g)(x) = f(x) + g(x), \quad \forall x \in S$$

Also, for $f \in S$, $\alpha \in \mathbb{R}$, define the scalar multiplication of f by α , by the rule

$$(\alpha f)(x) = \alpha f(x), \quad \forall x \in S$$

Verify that V is a vector space over \mathbb{R} with respect to addition and scalar multiplication defined above.

Solution:

1. If $f, g \in V$, then $f + g$, as defined above, is also a function from S to \mathbb{R} so that $f + g \in V$.

2. Let $f, g, h \in V$. For $x \in S$

$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) \\ &= (f(x) + g(x)) + h(x) \\ &= f(x) + (g(x) + h(x)) \quad (\text{using Associative law for addition in } \mathbb{R}) \\ &= f(x) + (g + h)(x) \\ &= (f + (g + h))(x) \end{aligned}$$

$$\therefore (f + g) + h = f + (g + h)$$

3. Let $f, g \in V$. For $x \in S$

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ &= g(x) + f(x) \\ &= (g + f)(x) \end{aligned}$$

Hence $f + g = g + f$.

4. If e is the zero function on S , defined by $e(x) = 0 \forall x \in S$, then $e \in V$. Also for any $f \in V$, and $x \in S$

$$\begin{aligned} (f + e)(x) &= f(x) + e(x) \\ &= f(x) + 0 \\ &= f(x) \end{aligned}$$

$$\therefore f + e = f$$

Thus e is the zero element of V .

5. Let $f \in V$. Define the function g on S by $g(x) = -f(x)$, $\forall x \in S$ then $g \in V$. For each $x \in S$

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ &= f(x) - f(x) \\ &= 0 \\ &= e(x) \end{aligned}$$

$$\therefore f + g = e$$

Thus g is the inverse of f in V .

6. If $f \in V$ and $\alpha \in \mathbb{R}$, then αf , is a well defined function on S , so that $\alpha f \in V$.

7. Let $f \in V$, $\alpha, \beta \in \mathbb{R}$. For each $x \in S$

$$\begin{aligned} ((\alpha\beta)f)(x) &= (\alpha\beta)f(x) \\ &= \alpha(\beta f(x)) \\ &= \alpha((\beta f)(x)) \\ &= (\alpha(\beta f))(x) \end{aligned}$$

$$\therefore (\alpha\beta)f = \alpha(\beta f)$$

8. Let $f \in V$, $\alpha, \beta \in \mathbb{R}$. For each $x \in S$

$$\begin{aligned} ((\alpha + \beta)(f))(x) &= (\alpha + \beta)f(x) \\ &= \alpha f(x) + \beta f(x) \quad (\text{using Distributive law in } \mathbb{R}) \\ &= (\alpha f)(x) + (\beta f)(x) \\ &= (\alpha f + \beta f)(x) \end{aligned}$$

$$\therefore (\alpha + \beta)f = \alpha f + \beta f$$

9. Let $f, g \in V$, $\alpha \in \mathbb{R}$. For each $x \in S$

$$\begin{aligned} (\alpha(f + g))(x) &= \alpha((f + g)(x)) \\ &= \alpha(f(x) + g(x)) \\ &= \alpha f(x) + \alpha g(x) \\ &= (\alpha f)(x) + (\alpha g)(x) \end{aligned}$$

$$\therefore \alpha(f + g) = \alpha f + \alpha g$$

10. Let $f \in V$, $1 \in \mathbb{R}$ such that

$$\begin{aligned} (1f)(x) &= 1f(x) \\ &= f(x) \end{aligned}$$

$$\therefore 1f = f$$

Thus V is a vector space over \mathbb{R} .

Problem 13.3. Let $V = \mathbb{R}^+$, the set of all positive real numbers. Define addition and scalar multiplication on V as follows:

For $x, y \in V$, $\alpha \in \mathbb{R}$, define

$$\begin{aligned} x + y &= xy \\ \alpha x &= x^\alpha. \end{aligned}$$

Show that V is a vector space over \mathbb{R} .

Solution: Clearly sum is well defined. Also if $x \in V, \alpha \in \mathbb{R}$, then $x^\alpha \geq 0$ so that $x^\alpha \in V$. The zero vector is 1, and the negative of $x \in V$ is $\frac{1}{x} \in V$. If $\alpha, \beta \in V$, then

$$\begin{aligned} (\alpha + \beta)x &= x^{\alpha+\beta} \\ &= x^\alpha x^\beta \\ &= x^\alpha + x^\beta \\ &= \alpha x + \beta x \end{aligned}$$

$$\therefore (\alpha + \beta)x = \alpha x + \beta x$$

Similarly the other axioms can be verified, so that V is a vector space over \mathbb{R} .

Problem 13.4. Let X be a non-empty set and V be the power set of X . Show that V is a vector space over the field $F = (\mathbb{Z}_2, \oplus_2, \odot_2)$ where $\mathbb{Z}_2 = \{0, 1\}$, with respect to the operations defined as follows:

For $A, B \in V$, $\alpha \in F$,

$$A + B = A \Delta B$$

$$\alpha A = \begin{cases} A, & \text{if } \alpha = 1; \\ \phi, & \text{if } \alpha = 0. \end{cases}$$

Solution: It has been proved in Example 5.26 that $(V, +)$ is a group. Clearly $\alpha A \in V$, $\forall A \in V$, $\alpha \in F$.

(i) Let $A, B \in V$, $\alpha \in F$. Then

$$\begin{aligned} \alpha(A + B) &= \alpha(A \Delta B) \\ &= \begin{cases} A \Delta B, & \text{if } \alpha = 1; \\ \phi, & \text{if } \alpha = 0. \end{cases} \end{aligned}$$

If $\alpha = 1$ then $\alpha A + \alpha B = A + B = A \Delta B$

If $\alpha = 0$ then $\alpha A + \alpha B = \phi + \phi = \phi \Delta \phi = \phi$

$$\therefore \alpha A + \alpha B = \begin{cases} A \Delta B, & \text{if } \alpha = 1; \\ \phi, & \text{if } \alpha = 0. \end{cases}$$

Hence $\alpha(A + B) = \alpha A + \alpha B$

(ii) Let $A, B \in V$, $\alpha \in F$

$$\text{Then } \alpha + \beta = \begin{cases} 1, & \text{if either } \alpha = 1, \text{ or } \beta = 1, \text{ but not both;} \\ 0, & \text{otherwise.} \end{cases}$$

There are four cases which are shown below.

α	β	$(\alpha + \beta)A$	$\alpha A + \beta A$
0	0	$0A = \phi$	$0A + 0B = \phi \Delta \phi = \phi$
1	0	$1A = A$	$1A + 0A = A + \phi = A \Delta \phi = A$
0	1	$1A = A$	$0A + 1A = \phi + A = \phi \Delta A = A$
1	1	$0A = \phi$	$1A + 1A = A + A = A \Delta A = \phi$

Hence $(\alpha + \beta)A = \alpha A + \beta A$.

(iii) Let $A \in V$, $\alpha, \beta \in F$

The different cases are shown below:

α	β	$(\alpha\beta)A$	βA	$\alpha(\beta A)$
0	0	$0A = \phi$	$0A = \phi$	$0\phi = \phi$
1	0	$0A = \phi$	$0A = \phi$	$1\phi = \phi$
0	1	$0A = \phi$	$1A = A$	$0A = \phi$
1	1	$1A = A$	$1A = A$	$1A = A$

Thus in all cases

$$(\alpha\beta)A = \alpha(\beta A)$$

(iv) Let $A \in V$. Then $1A = A$ by definition of scalar multiplication.

Thus $1A = A$, $\forall A \in V$

Hence V is a vector space over F .

13.2 Exercise

1. Verify that the vector spaces in Example 13.1 and Example 13.12 satisfy the axioms of a vector space.
2. Show that $V = M_{m \times n}(\mathbb{R})$ is a vector space over \mathbb{R} under matrix addition and scalar multiplication of matrices.
3. If A is a fixed $m \times n$ matrix over \mathbb{R} then prove that the set of all solutions of $AX = 0$ in \mathbb{R}^n is a vector space over \mathbb{R} , under the usual addition and scalar multiplication of matrices.
4. Let $V = \mathbb{R}^+$. Is V a vector space over \mathbb{R} with respect to usual addition in V and scalar multiplication defined by $\alpha x = x^\alpha$, for $x \in V, \alpha \in \mathbb{R}$.
5. Are the following true or false? Justify your answer.
 - (i) \mathbb{Q} is a vector space over \mathbb{R} .
 - (ii) \mathbb{R} is a vector space over \mathbb{C} .
 - (iii) \mathbb{Q} is a vector space over \mathbb{Q} .
 - (iv) \mathbb{R} is a vector space over \mathbb{Q} .
 - (v) \mathbb{C} is a vector space over \mathbb{Q} .
6. In each of the following justify why V is not a vector space over \mathbb{R} .
 - (i) $V = \mathbb{R}^3$ with the operations defined as follows: For $v_1, v_2 \in V, \alpha \in \mathbb{R}$
 If $v_1 = (x_1, y_1, z_1), v_2 = (x_2, y_2, z_2)$
 $v_1 + v_2 = (x_2, y_1 + y_2, z_2)$
 $\alpha v_1 = (\alpha x_1, \alpha y_1, \alpha z_1)$
 - (ii) $V = \mathbb{R}^3$, usual componentwise addition
 For $v = (x, y, z) \in \mathbb{R}^3, \alpha \in \mathbb{R}$
 $\alpha v = (x, 1, z) \cdots \cdots$
 - (iii) $V = \mathbb{R}^2, v_1, v_2 \in V, \alpha \in \mathbb{R}$
 $v_1 = (x_1, y_1), v_2 = (x_2, y_2)$
 $v_1 + v_2 = (x_1 + x_2 + 1, y_1 + y_2)$
 $\alpha(x_1, y_1) = (\alpha x_1, \alpha y_1)$
 - (iv) $V = \mathbb{R}^+$, with vector addition and scalar multiplication defined by $\alpha x = e^{\alpha x}$, for $\alpha \in \mathbb{R}, x \in V$
 - (v) $V = \mathbb{R}^+$, with usual addition and scalar multiplication defined as
 For $x \in V, \alpha \in \mathbb{R}, \alpha x = |\alpha|x$.
7. Let V be a vector space over \mathbb{C} . Define another system over \mathbb{C} as follows: $V_1 = V$ and addition in V_1 is same as in V , scalar multiplication in V_1 is defined as:
 For $v \in V, \alpha \in \mathbb{C}, \alpha V = (\operatorname{Re} \alpha)v$. Is V_1 a vector space over \mathbb{C} ?
8. Let V be a vector space over \mathbb{R} . If $v \in V$ and $\alpha \in \mathbb{R}$ such that $\alpha v = v$ then show that either $\alpha = 1$ or $v = 0$.
9. Let $V = \mathbb{Q}[\sqrt{5}] = \{a + b\sqrt{5} | a, b \in \mathbb{Q}\}$. Prove that V is a vector space over \mathbb{Q} with respect to usual addition of real numbers, and scalar multiplication defined by usual multiplication.

10. Show that the set of all real valued functions on $[0, 1]$ is a vector space over \mathbb{R} with respect to pointwise addition and scalar multiplication of functions.
11. Show that the set of all real valued continuous functions on $[0, 1]$ is a vector space over \mathbb{R} with respect to pointwise addition and scalar multiplication of functions.
12. Let S denote the set of all real valued functions defined on \mathbb{Z} . Prove that S is a vector space over \mathbb{R} under the operations defined by the following:
 For $f, g \in S, \alpha \in \mathbb{R}$
 $(f + g)(x) = f(x) + g(x), \forall x \in \mathbb{Z}$
 $(\alpha f)(x) = \alpha f(x), \forall x \in \mathbb{Z}$
 (such functions are called discrete or digital signals).
13. Let $V = \mathbb{R}$. For any $u, v \in V, \alpha \in \mathbb{R}$, define
 $u \oplus v = u + v + 1$
 $\alpha \odot u = \alpha u + \alpha - 1$
 Prove that V is a vector space over \mathbb{R} under these two operations.

13.3 Subspaces

Similar to the concept of subgroup and subrings, here also, subspaces mean vector spaces within vector spaces. If V is a vector space over a field F and W is a non-empty subset of V such that W is a vector space over F in its own right then W is called a subspace of V over F . Thus, we have the following definition

Definition 13.2. (Subspace): Let V be a vector space over a field F and W be a non-empty subset of V . Then W is called a subspace of V if W is a vector space over F , with respect to the operations of addition and scalar multiplication in V restricted to W .

For W to be a subspace of V , it is sufficient to check only 3 conditions. The rest are satisfied automatically. In this regard, we have the following theorem.

Theorem 13.2. Let V be a vector space over a field F and W a subset of V . Then W is a subspace of V if and only if

- (i) $0 \in W$
 (ii) $w_1 + w_2 \in W$ for all $w_1, w_2 \in W$
 (iii) $\alpha w \in W$ for all $\alpha \in F, w \in W$.

Proof: The conditions are necessary

Suppose W is a subspace of V , then W is a vector space over F , so that $0 \in W$. Hence (i) holds. The conditions (ii) and (iii) must be satisfied since a vector space is closed under vector addition and scalar multiplication.

The conditions are sufficient

Let the conditions hold.

Since $W \subseteq V$, \therefore the properties $A2$ to $A4$ and $M2$ to $M5$ hold in W as they hold in V .

We only need to show the existence of additive inverse in W . For $w_1, w_2 \in W$ and $\alpha \in F, w_1 + w_2, \alpha w_1 \in W$. \therefore For $\alpha = -1, w_1 \in W$ we get $(-1)w_1 \in$

$W \Rightarrow -w_1 \in W$. Hence all the axioms hold in W so that W is a vector space in its own right. Thus W is a subspace of V . \square

By virtue of this theorem, for a subset W of V to be a subspace of V , it is sufficient to check only these three conditions.

Example 13.16. 1. Consider the vector space $V(F)$. Let $W = \{0\}$. Then

- (i) $0 \in W$
- (ii) $0 + 0 = 0 \in W$
- (iii) if $\alpha \in F$, then, $\alpha 0 = 0 \in W$

Thus W is a subspace of V .

2. Let $V = \mathbb{R}^2$. Then $V(\mathbb{R})$ is a vector space. Let

$W = \{(x, 0) | x \in \mathbb{R}\}$, then W is a subspace of V . Since

- (i) $(0, 0) \in W$
- (ii) If $(x_1, 0), (x_2, 0) \in W$ then $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \in W$
- (iii) For $\alpha \in \mathbb{R}, w = (x_1, 0) \in W$, $\alpha w = (\alpha x_1, 0) \in W$.

3. Let $V = \mathbb{R}^2$. Then $V(\mathbb{R})$ is a vector space. Let $W = \{(x, y) \in \mathbb{R}^2 | y = 2x\}$.

(i) Since $0 = 2 \cdot 0$

$\therefore (0, 0) \in W$

(ii) If $w_1 = (x_1, y_1)$ and $w_2 = (x_2, y_2)$ belongs to W , then $y_1 = 2x_1, y_2 = 2x_2$

$\therefore y_1 + y_2 = 2(x_1 + x_2)$ so that $(x_1 + x_2, y_1 + y_2) \in W$

(iii) Let $w \in W, \alpha \in \mathbb{R}$, and $w = (x, y)$

$\therefore y = 2x \Rightarrow \alpha y = 2\alpha x$

$\Rightarrow (\alpha x, \alpha y) \in W$

But $\alpha(x, y) = (\alpha x, \alpha y) \in W$

Hence W is a subspace of V .

If $V(F)$ is a vector space, then $\{0\}$ and V are subspaces of V . These are called the trivial subspaces of V . Any subspace of V other than $\{0\}$ and V is called a non-trivial subspace of V .

To prove that a subset W is not a subspace of a vector space $V(F)$, it is sufficient to show that any one of the three conditions of Theorem 13.2 fails to hold.

Example 13.17. Consider the vector space $V(\mathbb{R})$, where $V = \mathbb{R}^2$. Let

(i) $W_1 = \{(x, y) \in V | 2x + 3y = 4\}$. Then $2(0) + 3(0) = 0 \neq 4 \therefore (0, 0)$ does not satisfy $2x + 3y = 4$, so that $(0, 0) \notin W_1$. Hence W_1 is not a subspace of V .

(ii) $W_2 = \{(x, y) \in V | x^2 = y^2\}$. Clearly $0 = (0, 0) \in W_2$. Take $w_1 = (1, -1), w_2 = (-2, -2)$. Then $w_1, w_2 \in W_2$ and $w_1 + w_2 = (-1, -3) \notin W_2$. Hence W_2 is not a subspace of V .

(iii) Let $W_3 = \{(x, y) \in V | 2x + 3y \geq 0\}$
Clearly $(0, 0) \in W_3$. If $w \in W_3$, then $-1w \notin W_3$

Therefore W is not closed under scalar multiplication. Hence W_3 is not a subspace of V .

To check whether a given subset of a vector space V is a subspace, the three conditions in Theorem 13.2 can be reduced to other equivalent conditions.

Theorem 13.3. *Let $V(F)$ be a vector space and W a subset of V . Then the following are equivalent:*

1. W is a subspace of V
2. $W \neq \phi$ and
(i) $u + v \in W, \forall u, v \in W$ (ii) $\alpha u \in W, \forall u \in W, \alpha \in F$
3. $W \neq \phi$ and
 $\alpha u + \beta v \in W, \forall \alpha, \beta \in F, u, v \in W$
4. $W \neq \phi$ and
 $\alpha u + v \in W \forall u, v \in W, \alpha \in F$

Proof: Left to the reader. □

We now give methods to create new subspaces from given subspaces.

Theorem 13.4. *The intersection of two subspaces is a subspace.*

Proof: Let $V(F)$ be a vector space and W_1, W_2 be two subspaces of V . Let $W = W_1 \cap W_2$. Since $0 \in W_1, 0 \in W_2, \therefore 0 \in W_1 \cap W_2 = W$ so that $W \neq \phi$. Let $u, v \in W, \alpha \in F$, then $u, v \in W_1, u, v \in W_2$. Since W_1, W_2 are subspaces of V

$\therefore \alpha u + v \in W_1$ and $\alpha u + v \in W_2$ so that $\alpha u + v \in W_1 \cap W_2 = W$

Hence W is a subspace of V . □

The union of two subspaces may fail to be a subspace. This is shown by the following example.

Example 13.18. *Consider the vector space $V(\mathbb{R})$, where $V = \mathbb{R}^2$. Let $W_1 = \{(x, y) \in \mathbb{R}^2 | x = y\}$, $W_2 = \{(x, y) \in \mathbb{R}^2 | y = -x\}$. Then W_1, W_2 are subspaces of V . $w_1 = (2, 2) \in W_1, w_2 = (-2, 2) \in W_2$. Thus $w_1, w_2 \in W_1 \cup W_2$, but $w_1 + w_2 = (0, 4) \notin W_1 \cup W_2$. Hence $W_1 \cup W_2$ is not a subspace of V .*

Remark 13.4.

1. The only subspaces of \mathbb{R} are $\{0\}$ and \mathbb{R} .
2. The only non-trivial subspaces of \mathbb{R}^2 are of the form $\{(x, y) \in \mathbb{R}^2 | y = ax\}$ for some $a \in \mathbb{R}$. Geometrically, these are lines passing through the origin.
3. The only non-trivial subspace of \mathbb{R}^3 is the form
 $\{(x, y, z) \in \mathbb{R}^3 | ax + by + cz = 0\}$ for some $a, b, c \in \mathbb{R}$
or $\{(x, y, z) \in \mathbb{R}^3 | \frac{x}{a} = \frac{y}{b} = \frac{z}{c}\}$ for some $a, b, c \in \mathbb{R}$

Geometrically, these are planes or lines in spaces. The proof is given in chapter 14.

It is important to note that the subspaces of $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3$ have a very specific form. This will help the reader to verify whether a given subset of \mathbb{R}^2 or \mathbb{R}^3 is a subspace or not.

Definition 13.3. (Sum of two subspaces): Let V be a vector space over a field F , and W_1, W_2 be two subspaces of V . Then $W = \{w_1 + w_2 | w_1 \in W_1, w_2 \in W_2\}$ is called the sum of the subspaces W_1 and W_2 and is denoted by $W_1 + W_2$.

Theorem 13.5. *If V is a vector space over F , and W_1, W_2 are two subspaces of V , then $W_1 + W_2$ is a subspace of V , containing W_1 and W_2 . Moreover it is the smallest subspace of V containing both W_1 and W_2 .*

Proof: Let $W = W_1 + W_2$. If $u, v \in W, \alpha \in F$.

Then $u = u_1 + u_2, v = v_1 + v_2$, where $u_1, v_1 \in W_1, u_2, v_2 \in W_2$

Then

$$\begin{aligned}\alpha u + v &= \alpha(u_1 + u_2) + (v_1 + v_2) \\ &= (\alpha u_1 + v_1) + (\alpha u_2 + v_2)\end{aligned}$$

Since W_1, W_2 are subspaces of V ,

$\therefore (\alpha u_1 + v_1) \in W_1, (\alpha u_2 + v_2) \in W_2$

$\Rightarrow (\alpha u_1 + v_1) + (\alpha u_2 + v_2) \in W$

$\Rightarrow \alpha u + v \in W$

$\Rightarrow W$ is a subspace of V .

Now, if $w_1 \in W_1$, then $w_1 = w_1 + 0 \in W_1 + W_2$ ($\because 0 \in W_2$)

$\Rightarrow w_1 \in W$

$\Rightarrow W_1 \subseteq W$. Similarly $W_2 \subseteq W$.

Let T be a subspace of V containing W_1 and W_2 .

Let $u \in W = W_1 + W_2$, $\therefore u = w_1 + w_2$ for some $w_1 \in W_1, w_2 \in W_2$.

Since $W_1 \subseteq T$ and $W_2 \subseteq T$ $\therefore w_1, w_2 \in T$ so that $w_1 + w_2 \in T$

$\Rightarrow u \in T \Rightarrow W_1 + W_2 \subseteq T$. Hence $W \subseteq T$, so that W is the smallest subspace of V containing both W_1 and W_2 . \square

$V = \mathbb{R}^3$, is a vector space over \mathbb{R} . Let

$W_1 = \{(x, y, 0) | x, y \in \mathbb{R}\}$, $W_2 = \{(0, y, z) | y, z \in \mathbb{R}\}$. Then W_1 and W_2 are subspaces of V . If $v \in V$, then

$$\begin{aligned}v &= (x, y, z) \text{ for some } x, y, z \in \mathbb{R} \\ &= (x, 1, 0) + (0, y - 1, z) \\ &= w_1 + w_2 \text{ (say)}\end{aligned}$$

Clearly $w_1 \in W_1$ and $w_2 \in W_2$. Thus every element of V is expressible as the sum of an element of W_1 and an element of W_2 . We now show that this representation is not unique. Consider $v = (2, 4, 3) \in V$ then

$$\begin{aligned}v = (2, 4, 3) &= (2, 2, 0) + (0, 2, 3) \\ &= w_1 + w_2, \text{ where } w_1 \in W_1, w_2 \in W_2\end{aligned}$$

$$\begin{aligned}\text{Also } v = (2, 4, 3) &= (2, 1, 0) + (0, 3, 3) \\ &= w'_1 + w'_2, \text{ where } w'_1 \in W_1 \text{ and } w'_2 \in W_2\end{aligned}$$

Thus v is expressed as the sum of an element of W_1 and an element of W_2 in two different ways.

We are interested in knowing the conditions on W_1 and W_2 so that the representation of an element of V as the sum of an element of W_1 and an element of W_2 is unique. The following theorem answers this question.

Theorem 13.6. *Let V be a vector space over a field F and W_1 and W_2 be two subspaces of V over F , such that $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$. Then every vector $v \in V$ is expressible uniquely as $w_1 + w_2$, for some $w_1 \in W_1, w_2 \in W_2$.*

Proof: From the definition it is clear that if $v \in V$, then there exist vectors $w_1 \in W_1, w_2 \in W_2$ such that $v = w_1 + w_2$.

Uniqueness

For some $v \in V$, suppose there are two expressions of v as the sum of an element of W_1 and an element of W_2 , i.e.

$v = w_1 + w_2$ and also

$v = w_1' + w_2'$ for some $w_1, w_1' \in W_1, w_2, w_2' \in W_2$. Then

$w_1 + w_2 = w_1' + w_2'$

$\Rightarrow w_1 - w_1' = w_2' - w_2 = x$ (say)

Since $w_1 - w_1' \in W_1, w_2' - w_2 \in W_2$

$\therefore x \in W_1 \cap W_2 = \{0\}$

$\Rightarrow x = 0$

$\Rightarrow w_1 - w_1' = 0 = w_2' - w_2$

$\Rightarrow w_1 = w_1'$ and $w_2 = w_2'$

\Rightarrow expression for v as an element of $W_1 + W_2$ is unique. \square

Definition 13.4. A vector space $V(F)$ is said to be the direct sum of two subspaces W_1 and W_2 , if

(i) $V = W_1 + W_2$

(ii) every element of V is expressed uniquely as the sum of an element of W_1 and an element of W_2 .

We write $V = W_1 \oplus W_2$.

Theorem 13.7. If $V(F)$ is a vector space and W_1, W_2 are two subspaces of V then $V = W_1 \oplus W_2 \Leftrightarrow V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$

Proof: Follows from Theorem 13.6. \square

Problem 13.5. Is $W = \{(x, y, z) \in V \mid y = -4x, z = 5x\}$ a subspace of $V(\mathbb{R})$? where $V = \mathbb{R}^3$.

Solution:

(i) Clearly $(0, 0, 0) \in W$, so that $W \neq \phi$

(ii) Let $u, v \in W, \alpha \in \mathbb{R}$

Then $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ for some $u_i, v_i \in \mathbb{R}, 1 \leq i \leq 3$ such that $u_2 = -4u_1, u_3 = 5u_1,$

$v_2 = -4v_1, v_3 = 5v_1$. Observe that $u_2 + v_2 = -4(u_1 + v_1), u_3 + v_3 = 5(u_1 + v_1)$ so that $(u_1 + v_1, u_2 + v_2, u_3 + v_3) \in W$, i.e $u + v \in W$.

(iii) For any $\alpha \in \mathbb{R}, \alpha u_2 = -4\alpha u_1, \alpha u_3 = 5\alpha u_1$

$\Rightarrow (\alpha u_1, \alpha u_2, \alpha u_3) \in W$

$\Rightarrow \alpha u \in W$

Hence W is closed under addition and scalar multiplication. Thus W is a subspace of V .

Problem 13.6. Let $V = M_3(\mathbb{R})$ be the vector space of all 3×3 matrices over \mathbb{R} . Is $W = \{A \in V \mid |A| \neq 0\}$ a subspace of V ?

Solution: $0 \notin W$. Hence W is not a subspace of V .

Problem 13.7. Let V be the vector space of all real valued functions on \mathbb{R} . If

(i) $W_1 = \{f \in V \mid f \text{ is a solution of } y'' + 3y' - 5y = 0\}$

(ii) $W_2 = \{f \in V \mid f \text{ is increasing}\}$

then, determine if they are subspaces of $V(\mathbb{R})$?

Solution:

(i) Clearly the zero element of V , namely the zero function belongs to W_1 .

Hence W_1 is non-empty.

Let $f, g \in W_1, \alpha, \beta \in \mathbb{R}$. Then

$$f'' + 3f' - 5f = 0 \quad (13.1)$$

$$g'' + 3g' - 5g = 0 \quad (13.2)$$

$$\text{Now, } (\alpha f + \beta g)'' = \alpha f'' + \beta g''$$

$$(\alpha f + \beta g)' = \alpha f' + \beta g'$$

$$\begin{aligned} & (\alpha f + \beta g)'' + 3(\alpha f + \beta g)' - 5(\alpha f + \beta g) \\ &= \alpha(f'' + 3f' - 5f) + \beta(g'' + 3g' - 5g) \\ &= 0 \quad (\text{using (13.1) and (13.2)}) \end{aligned}$$

$\therefore \alpha f + \beta g \in W_1$, so that W_1 is a subspace of V .

(ii) Consider the function f defined on \mathbb{R} by

$$f(x) = 2x,$$

Then f is an increasing function. $\therefore f \in W_2$. Consequently $-f$ is not an increasing function so that $-f \notin W$. Hence W is not a subspace of V .

Problem 13.8. For the vector space $V(\mathbb{R})$, where $V = M_n(\mathbb{R})$. Is the set of all $n \times n$ non-singular matrices a subspace of V ?

Solution: Let $W =$ set of all $n \times n$ non-singular matrices. Since the sum of two non-singular matrices may not be non-singular, this leads us to believe that W is not a subspace of V . Consider

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 1 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & -1 \end{pmatrix}$$

Thus $|A| = 1 \neq 0$, $|B| \neq (-1)^n \neq 0 \quad \therefore A, B \in W$

$$\text{But } A + B = \begin{pmatrix} 0 & \dots & \dots & 0 \\ 2 & \ddots & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 2 & \dots & 2 & 0 \end{pmatrix}, \quad \text{and } |A + B| = 0$$

so that $A + B \notin W$. Hence W is not a subspace of V .

Problem 13.9. Let V be a vector space over a field, and W_1, W_2 be two subspaces of V . Then $W_1 \cup W_2$ is a subspace of V if and only if either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Solution: *The condition is necessary.*

Let $W = W_1 \cup W_2$. Suppose that W is a subspace of V . Let, if possible, $W_1 \not\subseteq W_2$ and $W_2 \not\subseteq W_1$. Then $\exists w_1 \in W_1$ such that $w_1 \notin W_2$ and $w_2 \in W_2$ and such that $w_2 \notin W_1$.

Thus $w_1, w_2 \in W_1 \cup W_2$

$\Rightarrow w_1 + w_2 \in W_1 \cup W_2$ ($\because W_1 \cup W_2$ is a subspace)

$\Rightarrow w_1 + w_2 \in W_1$ or $w_1 + w_2 \in W_2$.

If $w_1 + w_2 \in W_1$, then $(w_1 + w_2) - w_1 \in W_1$, i.e. $w_2 \in W_1$, a contradiction to the fact that $w_2 \notin W_1$. Similarly $w_1 + w_2 \in W_2$ leads to a contradiction.

Hence our assumption is wrong so that we must have $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

The condition is sufficient.

Conversely, let $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$. If $W_1 \subseteq W_2$, then $W_1 \cup W_2 = W_2$ which is a subspace of V . Similarly if $W_2 \subseteq W_1$

$\Rightarrow W_1 \cup W_2 = W_1$, which is a subspace of V .

Hence the condition is sufficient.

Problem 13.10. Consider the vector space $V(\mathbb{R})$, where V = set of all functions from \mathbb{R} into \mathbb{R} .

Let W_1 be the set of all even functions and W_2 be the set of all odd functions.

Prove that

(i) W_1, W_2 are subspaces of V

(ii) $V = W_1 + W_2$

(iii) $W_1 \cap W_2 = \{0\}$

(iv) $V = W_1 \oplus W_2$.

Solution:

(i) Let $f, g \in W_1, \alpha \in \mathbb{R}$

$$\text{Then } f(-x) = f(x), \forall x \in \mathbb{R} \quad (1)$$

$$\text{and } g(-x) = g(x), \forall x \in \mathbb{R} \quad (2)$$

$$\begin{aligned} \text{Then } (\alpha f + g)(-x) &= (\alpha f)(-x) + g(-x) \\ &= \alpha f(-x) + g(-x) \\ &= \alpha f(x) + g(x) \quad \text{using (1) and (2)} \\ &= (\alpha f + g)(x) \end{aligned}$$

$\therefore \alpha f + g \in W_1$, so that W_1 is a subspace of V .

Similarly W_2 is a subspace of V .

(ii) Let $f \in V$. Define functions g and h as follows:

$$g(x) = \frac{1}{2}(f(x) + f(-x)), \forall x \in \mathbb{R}$$

$$h(x) = \frac{1}{2}(f(x) - f(-x)), \forall x \in \mathbb{R}$$

$$\begin{aligned} \text{Then } g(-x) &= \frac{1}{2}(f(-x) + f(x)) \\ &= g(x) \end{aligned}$$

$\therefore g \in W_1$.

$$\text{Similarly } h(-x) = -h(x)$$

$\therefore h \in W_2$.

Also

$$\begin{aligned} f(x) &= \frac{1}{2}(f(x) + f(-x)) + \frac{1}{2}(f(x) - f(-x)) \\ &= g(x) + h(x) \end{aligned}$$

Thus f is expressible as the sum of an element of W_1 and an element of W_2 .

(iii) $f \in W_1 \cap W_2$

$\Leftrightarrow f \in W_1$, and $f \in W_2$

$\Leftrightarrow f(-x) = f(x), \forall x \in \mathbb{R}$ and $f(-x) = -f(x), \forall x \in \mathbb{R}$

$\Leftrightarrow f(x) = -f(x), \forall x \in \mathbb{R}$

$\Leftrightarrow f(x) = 0, \forall x \in \mathbb{R}$

$\Leftrightarrow f$ is the zero function.

Hence $W_1 \cap W_2 = \{0\}$.

(iv) Follows from (i) to (iii).

13.4 Exercise

1. $V(\mathbb{R})$ is a vector space, where $V = \mathbb{R}^3$. Verify whether W is a subspace of V , where

(i) $W = \{(x_1, x_2, 0) \in V \mid x_1, x_2 \in \mathbb{R}\}$

(ii) $W = \{(x, x, x) \in V \mid x \in \mathbb{R}\}$

(iii) $W = \{(x, y, z) \in V \mid x + y + z = 0\}$

(iv) $W = \{(x, y, z) \in V \mid x + y + z = 1\}$

(v) $W = \{(x, y, z) \in V \mid x^2 + y^2 + z^2 = 9\}$.

2. $V(\mathbb{R})$ is a vector space, where $V = \mathbb{R}^2$. Verify whether W is a subspace of V , where

(i) $W = \{(x, y) \in V \mid x^2 + y^2 = 1\}$

(ii) $W = \{(x, y) \in V \mid 3x - 4y = 0\}$

(iii) $W = \{(x, y) \in V \mid \frac{x^2}{9} + \frac{y^2}{16} = 1\}$

(iv) $W = \{(x, y) \in V \mid xy = 0\}$

(v) $W = \{(0, y) \in V \mid y \in \mathbb{R}\}$

(vi) $W = \{(x, y) \in V \mid x \leq y\}$.

3. Determine, whether the following subsets of \mathbb{R}^3 are subspaces or not. Give reasons for your answer.

(i) $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 3z = 1\}$

(ii) $W = \{(x, y, z) \in \mathbb{R}^3 \mid z = 1\}$

(iii) $W = \{(x, y, z) \in \mathbb{R}^3 \mid y = 2x, z = -x\}$

(iv) $W = \{(x, 2x, 3x) \in \mathbb{R}^3 \mid x \in \mathbb{R}\}$

(v) $W = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\}$

(vi) $\{(x + y, x - y, x) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}$

(vii) $\{(x + 2y, x + 1, y) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}$

(viii) $\{(x, 5, y) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}$.

4. Determine, whether W is a subspace of \mathbb{R}^2 or not. Give reasons for your answer.
- $W = \{(x, y) \in V | xy \geq 0\}$
 - $W = \{(x, y) \in V | x = |y|\}$
 - $W = \{(x, y) \in V | y = 2x + 1\}$
 - $W = \{(x, y) \in V | y = x^2\}$
 - $W = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ for a fixed 2×2 matrix over \mathbb{R} .
5. $V(\mathbb{R})$ is a vector space where $V = \{a_0 + a_1x + a_2x^2 | a_0, a_1, a_2 \in \mathbb{R}\}$. Determine whether W is a subspace of V . Justify your answer.
- $W = \{a_0 + a_1x | a_0, a_1 \in \mathbb{R}\}$
 - $W = \{a_0 + a_1x + a_2x^2 | a_0 + a_1 + a_2 = 0\}$
 - $W = \{a_0 + a_1x + a_2x^2 | a_0 = 0\}$
 - $W = \{a_0 + x^2 | a_0 \in \mathbb{R}\}$
 - $W = \{ax^2 | a \in \mathbb{R}\}$
 - $W = \{p(x) \in V | p(0) = 5\}$
 - Let $\alpha \in \mathbb{R}$ be fixed. $W = \{p(x) \in V | p(\alpha) = 0\}$.
6. Let $V = \mathcal{M}_{2 \times 2}(\mathbb{R})$ be the vector space of 2×2 matrices over \mathbb{R} . Determine whether or not the following subsets of V are subspaces of V .
- $W = \{A \in V | |A| = 0\}$
 - $W = \{A \in V | |A| = 1\}$
 - $W = \{A \in V | |A| \neq 0\}$
 - $W = \{A \in V | A \text{ is a diagonal matrix}\}$
 - $W = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \in V | a \in \mathbb{R} \right\}$
 - $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V | a + d = b + c \right\}$.
7. Let V be the space of all real valued functions on $[0, 1]$. Check whether W is a subspace of V . Justify your answer.
- $W = \{f \in V | f \text{ is continuous}\}$
 - $W = \{f \in V | f \text{ is differentiable}\}$
 - $W = \{f \in V | f(1/2) = 2\}$
 - $W = \{f \in V | f \text{ is a solution of } y'' + y = 0\}$
 - $W = \{f \in V | f(0) = f(1)\}$
 - $W = \{f \in V | \text{Range } f \text{ of finite}\}$
 - $W = \{f \in V | f \text{ is non decreasing}\}$.
8. $V(\mathbb{R})$ is a vector space, where $V = \mathbb{C}$. Verify whether W is a subspace of V . Justify your answer.
- $W = \{v \in V | \text{Re } v = 0\}$
 - $W = \{v \in V | \text{Im } v = 0\}$
 - Let k be a fixed real number. $W = \{a + ib \in V | b = ka\}$
 - $W = \{v \in V | |v| = 1\}$
 - $W = \{v \in V | \text{Re } v \geq 0\}$.
9. $V(\mathbb{C})$ is a vector space, where $V = \mathbb{C}^n$. Do the following sets form a subspace of V ? Justify your answer.
- $W = \{(x_1, x_2, \dots, x_n) \in \mathbb{C}^n | x_1 \in \mathbb{R}\}$
 - $W = \{(x_1, x_2, \dots, x_n) \in \mathbb{C}^n | |x_1| = 1\}$.

10. Let A be a non-empty set and $V = \mathcal{P}(A)$. Then $V(\mathbb{Z}_2)$ is a vector space as defined in Problem 13.4. Verify that $\{\phi, A\}$ is a subspace of V .
11. Consider the vector space $V(\mathbb{R})$ where $V = M_n(\mathbb{R})$. Let $W_1 =$ set of all symmetric matrices of V and $W_2 =$ set of all skew symmetric matrices of V . Prove that $V = W_1 \oplus W_2$.
12. Let $V = \mathbb{C}^n$ then $V(\mathbb{C})$ is a vector space. Let $W = \{(x_1, x_2, \dots, x_n) \in \mathbb{C}^n \mid x_i \in \mathbb{R}\}$. Is W a subspace of V ? If not, can you modify the vector space $V(\mathbb{C})$, so that W becomes a subspace.

13.5 Linear Span of a Subset

In Chapter 12 we have defined a linear combination of vectors in \mathbb{R}^n . We would now like to extend this concept to a general vector space.

Definition 13.5. (Linear combination): Let $V(F)$ be a vector space and $S = \{v_1, v_2, \dots, v_n\}$ be a finite subset of V . Then the vector

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n, \quad \alpha_i \in F, \quad 1 \leq i \leq n$$

is called a linear combination of the elements of S . The above expression is also written as $\sum_{i=1}^n \alpha_i v_i$.

Definition 13.6. (Linear span): Let $V(F)$ be a vector space and S a non-empty subset of V . Then linear span of S is the set of all linear combinations of finitely many elements of S . The linear span of null set is $\{0\}$.

It is denoted by $\text{Span}(S)$ (or $L(S)$ or $[S]$). We shall use $\text{Span}(S)$.

Remark 13.5.

1. If S is a finite set then $\text{Span}(S)$ is the set of all linear combination of elements of S .
2. The linear span of a set S is also called the span of S .

Theorem 13.8. Let $V(F)$ be a vector space and S a finite subset of V . Then $\text{Span}(S)$ is the smallest subspace of V containing S .

Proof: Two cases arise:

Case 1. $S = \phi$. Then by definition $\text{Span}(S) = \{0\}$, which is a subspace of V . If W be a subspace of V containing S .

Then $\text{Span}(S) = \{0\} \subseteq W$

$\therefore \text{Span}(S)$ is the smallest subspace of V containing S .

Case 2. $S \neq \phi$, let $S = \{v_1, v_2, \dots, v_n\}$. Then $\text{Span}(S) = \{\sum_{i=1}^n \alpha_i v_i \mid \alpha_i \in F, 1 \leq i \leq n\}$.

Let $s \in S$, then $s = 1s, 1 \in F$ so that $s \in \text{Span}(S)$, $\therefore \text{Span}(S) \neq \emptyset$. Hence $S \subseteq \text{Span}(S)$.

Let $s_1, s_2 \in \text{Span}(S)$ and $\alpha \in F$. Then $s_1 = \sum_{i=1}^n \alpha_i v_i, s_2 = \sum_{i=1}^n \beta_i v_i$ for

some $\alpha_i, \beta_i \in F, 1 \leq i \leq n$.

$$\begin{aligned}\alpha s_1 + s_2 &= \alpha \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^n \beta_i v_i \\ &= \sum_{i=1}^n (\alpha \alpha_i + \beta_i) v_i \in \text{Span}(S) \text{ as } \alpha \alpha_i + \beta_i \in F \forall i\end{aligned}$$

$\therefore \text{Span}(S)$ is a subspace of V .

Let W be a subspace of V containing S . Let $u \in \text{Span}(S)$, so that $\exists \alpha_1, \dots, \alpha_n \in F$ such that $u = \sum_{i=1}^n \alpha_i v_i$.

Now $v_i \in W, \alpha_i \in F, 1 \leq i \leq n$, and W is a subspace.

$\therefore \sum_{i=1}^n \alpha_i v_i \in W$.

Hence $u \in W$, so that $\text{Span}(S) \subset W$. \square

Theorem 13.9. Let $V(F)$ be a vector space and S an infinite subset of V . Then $\text{Span}(S)$ is the smallest subspace of V containing S .

Proof: If $s \in S$ then $s = 1s$, so that $s \in \text{Span}(S)$. Therefore $S \subseteq \text{Span}(S)$.

Let $s_1, s_2 \in \text{Span}(S)$. Then

$s_1 = \sum_{i=1}^n \alpha_i u_i, s_2 = \sum_{j=1}^m \beta_j v_j$, where $u_i, v_j \in S, \alpha_i, \beta_j \in F, 1 \leq i \leq n, 1 \leq j \leq m$. If $\alpha \in F$, then $\alpha s_1 + s_2 = \sum_{i=1}^n \alpha \alpha_i u_i + \sum_{j=1}^m \beta_j v_j$.

Hence $\alpha s_1 + s_2$ is a linear combination of finitely many elements of S .

$\therefore \text{Span}(S)$ is a subspace of V .

Let W be a subspace of V containing S . Clearly $S \subseteq W$. Let $s \in \text{Span}(S)$

Then there exist finitely many vectors u_1, u_2, \dots, u_n of S and $\alpha_i \in F, 1 \leq i \leq n$ such that $s = \sum_{i=1}^n \alpha_i u_i$. Since $u_i \in W$ and $\alpha_i \in F, 1 \leq i \leq n, \therefore \sum_{i=1}^n \alpha_i u_i \in W$ (as W is a subspace) so that $s \in W$ i.e., $\text{Span}(S) \subset W$.

Thus $\text{Span}(S)$ is the smallest subspace of V containing S . \square

Remark 13.6. In view of the above theorem $\text{Span}(S)$ is also called the subspace of V generated by S , or spanned by S .

Definition 13.7. (Finitely generated): A subspace W of a vector space $V(F)$ is said to be finitely generated if there exist a finite subset S of W such that $W = \text{Span}(S)$.

Theorem 13.10. Let $V(F)$ be a vector space. If S is a subset of V , then

1. $\text{Span}(\text{Span}(S)) = \text{Span}(S)$.
2. If S is a subspace of V , then $\text{Span}(S) = S$

Proof: Follows immediately from Theorem 13.9. \square

Example 13.19.

1. Let $V = \mathbb{R}^2$. Take $v = (0, 0)$
 $\text{Span}(\{v\}) = \{\alpha v \mid \alpha \in F\} = \{\alpha(0, 0) \mid \alpha \in \mathbb{R}\} = \{(0, 0)\}$
 Thus the span of the zero vector is a zero space.
2. Let $V = \mathbb{R}^2$. Take $v = (1, 2)$.
 $\text{Span}(\{v\}) = \{\alpha v \mid \alpha \in F\} = \{\alpha(1, 2) \mid \alpha \in \mathbb{R}\} = \{(\alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$
 which represents the line passing through origin and the point $(1, 2)$, Hence we can say that the $\text{Span}(\{v\})$ is the line containing all scalar multiples of the vector $\overrightarrow{OA} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

3. Let $V = \mathbb{R}^2$. Take $v_1 = (1, 2), v_2 = (2, 3)$
 $Span(\{v_1, v_2\}) = \{\alpha_1 v_1 + \alpha_2 v_2 | \alpha_1, \alpha_2 \in \mathbb{R}\} = \{(\alpha_1 + 2\alpha_2, 2\alpha_1 + 3\alpha_2) | \alpha_1, \alpha_2 \in \mathbb{R}\}$.
 This is whole of \mathbb{R}^2 because every vector in \mathbb{R}^2 can be expressed in the form $\alpha v_1 + \beta v_2$. For, $(x_1, x_2) \in V$, $(x_1, x_2) = (2x_2 - x_1)v_1 + (x_1 - x_2)v_2 \Rightarrow V \subseteq span(\{v_1, v_2\}) \subseteq V \Rightarrow V = span(\{v_1, v_2\})$.

4. Let $V = \mathbb{R}^3$, $W = \{(\alpha + \beta, \alpha - \beta, \alpha) | \alpha, \beta \in \mathbb{R}\}$. Then

$$\begin{aligned}(\alpha + \beta, \alpha - \beta, \alpha) &= (\alpha, \alpha, \alpha) + (\beta, -\beta, 0) \\ &= \alpha(1, 1, 1) + \beta(1, -1, 0)\end{aligned}$$

Therefore $W = \{(\alpha + \beta, \alpha - \beta, \alpha) | \alpha, \beta \in \mathbb{R}\}$ is the linear span of v_1, v_2 , where $v_1 = (1, 1, 1), v_2 = (1, -1, 0)$. Hence W is a subspace of \mathbb{R}^3 .

Example 13.20. Consider $V(F)$, where $V = \mathbb{C}, F = \mathbb{C}$. Let $S = \{1\}$.

$$Span(S) = \{\alpha 1 | \alpha \in \mathbb{C}\} = \{\alpha | \alpha \in \mathbb{C}\} = \mathbb{C}.$$

Now take $F = \mathbb{R}$. Then V is a vector space over \mathbb{R} .

$$Span(S) = \{\alpha 1 | \alpha \in \mathbb{R}\} = \{\alpha | \alpha \in \mathbb{R}\} = \mathbb{R}.$$

Thus $Span(S)$ depends upon the field over which the vector space is considered.

What is $Span(S)$ when $F = \mathbb{Q}$?

The next theorem give us some basic properties of the span of a set.

Theorem 13.11. Let $V(\mathbb{F})$ be a vector space, and S_1, S_2 two subsets of V .

- (i) If $S_1 \subseteq S_2$ then $Span(S_1) \subseteq Span(S_2)$
- (ii) $Span(S_1) \cup Span(S_2) \subseteq Span(S_1 \cup S_2)$
- (iii) $Span(S_1 \cap S_2) \subseteq Span(S_1) \cap Span(S_2)$
- (iv) $Span(S_1 \cup S_2) = Span(S_1) + Span(S_2)$

Proof: Proof left to the reader. □

Example 13.21. Consider the vector space $V(\mathbb{R})$, where $V = \mathcal{P}_3$. Let $S_1 = \{1, x\}, S_2 = \{1, 1 + x\}$

$$\begin{aligned}Span(S_1) &= \{\alpha 1 + \beta x | \alpha, \beta \in \mathbb{R}\} \\ &= \{\alpha + \beta x | \alpha, \beta \in \mathbb{R}\}\end{aligned}$$

Now $1 = 1 + 0 \cdot x \in Span(S_1), 1 + x = 1 + 1 \cdot x \in Span(S_1)$,

$\therefore S_2 \subseteq Span(S_1)$, so that

$$Span(S_2) \subseteq Span(Span(S_1)) = Span(S_1) \tag{13.3}$$

$Span(S_2) = \{\alpha' 1 + \beta' (1 + x) | \alpha', \beta' \in \mathbb{R}\}$
 $1 = 1 \cdot 1 + 0(1 + x) \in Span(S_2), x = (-1) \cdot 1 + 1(1 + x) \in Span(S_2)$
 $\therefore S_1 \subseteq Span(S_2)$, so that as in (13.3)

$$Span(S_1) \subseteq Span(S_2) \tag{13.4}$$

From (13.3) and (13.4), $Span(S_1) = Span(S_2)$.

Note that $S_1 \neq S_2$, thus two different sets may have the same span. This important observation will be used later.

13.6 Column Space

Definition 13.8. Let $A = [C_1 \ C_2 \ \dots \ C_n]$ be a $m \times n$ matrix. The set of all linear combinations of the columns of A is called the column space of A and is written as $\text{Col } A$. Thus

$$\text{Col } A = \text{Span}\{C_1, C_2, \dots, C_n\}$$

Theorem 13.12. The column space of a $m \times n$ matrix A is a subspace of \mathbb{R}^m .

Proof: Let $A = [C_1 \ C_2 \ \dots \ C_n]$, where $C_i \in \mathbb{R}^m$. $\text{Col } A = \text{Span}(\{C_1, C_2, \dots, C_n\})$. Since $\text{Span}(S)$ is a subspace, therefore $\text{Col } A$ is a subspace of \mathbb{R}^m . \square

Another description of Col A

If A is a $m \times n$ matrix, with columns C_1, C_2, \dots, C_n then a typical element of

$$\text{Col } A \text{ is } \alpha_1 C_1 + \alpha_2 C_2 + \dots + \alpha_n C_n = [C_1 \ C_2 \ \dots \ C_n] \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = Au, \text{ where } u =$$

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n. \text{ Thus}$$

$$\begin{aligned} \text{Col } A &= \{Au \mid u \in \mathbb{R}^n\} \\ &= \{b \in \mathbb{R}^m \mid AX = b \text{ for some } X \in \mathbb{R}^n\} \\ &= \{b \in \mathbb{R}^m \mid AX = b \text{ has a solution}\} \end{aligned}$$

In particular, if $AX = b$ has a solution for every $b \in \mathbb{R}^m$, then $\text{Col } A = \mathbb{R}^m$.
 \Leftrightarrow every row of A has a pivot.

Example 13.22. 1. If $A = \begin{pmatrix} 2 & -1 \\ 3 & 0 \\ 4 & 1 \end{pmatrix} = [C_1 \ C_2]$.

$$\begin{aligned} \text{Then } \text{Col } A &= \{\alpha_1 C_1 + \alpha_2 C_2, \alpha_1, \alpha_2 \in \mathbb{R}\} \\ &= \left\{ \begin{pmatrix} 2\alpha_1 - \alpha_2 \\ 3\alpha_1 \\ 4\alpha_1 + \alpha_2 \end{pmatrix}, \alpha_1, \alpha_2 \in \mathbb{R} \right\} \end{aligned}$$

2. If $A = \begin{pmatrix} -6 & 12 \\ -3 & 6 \end{pmatrix}, w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Does $w \in \text{Col } A$?

$$\text{Here } A = [v_1 \ v_2], \text{ where } v_1 = \begin{pmatrix} -6 \\ -3 \end{pmatrix}, v_2 = \begin{pmatrix} 12 \\ 6 \end{pmatrix}$$

$$w \in \text{Col } A \text{ if } w = \alpha_1 v_1 + \alpha_2 v_2 \text{ for some } \alpha_1, \alpha_2 \in \mathbb{R}, \text{ i.e. } [v_1 \ v_2] \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = w$$

i.e. $AX = w$ (1) has a solution $(\alpha_1, \alpha_2)^t$

Here $[A \mid w] \sim \left(\begin{array}{cc|c} -6 & 12 & 2 \\ -3 & 6 & 1 \end{array} \right)$. Since the augmented column is not a pivot column, \therefore (1) has a solution. Hence $w \in \text{Col } A$.

Definition 13.9. Let A be a $m \times n$ matrix. The set of all solutions of $AX = 0$ is called the null space of A and is written as $\text{Nul } A$. Thus, $\text{Nul } A = \{X \in \mathbb{R}^n \mid AX = 0\}$.

Theorem 13.13. *The null space of a $m \times n$ matrix A is a subspace of \mathbb{R}^n .*

Proof: Since $A \cdot 0 = 0$, $\therefore 0 \in \text{Nul } A$ so that $\text{Nul } A \neq \emptyset$. Let $X_1, X_2 \in \mathbb{R}^n, \alpha \in \mathbb{R}$. Then $AX_1 = 0, AX_2 = 0$.

$$\begin{aligned} A(\alpha X_1 + X_2) &= \alpha AX_1 + AX_2 \\ &= \alpha 0 + 0 \\ &= 0 \end{aligned}$$

Hence $\text{Nul } A$ is a subspace of \mathbb{R}^n . □

Example 13.23. Let $A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 3 \end{pmatrix}$. Find $\text{Nul } A$. Does $\begin{pmatrix} 0 \\ 3 \end{pmatrix} \in \text{Nul } A$?

Solution: $\text{Nul } A = \{X \in \mathbb{R}^2 \mid AX = 0\}$

$$\begin{aligned} \text{Now } X &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \text{Nul } A \\ \Leftrightarrow \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

$$\therefore 2x_1 = 0$$

$$-x_1 + x_2 = 0$$

$$3x_2 = 0$$

$$\Rightarrow x_1 = 0, x_2 = 0.$$

Hence $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Thus $\text{Nul } A = \{0\}$. Obviously $\begin{pmatrix} 0 \\ 3 \end{pmatrix} \notin \text{Nul } A$.

Problem 13.11. In the vector space $V(\mathbb{R}), V = \mathbb{R}^3$, find two generating sets for $W = \{(\alpha, \alpha + \beta, -\alpha + 4\beta) \mid \alpha, \beta \in \mathbb{R}\}$.

Solution: If $w \in W$, then

$$\begin{aligned} w &= (\alpha, \alpha + \beta, -\alpha + 4\beta) \text{ for some } \alpha, \beta \in \mathbb{R} \\ &= \alpha(1, 1, -1) + \beta(0, 1, 4) \end{aligned}$$

Thus if, $u_1 = (1, 1, -1), u_2 = (0, 1, 4)$, then $\{u_1, u_2\}$ spans W . To find another spanning set we need to only change the parameters, α and β .

Let $\alpha = \alpha, \gamma = \alpha + \beta$. Then if $w \in W$, we write w in terms of α and γ .

$$\begin{aligned} w &= (\alpha, \gamma, -\alpha + 4(\gamma - \alpha)) \\ &= (\alpha, \gamma, 4\gamma - 5\alpha) \\ &= \alpha(1, 0, -5) + \gamma(0, 1, 4) \end{aligned}$$

Hence if $v_1 = (1, 0, -5), v_2 = (0, 1, 4)$ then $\text{Span}(\{v_1, v_2\}) = W$.

Problem 13.12. In the vector space $\mathbb{R}^3(\mathbb{R})$, consider

$$W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x + 2y + 3z = 0 \right\}.$$

Give three different sets of vectors $S = \{u, v\}$ such that $W = \text{Span}(S)$.

Solution: Let $w \in W$. Then for some $x, y, z \in \mathbb{R}$

$$\begin{aligned} w &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ such that } x + 2y + 3z = 0 & (1) \\ &= \begin{pmatrix} x \\ y \\ -\frac{x+2y}{3} \end{pmatrix} \text{ using (1)} \\ &= x \begin{pmatrix} 1 \\ 0 \\ -1/3 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -2/3 \end{pmatrix} \\ &= xu + zv, \text{ where } u = \begin{pmatrix} 1 \\ 0 \\ -1/3 \end{pmatrix}, v = \begin{pmatrix} 0 \\ 1 \\ -2/3 \end{pmatrix} \end{aligned}$$

Thus $W = \text{Span}(S)$, where $S = \{u, v\}$.

If

$$\begin{aligned} w &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} x \\ -\frac{x+3z}{2} \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ -1/2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ -3/2 \\ 1 \end{pmatrix} \\ &= xu_1 + zv_1, \text{ where } u_1 = \begin{pmatrix} 1 \\ -1/2 \\ 0 \end{pmatrix}, v_1 = \begin{pmatrix} 0 \\ -3/2 \\ 1 \end{pmatrix} \end{aligned}$$

and $W = \text{Span}(S)$, where $S = \{u_1, v_1\}$. Similarly if we substitute for x in terms of y and z , we get

$$\begin{aligned} w &= \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} y + \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} z \\ &= u_2 y + v_2 z \end{aligned}$$

where $u_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$. Hence $W = \text{Span}(S)$, where $S = \{u_2, v_2\}$.

Problem 13.13. Prove that W is a subspace of $\mathbb{R}^4(\mathbb{R})$,

$$\text{where } W = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mid a - 2b = 4c, 2a = c + 3d \right\}.$$

Solution: If we can find a subset S such that $W = \text{Span}(S)$ then W will be a subspace, since $\text{span}(S)$ is always a subspace. Now $a - 2b = 4c, 2a = c + 3d \Leftrightarrow$

$$b = \frac{1}{2}(a - 4c), d = \frac{1}{3}(2a - c)$$

$$\begin{aligned} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} &= \begin{pmatrix} a \\ \frac{1}{2}a - 2c \\ c \\ \frac{2}{3}a - \frac{1}{3}c \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \\ 0 \\ 2/3 \end{pmatrix} a + \begin{pmatrix} 0 \\ -2 \\ 1 \\ -1/3 \end{pmatrix} c \\ &= u_1 a + u_2 c \quad (\text{say}) \end{aligned}$$

Thus

$$\begin{aligned} W &= \{u_1 a + u_2 c \mid a, c \in \mathbb{R}\} \\ &= \text{Span}(\{u_1, u_2\}) \end{aligned}$$

Hence W is a subspace of $\mathbb{R}^4(\mathbb{R})$.

Problem 13.14. Does $(2, -5, 3)^t$ lie in the subspace spanned by $\{(1, -3, 2)^t, (2, -4, -1)^t, (1, -5, 7)^t\}$?

Solution: Let $v_1 = (1, -3, 2)^t, v_2 = (2, -4, -1)^t, v_3 = (1, -5, 7)^t, u = (2, -5, 3)^t$

$u \in \text{Span}(\{v_1, v_2, v_3\}) \Leftrightarrow \exists \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ such that $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = u$

$$\Leftrightarrow [v_1 \ v_2 \ v_3] \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = u$$

$$\Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ -3 & -4 & -5 \\ 2 & -1 & 7 \end{pmatrix} X = \begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix}, \text{ where } X = (\alpha_1 \ \alpha_2 \ \alpha_3)^t$$

$\Leftrightarrow X$ is the solution of $AX = u$, where $A = [v_1 \ v_2 \ v_3]$

$$\text{But } [A|u] \sim \left(\begin{array}{ccc|c} \mathbf{1} & 2 & 1 & 2 \\ 0 & \mathbf{1} & -1 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{3}{10} \end{array} \right)$$

Since the augmented column has a pivot. $\therefore AX = u$ does not have a solution. So $u \notin \text{Span}(\{v_1, v_2, v_3\})$.

Problem 13.15. Let $W = \left\{ \begin{pmatrix} -2l + 3n \\ l + m + n \\ 2m - n \\ n \end{pmatrix} \mid l, m, n \in \mathbb{R} \right\}$

Find a matrix A whose column space is W .

Solution:

$$\begin{aligned} \begin{pmatrix} -2l + 3n \\ l + m + n \\ 2m - n \\ n \end{pmatrix} &= \begin{pmatrix} -2l \\ l \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ m \\ 2m \\ 0 \end{pmatrix} + \begin{pmatrix} 3n \\ n \\ -n \\ n \end{pmatrix} \\ &= l \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + m \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + n \begin{pmatrix} 3 \\ 1 \\ -1 \\ 1 \end{pmatrix} \\ &= lu_1 + mu_2 + nu_3 \quad (\text{say}) \end{aligned}$$

$$\begin{aligned} \therefore W &= \{lu_1 + mu_2 + nu_3 | l, m, n \in \mathbb{R}\} \\ &= \text{Span}(\{u_1, u_2, u_3\}) \end{aligned}$$

Thus if $A = [u_1 \ u_2 \ u_3]$ then the columns of A span W , i.e. $\text{Col } A = W$.

Problem 13.16. Find the Null space of A , for $A = \begin{pmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 4 & -2 \end{pmatrix}$

Solution: The null space of A is the solution set of $AX = 0$.
Reduce $[A|0]$ to reduced echelon form

$$\begin{aligned} [A|0] &= \left(\begin{array}{cccc|c} 1 & 3 & 5 & 0 & 0 \\ 0 & 1 & 4 & -2 & 0 \end{array} \right) \\ &\sim \left(\begin{array}{cccc|c} 1 & 0 & -7 & 6 & 0 \\ 0 & 1 & 4 & -2 & 0 \end{array} \right) \end{aligned}$$

Solution is given by

$$x_1 - 7x_3 + 6x_4 = 0, \quad x_2 + 4x_3 - 2x_4 = 0$$

We get

$$x_1 = 7x_3 - 6x_4, \quad x_2 = -4x_3 + 2x_4, \quad x_3 = x_3, \quad x_4 = x_4$$

$$\therefore X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -4 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix} x_4.$$

$$\text{Hence } \text{Nul } A = \text{Span} \left(\left\{ \begin{pmatrix} 7 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right\} \right)$$

Problem 13.17. If $A = \begin{pmatrix} 1 & 4 & 2 \\ 2 & 5 & 1 \\ 3 & 6 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$, does $b \in \text{Nul } A$?

Solution: $b \in \text{Nul } A \Leftrightarrow Ab = 0$

$$\text{Now } Ab = \begin{pmatrix} 1 & 4 & 2 \\ 2 & 5 & 1 \\ 3 & 6 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \therefore b \in \text{Nul } A.$$

Problem 13.18. Describe $\text{Nul } A$ by finding a spanning set for it, where

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 1 & 4 & 9 & -2 \end{pmatrix}$$

Solution: $\text{Nul } A = \{X \in \mathbb{R}^4 | AX = 0\}$. Thus $\text{Nul } A$ is the solution set of

$$AX = 0 \quad (1)$$

Augmented matrix of (1) is

$$\begin{pmatrix} 1 & 2 & 1 & 2 & | & 0 \\ 1 & 4 & 9 & -2 & | & 0 \end{pmatrix} \\ \sim \begin{pmatrix} \mathbf{1} & 0 & -7 & 6 & | & 0 \\ 0 & \mathbf{1} & 4 & -2 & | & 0 \end{pmatrix}$$

Solution of (1) is $x_1 = 7x_3 - 6x_4$, $x_2 = -4x_3 + 2x_4$. x_3, x_4 are free variables.

$$\therefore X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -4 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix} x_4. \quad \text{Thus, Nul } A \text{ is spanned by}$$

$$\begin{pmatrix} 7 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

Problem 13.19. If A and B are sets such that $\text{Span}(A) \subseteq \text{Span}(B)$ then can we say that $A \subseteq B$? Justify your answer.

Problem 13.20. If A and B are the subsets of a vector space V , then

$$\text{Span}(A \cup B) = \text{Span}(A) + \text{Span}(B)$$

Problem 13.21. Let $A = \{1, 2, 3, 4, 5\}$ and $V = \mathcal{P}(A)$, the power set of A . Then V is a vector space over $F = (\mathbb{Z}_2, \oplus_2, \odot_2)$ where $\mathbb{Z}_2 = \{0, 1\}$. Find the subspace generated by S , where

$$(i) S = \{\{1, 5\}, \{2, 3, 4\}\}$$

$$(ii) S = \{\{1, 2\}, \{2, 3\}, \{4, 5\}\}$$

13.7 Exercise

- Let $V(F)$ be a vector space W , a subspace of V and $S \subseteq W$. Prove that $\text{Span}(S)$ is a subspace of W .
- Consider the vector space \mathbb{R} over \mathbb{Q} .
 - Find $\text{Span}(S)$, where
 - $S = \{\sqrt{2}\}$
 - $S = \{2\}$
 - $S = \{1, \sqrt{2}\}$
 - $S = \{1, \sqrt{2}, \sqrt{3}\}$
 - $S = \{\frac{2}{3}, \frac{5}{6}\}$
 - Can you find a finite subset S of \mathbb{R} such that $\text{Span}(S) = \mathbb{R}$?
- Consider the vector space $\mathbb{C}(\mathbb{R})$. Find $\text{Span}(S)$ where
 - $S = \{2\}$
 - $S = \{-1\}$
 - $S = \{i\}$
 - $S = \{1+i\}$
 - $S = \{1, i\}$.

4. Consider $V(\mathbb{R})$, where $V = \mathbb{R}^4$. Let $u = \begin{pmatrix} 2 \\ 0 \\ 3 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 2 \\ 3 \\ 0 \\ 1 \end{pmatrix}$, $w = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}$. Show that $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \text{Span}(\{u, v, w\})$, if $x_1 = 2x_4$, $x_3 = -x_2 + 3x_4$.
5. Consider $V(\mathbb{R})$, where $V = \mathbb{R}[x]$. Prove that $\text{Span}\{1, x, x^2\} = \text{Span}\{1, 1+x, 1+x^2\} = \text{Span}\{1+x, x+x^2, 1+x^2\}$.
6. Consider the vector space $\mathbb{C}(\mathbb{R})$. Show that $\text{Span}\{\frac{1+i\sqrt{3}}{2}, \frac{1-i\sqrt{3}}{2}\} = \text{Span}\{1, \sqrt{3}\}$. Hence, write any element in $\text{Span}\{\frac{1+i\sqrt{3}}{2}, \frac{1-i\sqrt{3}}{2}\}$ in the simplest form.
7. Let V be a vector space over F . If $u, v \in V$, prove that $\text{Span}(\{u, v\}) = \text{Span}(\{u+v, u-v\})$.
8. In the vector space $V(\mathbb{R})$, $V = \mathbb{R}^2$, find two different generating sets for $W = \{(2\alpha + 3\beta, \alpha - \beta) | \alpha, \beta \in \mathbb{R}\}$.
9. Consider the vector space of all 2×2 matrices over \mathbb{R} . Let $W = \left\{ \begin{pmatrix} 0 & b \\ a & 0 \end{pmatrix} | a, b \in \mathbb{R} \right\}$. Find a subset S of W such that $\text{Span}(S) = W$. Hence deduce that W is a subspace of V .
10. Let $W = \left\{ \begin{pmatrix} a+b & 0 \\ a-b & c \end{pmatrix} | a, b, c \in \mathbb{R} \right\}$. Find a subset S of W such that $W = \text{Span}(S)$. Hence deduce that W is a subspace of $V(\mathbb{R})$, where $V = M_2(\mathbb{R})$.
11. Let $W = \{a + (a+b)x + (2a-b)x^2 | a, b \in \mathbb{R}\}$. Prove that W is a subspace of \mathcal{P}_2 by showing that W is the Span of some subset of \mathcal{P}_2 .
12. Let W be the space of all 3×3 scalar matrices over \mathbb{R} . Find a set S of matrices such that $W = \text{Span}(S)$.
13. Let W be the space of all 4×4 diagonal matrices over \mathbb{R} . Find a smallest generating set for W .
14. Let W be the space of all 3×3 scalar matrices over \mathbb{R} . Find a set S of matrices such that $W = \text{Span}(S)$.
15. Let $W = \{A \in \mathcal{M}_2(\mathbb{R}) | \text{tr}(A) = 0\}$. Find a smallest generating set for W .
16. Describe $\text{Span}\{v_1, v_2, v_3\}$ in $\mathbb{R}^5(\mathbb{R})$, where $v_1 = (1, 2, 0, 3, 0)$, $v_2 = (0, 0, 1, 4, 0)$, $v_3 = (0, 0, 0, 0, 1)$. Also show that it contains $(-3, -6, 1, -5, 2)$ but not $(2, 4, 6, 7, 8)$.
17. Describe $\text{Span}\{v_1, v_2, v_3\}$ in \mathbb{R}^4 , where $v_1 = (2, -1, 3, 2)$, $v_2 = (-1, 1, 1, 3)$, $v_3 = (1, 1, 9, -5)$.

18. In $\mathcal{P}_2(\mathbb{R})$, let $p_1(x) = x^2 + x + 1$, $p_2(x) = x^2 + 1$, $p_3(x) = x$. Which of the following polynomials are linear combinations of $p_1(x), p_2(x), p_3(x)$?
- (i) $x^2 + 2x + 1$
 - (ii) $2x^2 - 3x + 3$
 - (iii) $x^2 + 2x + 2$
 - (iv) $3x^2 + 2x + 3$.
19. Find subsets A and B of a vector space V such that
- (i) $\text{Span}(A \cup B) \neq \text{Span}(A) \cup \text{Span}(B)$
 - (ii) $\text{Span}(A \cap B) \neq \text{Span}(A) \cap \text{Span}(B)$
 - (iii) Can you always find a set S such that $\text{Span}(A) \cup \text{Span}(B) = \text{Span}(S)$
20. Let $W = \left\{ \begin{pmatrix} a - b \\ b - c \\ c - a \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}$. Find a subset of W which spans W .
Hence show that W is a subspace of \mathbb{R}^3 .
21. Let $W = \left\{ \begin{pmatrix} a \\ b \\ c \\ a + b + c \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}$. Give a set of vectors u, v, w in W such that $W = \text{Span}(\{u, v, w\})$.
22. Let $W = \left\{ \begin{pmatrix} 2a + b \\ a - b \\ b \\ a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$ is a subspace of \mathbb{R}^4 by showing that $W = \text{Span}(S)$ for some subset S of W .
23. If $A = \begin{pmatrix} -8 & -2 & -9 \\ 6 & 4 & 8 \\ 4 & 0 & 4 \end{pmatrix}$, $u = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$, does $u \in \text{Col}A$?
24. Find $\text{Nul } A$, where $A = \begin{pmatrix} 1 & -2 & -4 & 0 \\ 2 & 0 & -1 & -3 \end{pmatrix}$.
25. Express $(1, -2, 5)$ as a linear combination of $(1 \ 1 \ 1)$, $(1 \ 2 \ 3)$, and $(2 \ -1 \ 1)$.

13.8 Solved Problems

Problem 13.22. Let $V = \mathbb{R}^3$. For $v_1, v_2 \in V$, $\alpha \in \mathbb{R}$, with $v_1 = (x_1, x_2, x_3)$, $v_2 = (y_1, y_2, y_3)$ define

$$v_1 \oplus v_2 = (x_1 + y_1 + 1, x_2 + y_2, x_3 + y_3)$$

$$\alpha \odot v_1 = (\alpha x_1 + \alpha - 1, \alpha x_2, \alpha x_3)$$

Prove that V is a vector space over \mathbb{R} under the given operations.

Solution: Clearly $v_1 + v_2, \alpha v_1 \in V$ for all $v_1, v_2 \in V, \alpha \in \mathbb{R}$. Associative and commutative laws can be verified easily. Also $e = (-1, 0, 0) \in V$ is such that for any $v = (x_1, x_2, x_3) \in V$.

$$\begin{aligned} e \oplus v &= (x_1 - 1 + 1, x_2 + 0, x_3 + 0) \\ &= (x_1, x_2, x_3) = v \end{aligned}$$

Thus e is the zero element of V .

Also the negative of $(x_1, x_2, x_3) \in V$ is $(-x_1 - 2, -x_2, -x_3) \in V$.

Let $u, v \in V, \alpha, \beta \in \mathbb{R}$ and let $u = (x_1, x_2, x_3), v = (y_1, y_2, y_3)$.

$$\begin{aligned} \text{(i)} \quad \alpha \odot (u \oplus v) &= \alpha \odot (x_1 + y_1 + 1, x_2 + y_2, x_3 + y_3) \\ &= (\alpha(x_1 + y_1 + 1) + \alpha - 1, \alpha(x_2 + y_2), \alpha(x_3 + y_3)) \\ &= (\alpha x_1 + \alpha y_1 + 2\alpha - 1, \alpha x_2 + \alpha y_2, \alpha x_3 + \alpha y_3) \end{aligned}$$

$$\begin{aligned} \alpha \odot u \oplus \alpha \odot v &= (\alpha x_1 + \alpha - 1, \alpha x_2, \alpha x_3) \oplus (\alpha y_1 + \alpha - 1, \alpha y_2, \alpha y_3) \\ &= (\alpha x_1 + \alpha - 1 + \alpha y_1 + \alpha - 1 + 1, \alpha x_2 + \alpha y_2, \alpha x_3 + \alpha y_3) \\ &= (\alpha x_1 + \alpha y_1 + 2\alpha - 1, \alpha x_2 + \alpha y_2, \alpha x_3 + \alpha y_3) \end{aligned}$$

Hence $\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v$

$$\begin{aligned} \text{(ii)} \quad \alpha \odot u \oplus \beta \odot u &= (\alpha x_1 + \alpha - 1, \alpha x_2, \alpha x_3) + (\beta x_1 + \beta - 1, \beta x_2, \beta x_3) \\ &= (\alpha x_1 + \alpha - 1 + \beta x_1 + \beta - 1 + 1, \alpha x_2 + \beta x_2, \alpha x_3 + \beta x_3) \\ &= ((\alpha + \beta)x_1 + \alpha + \beta - 1, (\alpha + \beta)x_2, (\alpha + \beta)x_3) \\ &= (\alpha + \beta) \odot u. \end{aligned}$$

Hence

$$(\alpha + \beta) \odot u = \alpha \odot u \oplus \beta \odot u$$

$$\text{(iii)} \quad (\alpha\beta) \odot u = (\alpha\beta x_1 + \alpha\beta - 1, \alpha\beta x_2, \alpha\beta x_3)$$

$$\begin{aligned} \alpha \odot (\beta \odot u) &= \alpha \odot (\beta x_1 + \beta - 1, \beta x_2, \beta x_3) \\ &= (\alpha(\beta x_1 + \beta - 1) + \alpha - 1, \alpha(\beta x_2), \alpha(\beta x_3)) \\ &= (\alpha\beta x_1 + \alpha\beta - 1, \alpha\beta x_2, \alpha\beta x_3) \end{aligned}$$

$$\therefore (\alpha\beta) \odot u = \alpha \odot (\beta \odot u)$$

$$\begin{aligned} \text{(iv)} \quad 1u &= (1x_1 + 1 - 1, 1x_2, x_3) \\ &= (x_1, x_2, x_3) = u \end{aligned}$$

Hence V is a vector space over \mathbb{R} .

Problem 13.23. Prove that the only non-trivial subspaces of \mathbb{R}^3 are the planes and lines through the origin.

Solution: Let W be a subspace of \mathbb{R}^3 . If $W = \{0\}$ then it is trivial subspace, if $W \neq \{0\}$, let $w_1 \neq 0 \in W$.

Consider $W_1 = \{\alpha w_1 | \alpha \in \mathbb{R}\} = \text{Span}(\{w_1\})$. Then W_1 is a subspace of \mathbb{R}^3 . Also W_1 is a line through the origin containing w_1 . If $W_1 = W$, then W is a

line through the origin. If $W_1 \neq W$, then there exist a vector $w_2 \in W$ such that $w_2 \notin W_1$. Clearly w_1, w_2 are not collinear.

Let

$$\begin{aligned} W_2 &= \text{Span}(\{w_1, w_2\}) \\ &= \{\alpha w_1 + \beta w_2 \mid \alpha, \beta \in \mathbb{R}\} \end{aligned}$$

W_2 is a plane determined by the vectors w_1 and w_2 . Since $0 = 0w_1 + 0w_2$, $\therefore W_2$ passes through the origin. If $W_2 = W$, proof is complete. If $W_2 \neq W$, then there exist $w_3 \in W$ such that $w_3 \notin W_2$. Clearly w_1, w_2, w_3 are non coplanar vectors. Then $\text{Span}(\{w_1, w_2, w_3\}) = \mathbb{R}^3$ because any vector in \mathbb{R}^3 can be expressed uniquely as a linear combinations of 3 non-coplanar vectors. This is the trivial subspace. Thus, the only non-trivial subspaces are lines or planes through the origin.

Problem 13.24. Prove that \mathbb{R} is a subspace of $\mathbb{C}(\mathbb{R})$. Does there lie a subspace W of $\mathbb{C}(\mathbb{R})$ such that $\mathbb{R} \subset W \subset \mathbb{C}$?

Solution: Clearly $\mathbb{R} \neq \phi$. Let $u, v \in \mathbb{R}, \alpha \in \mathbb{R}$. Then $\alpha u + v$ is also a real number i.e. $\alpha u + v \in \mathbb{R}$. Hence \mathbb{R} is a subspace of $\mathbb{C}(\mathbb{R})$. Let if possible W be a subspace of $\mathbb{C}(\mathbb{R})$ such that

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

Then, there exist $w \in W \setminus \mathbb{R}$. Let $w = a + ib$, where $a, b \in \mathbb{R}$. Since $w \notin \mathbb{R} \therefore b \neq 0$.

$a \in \mathbb{R} \Rightarrow a \in W \Rightarrow -a \in W$. Then $(-a) + (a + ib) \in W \Rightarrow ib \in W$. Also $0 \neq b \in \mathbb{R} \Rightarrow b^{-1} \in \mathbb{R} \subset W$, therefore $(ib)b^{-1} \in W$ ($\because W$ is subspace), i.e. $i \in W$. Let $x, y \in \mathbb{R}$, then $x + iy \in W$ so that $\mathbb{C} \subset W$. Thus we get $W = \mathbb{C}$.

13.9 Exercise

1. State whether the following statements are true or false.
 - (i) \mathbb{R} is a vector space over \mathbb{Q} .
 - (ii) \mathbb{R} is a vector space over \mathbb{C} .
 - (iii) The set of all polynomials of degree 6, with real coefficients with respect to pointwise addition and scalar multiplication is a vector space over \mathbb{R} .
 - (iv) The set of all 4×4 matrices over \mathbb{R} is a vector space over \mathbb{Q} with respect to matrix addition and scalar multiplication.
 - (v) The set of all 3×3 complex Hermitian matrices over \mathbb{R} is a vector space over \mathbb{Q} with respect to matrix addition and scalar multiplication.
 - (vi) The set of all 2×2 complex skew Hermitian matrices over \mathbb{R} is a vector space over \mathbb{Q} with respect to matrix addition and scalar multiplication.
 - (vii) The set of all matrices of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}, a, b \in \mathbb{C}$ is a vector space over \mathbb{R} with respect to matrix addition and scalar multiplication.

- (viii) The set of all orthogonal matrices is a subspace of the vector space of all 3×3 real matrices over \mathbb{R} .
- (ix) $W = \{(x, y, z) | x, y, z \in \mathbb{Z}\}$ is a subspace of $\mathbb{R}^3(\mathbb{R})$.
- (x) If S and T are non-empty subsets of a vector space V , then $Span(S \cup T) = Span(S) \cup Span(T)$.
2. Prove the following:
- (i) The set of all $m \times n$ matrices over \mathbb{C} is a vector space over \mathbb{C} with respect to matrix addition and scalar multiplication.
- (ii) The set of all $m \times n$ diagonal matrices over \mathbb{R} is a vector space over \mathbb{R} with respect to matrix addition and scalar multiplication.
- (iii) The set of all $m \times n$ matrices over \mathbb{C} is a vector space over \mathbb{R} with respect to matrix addition and scalar multiplication.
- (iv) The set of all polynomials over \mathbb{C} is a vector space over \mathbb{C} with respect to usual addition and scalar multiplication of polynomials.
- (v) The set of all polynomials over \mathbb{R} which vanish at 1 is a vector space over \mathbb{R} .
- (vi) The set of all matrices of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, $a, b \in \mathbb{C}$ is a vector space over \mathbb{C} .
3. Prove that the set in Q2(vi) is a subspace of the vector space of all 2×2 matrices over \mathbb{C} .
4. Prove that the set in Q2(v) is a subspace of the vector space of all polynomials over \mathbb{R} .
5. Let V be the set of all complex valued functions on the real line, such that $f(-x) = \overline{f(x)}$, for all $x \in \mathbb{R}$. Show that V is a vector space over \mathbb{R} , with respect to the operations defined by

$$(f + g)(x) = f(x) + g(x), \quad \forall x \in \mathbb{R} \quad (\alpha f)(x) = \alpha f(x), \quad \forall x \in \mathbb{R}$$

Give an example of a function in V which is not real valued.

6. Prove or disprove the following:
- (i) The set of all $n \times n$ skew Hermitian matrices with complex entries is a vector space over \mathbb{C} with respect to matrix addition and scalar multiplication.
- (ii) \mathbb{R}^2 is a vector space over \mathbb{R} with respect to operations defined as follows:
For $u, v \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$, $u = (x_1, y_1)$, $v = (x_2, y_2)$
- $$u \oplus v = (x_1 + y_1 + 1, x_2 + y_2) \alpha u = (\alpha x_1 + \alpha - 1, \alpha x_2)$$
- (iii) $V = \mathbb{R}^2$ is a vector space over \mathbb{R} with respect to usual addition and scalar multiplication defined as follows:

$$(x, y) = v \in \mathbb{R}^2, \quad \alpha \in \mathbb{R} \quad \alpha v = (0, 0)$$

- (iv) $V = \mathbb{R}$ is a vector space over \mathbb{R} with respect to usual addition and scalar multiplication defined as follows:

For $u, v \in V, \alpha \in \mathbb{R}$

$$u + v = 2u - v \quad \alpha u = \text{usual multiplication of real numbers}$$

Then V is a vector space over \mathbb{R} .

7. Let $V = \mathbb{C}$. Is V a vector space over \mathbb{C} with respect to the usual addition in \mathbb{C} , and scalar multiplication defined as follows:
For $v \in V, \alpha \in \mathbb{C}, \alpha v = (\text{Im } \alpha)v$
8. Let V be a vector space over \mathbb{C} . Then prove that V is also a vector space over \mathbb{R} .
9. Let V be a vector space over \mathbb{C} . Define another system over \mathbb{C} as follows:

$$V_1 = V$$

Addition in V_1 is the same as the addition in V . Scalar multiplication in V_1 is defined as follows:

For $v \in V_1, \alpha \in \mathbb{C} \quad \alpha v = \bar{\alpha}v$.

Is V_1 a vector space over \mathbb{C} ?

10. Let $V_1 = \mathbb{C}(\mathbb{C})$ and $V_2 = \mathbb{C}(\mathbb{R})$. Find a $\text{Span}(A)$ in V_1 and $\text{Span}(A)$ in V_2 , where
- (i) $A = \{1\}$
 - (ii) $A = \{i\}$
 - (iii) $\{1 + 2i\}$
 - (iv) $\{-1\}$.
11. Prove that the only subspaces of \mathbb{R} are \mathbb{R} and the zero subspace.
12. Prove that the only subspaces of \mathbb{R}^2 are \mathbb{R}^2 or the zero subspace, or consists of all scalar multiples of some fixed vector of \mathbb{R}^2 .
13. Describe the subspace of \mathbb{R}^3 .

14. If $V(\mathbb{R})$ is the vector space of all real valued functions on \mathbb{R} . Then, prove that

$$V = W_0 \oplus W_e$$

where W_0 = subspace of all odd functions, W_e = subspace of all even functions.

15. Consider the vector space $V(\mathbb{R})$ where $V = R[x]$, the set of all polynomials in x over \mathbb{R} . Is W a subspace of V , where
- (i) $W = \{p(x) \in V | p(x) = p(-x)\}$
 - (ii) $W = \{p(x) \in V | p(\alpha) = 0\}$ for some $\alpha \in [0, 1]$
 - (iii) $W = \{p(x) \in V | p'(\alpha) = 0\}$.

16. Let W be a subspace of a vector space $V(F)$. If $u, v \in V$ are such that $\alpha u + \beta v \in W$ for some non-zero scalars α, β , then show that either both u , and v belong to W or neither of them belongs to W . Give an example to show that $u \in W, v \notin W$ but $\alpha u + \beta v \in W, 0 \neq \alpha, \beta \in F$ is not possible.

17. Give 3 subspaces of the vector space V over \mathbb{R} , where

(i) $V = \text{space of all polynomials over } \mathbb{R}$.

(ii) $V = \mathbb{C}^n$.

(iii) $V = \text{set of all } 3 \times 4 \text{ matrices over } \mathbb{R}$.

18. Let $V = \mathbb{R}^2$. For any $u, v \in V, \alpha \in \mathbb{R}$, where $u = (x_1, x_2), v = (y_1, y_2)$, define

$$u + v = (x_1 + y_1 + 1, x_2 + y_2 + 1)\alpha u = (\alpha x_1 + \alpha - 1, \alpha x_2 + \alpha - 1)$$

Prove that V is a vector space over \mathbb{R} .

19. Consider the vector space \mathbb{C} over \mathbb{C} . Let $W = \text{Span}\{1 + i\}$. Do $1, i \in W$? What happens when \mathbb{C} is regarded as a vector space over \mathbb{R} ?

20. $V(\mathbb{R})$ is the vector space, where V is the set of all $n \times n$ matrices over \mathbb{R} . Verify whether the following subsets of V are subspaces or not. Give justification for your answer.

(i) $W = \text{set of all upper triangular matrices}$.

(ii) $W = \text{set of all symmetric matrices}$.

(iii) $W = \text{set of all skew symmetric matrices}$.

(iv) $\text{Set of all orthogonal matrices}$.

(v) $\text{Set of all matrices whose trace is zero}$.

21. $V(\mathbb{R})$ is a vector space, where $V = \mathcal{P}_n$, the set of all polynomials over \mathbb{R} of degree at most n . Is W a subspace of V , where

(i) $W = \{\sum_{i=0}^n a_i x^i \in V | a_0 + a_1 + \dots + a_n = 0\}$.

(ii) If α is any real number, and $W = \{p(x) \in V | p(\alpha) = 0\}$.

(iii) If α is any real number, and $W = \{p(x) \in V | p(0) = \alpha\}$.

Let $n \geq 1$ be a fixed integer. Prove that \mathcal{P}_m is a subspace of $\mathcal{P}_n(\mathbb{R})$ for all $0 \leq m \leq n$.

22. Let V be the set of all real valued functions on \mathbb{R} . Then prove that $V(\mathbb{R})$ is a vector space with respect to the usual addition and scalar multiplication of functions. Determine, whether W is a subspace of V , where

(i) $W = \{f \in V | f \text{ is an increasing}\}$.

ii $W = \{f \in V | f \text{ is an even function}\}$.

(iii) $W = \{f \in V | f \text{ is an odd function}\}$.

23. For $a \in \mathbb{R}$, define

$$f_a : \mathbb{R} \rightarrow \mathbb{R} \text{ by } f_a(x) = x + a, x \in \mathbb{R}$$

Let $V = \{f_a | a \in \mathbb{R}\}$. On V , define addition and scalar multiplication as follows: For $f_a, f_b \in V, \alpha \in \mathbb{R}$,

$$f_a + f_b = f_{a+b}, \alpha f_a = f_{\alpha a}$$

Prove that V is a vector space over \mathbb{R} .

24. Let V be a vector space over a field F , and let $S = \{v_1, v_2, \dots, v_m\}$.
 If v_1 is a linear combination of v_2, \dots, v_m , then $\text{Span}(\{v_2, \dots, v_m\}) = \text{Span}\{(v_1, v_2, \dots, v_m)\}$.

13.10 Answers to Exercises

Exercise - 13.2

- 5.
- (i) False; $1 \in \mathbb{Q}$, $\sqrt{2} \in \mathbb{R}$ but $1\sqrt{2} = \sqrt{2} \notin \mathbb{Q}$
 - (ii) False
 - (iii) True
 - (iv) True
 - (v) True
- 6.
- (i) $(V, +)$ is not a group.
 - (ii) $1v = v$ is not satisfied for all $v \in V$.
 - (iii) $(V, +)$ is not a group.
 - (iv) $(\alpha + \beta)x = \alpha x + \beta x$ does not hold
 - (v) $(\alpha + \beta)x = \alpha x + \beta x$ does not hold
7. No, $\because (\alpha\beta)v = \alpha(\beta v)$ is not satisfied.

Exercise - 13.4

1. (i) Yes (ii) Yes (iii) Yes (iv) No (v) No
2. (i) No (ii) Yes (iii) No (iv) No (v) Yes (vi) No
3. (i) No (ii) No (iii) Yes (iv) Yes (v) No (vi) Yes (vii) No (viii) No
4. (i) No (ii) No (iii) No (iv) No (v) Yes
5. (i) Yes (ii) Yes (iii) Yes (iv) No (v) Yes (vi) No (vii) Yes
6. (i) No (ii) No (iii) No (iv) Yes (v) Yes (vi) Yes
7. (i) Yes (ii) Yes (iii) No (iv) Yes (v) Yes (vi) Yes (vii) No
8. (i) Yes (ii) Yes (iii) Yes (iv) No (v) No
9. (i) No (ii) No
12. No. Consider $V(\mathbb{R})$.

Exercise - 13.7

5. *Hint:* If $S = \{1, x, x^2\}$, $T = \{1, 1 + x, 1 + x^2\}$, prove that $S \subseteq \text{Span}(T)$ and $T \subseteq \text{Span}(S)$.
8. $\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\};$
 $\left\{ \begin{pmatrix} 5 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\}$

19. Consider \mathbb{R}^3 over \mathbb{R} . $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$.
- (i) $A = \{e_1\}$, $B = \{e_2\}$. $\text{Span } A = X\text{-axis}$, $\text{Span } B = Y\text{-axis}$
 $\text{Span}(A \cup B) = XY\text{-plane}$.
 - (ii) $A = \{e_1, e_2\}$, $B = \{u, v\}$, $u = (1, 0, 1)$ $v = (1, 0, -1)$
 $\text{Span } A = xy\text{-plane}$, $\text{Span } B = xz\text{-plane}$.
 - (iii) Not always. As in (i) $\text{Span } A \cup \text{Span } B = X\text{-axis} \cup Y\text{-axis}$, which is not a subspace. But $\text{Span}(S)$ is always a subspace.
23. Yes
- $$u = -\frac{1}{2} \begin{pmatrix} -8 \\ 6 \\ 4 \end{pmatrix} + \begin{pmatrix} -2 \\ 4 \\ 0 \end{pmatrix}$$
- Other solutions are also possible.
24. $\text{Nul } A \text{ Span}(u_1, u_2)$, where $u_1 = (2 \ -7 \ 4 \ 0)^t$, $u_2 = (6 \ 3 \ 0 \ 4)^t$
25. $(1 \ -2 \ 5)^t = -6(1 \ 1 \ 1)^t + 3(1 \ 2 \ 3)^t + 2(2 \ -1 \ 1)^t$

Exercise - 13.9

5. $f : \mathbb{R} \rightarrow \mathbb{C}$
 $f(x) = x + ix^2$
 then $f \in \mathbb{V}$, f is not real valued.
7. No. $1v \neq v$
10. (i) \mathbb{C}, \mathbb{R}
 (ii) $\mathbb{C}, \{\alpha i : \alpha \in \mathbb{R}\}$
 (iii) $\mathbb{C}, \{\alpha + 2i\alpha | \alpha \in \mathbb{R}\}$
 (iv) \mathbb{C}, \mathbb{R}

Chapter 14

Basis and Dimension

In the previous chapters we have seen that given a vector space V for example \mathbb{R}^n , we can find a subset S of V such that S spans V . In this chapter we are interested in finding a subset S of V which spans V and no proper subset of S can span V . Such a set is called a minimal spanning set. We will show that S is such a set if no element of S is a linear combinations of the remaining elements.

14.1 Linearly Dependent Sets

In the vector space \mathbb{R}^3 over \mathbb{R} , consider

$$v_1 = (1, 1, 0), v_2 = (1, 0, 1), v_3 = (2, 1, 1), v_4 = (0, 1, 1)$$

Let $S = \{v_1, v_2, v_3, v_4\}$. Then S spans \mathbb{R}^3 . We see that $v_3 = v_1 + v_2$, i.e.

$$v_3 = v_1 + v_2 + 0 \cdot v_4 \tag{14.1}$$

Hence v_3 is a linear combination of v_1, v_2 and v_4 , so that any linear combination of v_1, v_2, v_3, v_4 is also a linear combination of v_1, v_2, v_4 . Thus, we can say that v_3 is not required to span \mathbb{R}^3 . Hence $\text{Span}(S) = \text{Span}(S \setminus \{v_3\})$, so that a proper subset of S spans \mathbb{R}^3 . Observe that in (Eq. 14.1) since the coefficient of v_4 is zero, therefore v_4 can not be expressed as a linear combination v_1, v_2 and v_3 so that v_4 cannot be removed. In fact v_1 is a linear combination of v_2, v_3, v_4 as $v_1 = -v_2 + v_3 + 0 \cdot v_4$ so that v_1 can also be removed from S instead of v_3 i.e.

$$\begin{aligned} \text{Span}(S) &= \text{Span}(S \setminus \{v_1\}) \\ \text{Similarly } \text{Span}(S) &= \text{Span}(S \setminus \{v_2\}). \end{aligned}$$

This leads us to the following definition.

Definition 14.1. (*Linearly dependent vectors*): Let $V(F)$ be a vector space. Vectors v_1, v_2, \dots, v_n of V are said to be linearly dependent over F if there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_n \in F$, not all zero such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

If v_1, v_2, \dots, v_n are not linearly dependent over F , they are said to be linearly independent over F .

That is for any $\alpha_1, \alpha_2, \dots, \alpha_n \in F$,

$$\begin{aligned} \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n &= 0 \\ \implies \alpha_1 = \alpha_2 = \dots = \alpha_n &= 0. \end{aligned}$$

A finite set S of vectors is linearly dependent (linearly independent) over F according as the elements of S are linearly dependent (linearly independent) over F . An infinite set S of vectors is said to be

- (i) linearly independent over F if every finite subset of S is linearly independent over F .
- (ii) linearly dependent over F if some finite non empty subset of S is linearly dependent over F .

Consequently the null set is not linearly dependent, as it does not contain any element. Thus the null set is linearly independent.

Remark 14.1. When F is understood from the context, we simply say S is a linearly independent or dependent set.

Example 14.1. In $\mathbb{R}^3(\mathbb{R})$, consider $v_1 = (1, 2, 3), v_2 = (0, 1, 2), v_3 = (2, 3, 1)$. To check whether v_1, v_2, v_3 are linearly dependent over \mathbb{R} , we solve

$$\begin{aligned} \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 &= 0, \text{ where } \alpha_i \in \mathbb{R} \text{ (14.2)} \\ (\alpha_1 + 2\alpha_3, 2\alpha_1 + \alpha_2 + 3\alpha_3, 3\alpha_1 + 2\alpha_2 + \alpha_3) &= 0 \end{aligned}$$

so that

$$\begin{aligned} \alpha_1 + 2\alpha_3 &= 0 \\ 2\alpha_1 + \alpha_2 + 3\alpha_3 &= 0 \\ 3\alpha_1 + 2\alpha_2 + \alpha_3 &= 0 \end{aligned}$$

Solving for $\alpha_1, \alpha_2, \alpha_3$, we get, $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Since the only solution of (14.2) is the zero solution, therefore v_1, v_2, v_3 are linearly independent over \mathbb{R} .

Example 14.2. In $V(\mathbb{R})$, where $V = \mathbb{P}_3(x)$, consider the vectors

$$\begin{aligned} v_1 &= 5 - 2x + 3x^2 + 10x^3 \\ v_2 &= 7 - 5x + 4x^2 + 20x^3 \\ v_3 &= 4 - 2x + 4x^2 + 7x^3 \\ v_4 &= 10 + 7x - 11x^2 + 13x^3 \end{aligned}$$

Let $\alpha_i \in \mathbb{R}$, $i = 1, 2, 3, 4$, such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = 0 \quad (14.3)$$

Then

$$(5\alpha_1 + 7\alpha_2 + 4\alpha_3 + 10\alpha_4) + (-2\alpha_1 - 5\alpha_2 - 2\alpha_3 + 7\alpha_4)x + (3\alpha_1 + 4\alpha_2 + 4\alpha_3 - 11\alpha_4)x^2 + (10\alpha_1 + 20\alpha_2 + 7\alpha_3 + 13\alpha_4)x^3 = 0$$

so that

$$\begin{aligned} 5\alpha_1 + 7\alpha_2 + 4\alpha_3 + 10\alpha_4 &= 0 \\ -2\alpha_1 - 5\alpha_2 - 2\alpha_3 + 7\alpha_4 &= 0 \\ 3\alpha_1 + 4\alpha_2 + 4\alpha_3 - 11\alpha_4 &= 0 \\ 10\alpha_1 + 20\alpha_2 + 7\alpha_3 + 13\alpha_4 &= 0 \end{aligned}$$

On solving we get $\alpha_1 = 15, \alpha_2 = -3, \alpha_3 = -11, \alpha_4 = -1$.

Since there is a non-zero solution of (Eq. 14.3), the vectors v_1, v_2, v_3, v_4 are linearly dependent over \mathbb{R} .

Example 14.3. In $V(\mathbb{R})$, where $V = \mathbb{R}[x]$, let $S = \{1, x, x^2, \dots\}$. Consider any finite subset of S , namely $\{x^{i_1}, x^{i_2}, \dots, x^{i_n}\}$, where $i_k \in \mathbb{N} \cup \{0\}$, are all distinct. Let $\alpha_j \in \mathbb{R}$ $1 \leq j \leq n$ such that

$$\alpha_1 x^{i_1} + \alpha_2 x^{i_2} + \dots + \alpha_n x^{i_n} = 0$$

Then $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, so that S is linearly independent. As every finite subset of S is linearly independent, $\therefore S$ is linearly independent.

Example 14.4. Let $S = \{1, \sqrt{2}\} = \{v_1, v_2\}$ (say). Then S is linearly dependent over \mathbb{R} as $2, -\sqrt{2} \in \mathbb{R}$ such that

$$2v_1 - \sqrt{2}v_2 = 0$$

We now prove that S is linearly independent over \mathbb{Q} . Let $\alpha_1, \alpha_2 \in \mathbb{Q}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 = 0$$

Then

$$\begin{aligned} \alpha_1 + \sqrt{2}\alpha_2 &= 0 = 0 + \sqrt{2} \cdot 0 \\ \implies \alpha_1 &= \alpha_2 = 0 \end{aligned}$$

Thus S is linearly independent over \mathbb{Q} . This shows that a set of vectors may be linearly independent over one field but may be linearly dependent over another field. Hence the field over which linear independence is being checked is important.

Example 14.5. In $V(\mathbb{R})$, $V = M_2(\mathbb{R})$, consider the set S of all scalar matrices. Then S is an infinite linearly dependent set.

For let $A_1, A_2 \in S$, where $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $S_1 = \{A_1, A_2\}$.

Then S_1 is a finite subset of S and $2A_1 - A_2 = 0$ so that S_1 is a linearly dependent set. Thus S has a finite linearly dependent subset, therefore S is linearly dependent.

The next theorem gives some important facts regarding linearly independent/dependent set of vectors.

Theorem 14.1. In any vector space $V(F)$ the following holds:

- (i) The zero vector is linearly dependent.
- (ii) Any non-zero vector is linearly independent.
- (iii) Every superset of a linearly dependent set is linearly dependent.
- (iv) Every subset of a linearly independent set is linearly independent.
- (v) Any set containing the zero vector is linearly dependent.

Proof:

- (i) If $0 \neq \alpha$ in F and $0 \in V$, then

$$\alpha 0 = 0 \tag{14.4}$$

So that $\{0\}$ is linearly dependent set.

(ii) Let $0 \neq \nu \in V$. Let $S = \{\nu\}$. If $\alpha \in F$ such that

$$\alpha\nu = 0 \quad (14.5)$$

then $\alpha = 0$ or $\nu = 0$ by Theorem (1.1). Since $\nu \neq 0$ therefore $\alpha = 0$. Hence $\{\nu\}$ is a linearly independent set.

(iii) Let S be a linearly dependent set and T be superset of S .

Then S has a finite linearly dependent subset, say U .

Hence $U \subseteq S \subseteq T$, so that U is a finite linearly dependent subset of T .

Thus T is a linearly dependent set.

(iv) Let S be a linearly independent set and T a subset of S . Let, if possible T be a linearly dependent set. By (iii) S is linearly dependent, which is a contradiction. Hence our assumption is wrong so that T is a linearly independent set.

(v) Let S be a set containing the zero vector. Then $\{0\} \subseteq S$. By (i) $\{0\}$ is a linearly dependent set. By (iii) S is a linearly dependent set. \square

Let S be any set. If the elements of S are labeled as $\{s_1, s_2, s_3 \dots\}$, then S is called an indexed set.

The following theorem proves that if the elements of a linearly dependent set S are ordered, then some element of S can be expressed as a linear combination of the preceding ones.

Theorem 14.2. *An indexed set of non-zero vectors $\{v_1, v_2, \dots, v_m\}$, $m \geq 2$ is linearly dependent if and only if some $v_j, j \geq 2$ is linear combination of v_1, v_2, \dots, v_{j-1} .*

Proof: Let $\{v_1, v_2, \dots, v_m\}$, $m \geq 2$ be an indexed set of non zero vectors.

Condition is necessary.

Let $\{v_1, v_2, \dots, v_m\}$ be linearly dependent. Then there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_m$, not all zero such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = 0 \quad (14.6)$$

Let j be the largest suffix such that $\alpha_j \neq 0$. If $j = 1$ then

$$\begin{aligned} \alpha_1 v_1 &= 0 \\ \Rightarrow v_1 &= 0 \quad \text{as } \alpha_1 \neq 0 \end{aligned}$$

a contradiction to the fact that $v_1 \neq 0$. Thus $j \geq 2$.

(14.6) gives

$$\begin{aligned} \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_j v_j &= 0 \\ \Rightarrow \alpha_j v_j &= -(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{j-1} v_{j-1}) \\ \Rightarrow v_j &= -\alpha_j^{-1}(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{j-1} v_{j-1}) \end{aligned}$$

Thus v_j is a linear combination of $\{v_1, v_2, \dots, v_{j-1}\}$ for some $j \geq 2$.

Condition is sufficient.

Let the condition hold. Then there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_{j-1}$ such that

$$v_j = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_{j-1} v_{j-1}$$

$$\therefore \alpha_1 v_1 + \dots + \alpha_{j-1} v_{j-1} + (-1)v_j + 0v_{j+1} + \dots + 0v_m = 0$$

$$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = 0$$

Since $\alpha_j = -1 \neq 0$, $\therefore v_1, v_2, \dots, v_m$ are linearly dependent. \square

Corollary 14.3. *A subset S of non-zero vectors of V is linearly dependent if and only if $\text{Span}(S) = \text{Span}(S \setminus \{v\})$, for some $v \in S$.*

Proof: Let S be linearly dependent. Then some finite subset S_1 of S is linearly dependent.

Let $S_1 = \{v_1, v_2, \dots, v_m\}$ be indexed. Since S_1 is linearly dependent, therefore S_1 has at least two elements so that $m \geq 2$. By the above theorem, there exists some $j \geq 2$ such that v_j is a linear combination of v_1, v_2, \dots, v_{j-1} . Hence v_j is a linear combination of a finite number of elements of S , so that

$$\text{Span}(S) = \text{Span}(S \setminus \{v_j\})$$

Conversely, suppose there exists $v \in S$ such that $\text{Span}(S) = \text{Span}(S \setminus \{v\})$

$$\begin{aligned} v \in S &\Rightarrow v \in \text{Span}(S) \\ &\Rightarrow v \in \text{Span}(S \setminus \{v\}) \end{aligned}$$

$\Rightarrow v$ is linear combination of a finite number of elements of $\text{Span}(S \setminus \{v\})$, say v_1, v_2, \dots, v_k

$\Rightarrow \exists$ scalars $\alpha_1, \alpha_2, \dots, \alpha_k$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k$$

$$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k + (-1)v = 0$$

$$\Rightarrow S_1 = \{v_1, v_2, \dots, v_k, v\} \text{ is linearly dependent.}$$

$\Rightarrow S$ is linearly dependent, as S has a finite linearly dependent subset S_1 . \square

Remark 14.2. *If a vector v is such that $\text{Span}(S) = \text{Span}(S \setminus \{v\})$ then v is called a redundant vector in S .*

Example 14.6. 1. In \mathbb{R}^4 consider $S = \{u_1, u_2, u_3, u_4\}$ where $u_1 = (1, 1, 1, 1)$, $u_2 = (1, 0, 1, 0)$, $u_3 = (-1, 1, -1, 1)$, $u_4 = (0, 0, 1, 1)$. It can be easily seen that

$$u_1 - 2u_2 - u_3 + 0u_4 = 0 \tag{14.7}$$

so that S is a linearly dependent set. The last non-zero coefficient in (14.7) is that of u_3 . Thus (14.7) can be written as

$$u_1 - 2u_2 - u_3 = 0$$

$$\Rightarrow u_3 = u_1 - 2u_2$$

$\Rightarrow u_3$ is a linear combination of u_1, u_2 .

Hence $\text{Span}(S) = \text{Span}(S \setminus \{u_3\})$.

If S is any set, then we are looking for some subset S_1 of S such that $\text{Span}(S) = \text{Span}(S_1)$ and S_1 contains no redundant vectors. This is done by weeding out all redundant vectors from S .

Theorem 14.4. If v_1, v_2, \dots, v_n are non-zero vectors in a vector space V and $S = \{v_1, v_2, \dots, v_n\}$, then

$$\text{Span}(S) = \text{Span}(S \setminus \{v_i\})$$

if and only if v_i is a linear combination of $v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n$.

Proof: Condition is necessary

$v_i \in S \Rightarrow v_i \in \text{Span}(S) = \text{Span}(S \setminus \{v_i\}) \Rightarrow v_i$ is a linear combination of $v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n$.

Condition is sufficient

Let v_i be a linear combination of $v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. Let $v \in \text{Span}(S)$. $\therefore v$ is a linear combination of v_1, v_2, \dots, v_n . But v_i is a linear combination of $v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. Hence v is a linear combination of $v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. So that

$$\text{Span}(S) \subseteq \text{Span}(S \setminus \{v_i\}) \quad (14.8)$$

Since $S \setminus \{v_i\} \subseteq S$

$$\therefore \text{Span}(S \setminus \{v_i\}) \subseteq \text{Span}(S) \quad (14.9)$$

From (14.8) and (14.9)

$$\text{Span}(S) = \text{Span}(S \setminus \{v_i\}).$$

□

The following table summarizes the problems of vector space studied in this section, for the vector space \mathbb{R}^m , their equivalent problem in system of linear equations and its solution in terms of the echelon form of a suitable matrix.

If $v_1, v_2, \dots, v_n \in \mathbb{R}^m$, let $S = \{v_1, v_2, \dots, v_n\}$ and $A = [v_1, v_2, \dots, v_n]$

	Problem in Vector Space \mathbb{R}^m	Equivalent problem in System of linear equations	Solution
1.	S is linearly independent	$AX = 0$ has only trivial solution	Every column of A is pivot column or $ A \neq 0$ if A is a square matrix.
2.	S is linearly dependent	$AX = 0$ has non-trivial solutions	A has non-pivot columns or $ A = 0$ if A is a square matrix.
3.	$b \in \text{Span}(S)$	$AX = b$ has a solution	In $[A : b]$, the augmented column does not have a pivot.
4.	if $b \in \text{Span}(S)$, find α_i 's $\in F$, $1 \leq i \leq n$ such that $b = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$	To find solution of $AX = b$	In $[A : b]$, the augmented column is not a pivot column. A solution of $Ax = b$ is $(\alpha_1, \alpha_2, \dots, \alpha_n)^t$.
5.	$\text{Span}(S) = \mathbb{R}^m$	$AX = b$ has a solution for every $b \in \mathbb{R}^m$	Every row of A has a pivot.
6.	T is a smallest subset of S such that $\text{Span}(T) = \text{Span}(S)$		T is the set of pivot columns of A .

14.2 Solved Problems

Problem 14.1. Show that the vectors $(3, 0, 0, 0)$, $(3, -1, 1, 0)$, $(5, -1, 1, 3)$, $(6, 0, 1, 3)$ are linearly independent over \mathbb{R} .

Solution:

Let $v_1 = (3, 0, 0, 0)$, $v_2 = (3, -1, 1, 0)$, $v_3 = (5, -1, 1, 3)$, $v_4 = (6, 0, 1, 3)$. v_1, v_2, v_3, v_4 are linearly independent if and only if for $\alpha_i \in \mathbb{R}$, $1 \leq i \leq 4$, the vector equation

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = 0 \quad (14.10)$$

has only the trivial solution. The corresponding augmented matrix is

$$\begin{aligned} [A \mid 0] &= [\ v_1 \ v_2 \ v_3 \ v_4 \ \vdots \ 0 \] \\ &= \begin{pmatrix} 3 & 3 & 5 & 6 & \vdots & 0 \\ 0 & -1 & -1 & 0 & \vdots & 0 \\ 0 & 1 & 1 & 1 & \vdots & 0 \\ 0 & 0 & 3 & 3 & \vdots & 0 \end{pmatrix} \\ &\sim \begin{pmatrix} 3 & 3 & 5 & 6 & \vdots & 0 \\ 0 & -1 & -1 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 1 & \vdots & 0 \\ 0 & 0 & 3 & 3 & \vdots & 0 \end{pmatrix} \text{ applying } R_3 \rightarrow R_3 + R_2 \\ &\sim \begin{pmatrix} \mathbf{3} & 3 & 5 & 6 & \vdots & 0 \\ 0 & \mathbf{-1} & -1 & 0 & \vdots & 0 \\ 0 & 0 & \mathbf{3} & 3 & \vdots & 0 \\ 0 & 0 & 0 & \mathbf{1} & \vdots & 0 \end{pmatrix} \text{ applying } R_3 \leftrightarrow R_4 \end{aligned}$$

Since every column has a pivot, \therefore the corresponding vector equation has a unique solution, namely the trivial solution.

$\therefore v_1, v_2, v_3, v_4$ are linearly independent over \mathbb{R} .

Problem 14.2. In $V(\mathbb{R})$, where $V = P_3(x)$, let

$$\begin{aligned} v_1 &= 1 + x + x^3 \\ v_2 &= 1 + x^2 - x^3 \\ v_3 &= x + x^2 + x^3 \\ v_4 &= 1 + 2x + 3x^2 \end{aligned}$$

Prove that v_1, v_2, v_3, v_4 are linearly independent.

Solution:

Let $\alpha_i \in \mathbb{R}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = 0$$

Then

$$(\alpha_1 + \alpha_2 + \alpha_4) + (\alpha_1 + \alpha_3 + 2\alpha_4)x + (\alpha_2 + \alpha_3 + 3\alpha_4)x^2 + (\alpha_1 - \alpha_2 + \alpha_3)x^3 = 0$$

so that,

$$\begin{aligned}\alpha_1 + \alpha_2 + \alpha_4 &= 0 \\ \alpha_1 + 2\alpha_4 + \alpha_3 &= 0 \\ \alpha_2 + \alpha_3 + \alpha_4 &= 0 \\ \alpha_1 - \alpha_2 + \alpha_3 &= 0\end{aligned}$$

Solving, we get

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$$

Hence the vectors v_1, v_2, v_3, v_4 are linearly independent.

Problem 14.3. Check the vectors $e^{2x} \sin x, e^{2x} \cos x$ for linear independence over \mathbb{R} .

Solution:

Let $v_1 = e^{2x} \sin x, v_2 = e^{2x} \cos x$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 = 0 \tag{14.11}$$

ie.

$$\begin{aligned}\alpha_1 e^{2x} \sin x + \alpha_2 e^{2x} \cos x &= 0 \\ \implies e^{2x}(\alpha_1 \sin x + \alpha_2 \cos x) &= 0 \\ \implies \alpha_1 \sin x + \alpha_2 \cos x &= 0 \quad \because e^{2x} \neq 0\end{aligned} \tag{14.12}$$

Differentiating with respect to x , we get

$$\alpha_1 \cos x - \alpha_2 \sin x = 0 \tag{14.13}$$

We are required to solve (14.12) and (14.13) for α_1 and α_2 .

Since

$$\begin{vmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{vmatrix} \neq 0$$

\therefore (14.12) and (14.13) have a unique solution, namely the zero solution.

$\therefore \alpha_1 = \alpha_2 = 0$.

Hence v_1, v_2 are linearly independent.

Problem 14.4. For what values of a is the set

$$\{(1, 1, 1 + a), (2, 2 + a, 2 + a), (3 + a, 3 + a, 3 + a)\}$$

linearly independent?

Solution:

Let $v_1 = (1, 1, 1+a)^t$, $v_2 = (2, 2+a, 2+a)^t$, $v_3 = (3+a, 3+a, 3+a)^t$. Let $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0 \quad (14.14)$$

We know that 3 vectors in \mathbb{R}^3 are linearly independent if and only if

$$\begin{aligned} | [v_1 \ v_2 \ v_3] | &\neq 0 \\ | [v_1 \ v_2 \ v_3] | &= \begin{vmatrix} 1 & 2 & 3+a \\ 1 & 2+a & 3+a \\ 1+a & 2+a & 3+a \end{vmatrix} \\ &= \begin{vmatrix} 1 & 0 & a \\ 1 & a & a \\ 1+a & -a & -2a \end{vmatrix} [C_2 \rightarrow C_2 - 2C_1, \ C_3 \rightarrow C_3 - 3C_1] \\ &= \begin{vmatrix} 1 & 0 & a \\ 0 & a & 0 \\ 1+a & -a & -2a \end{vmatrix} [R_2 \rightarrow R_2 - R_1] \\ &= \begin{vmatrix} 1 & 0 & a \\ 0 & a & 0 \\ 3+a & -a & -0 \end{vmatrix} [R_3 \rightarrow R_3 + 2R_1] \\ &= -a^2(3+a) = 0, \text{ if } a = 0 \text{ or } a = -3. \end{aligned}$$

Thus if $a \neq 0, -3$, then v_1, v_2, v_3 are linearly independent.

Remark 14.3. This method of showing that a system of vectors is linearly independent is used when the number of vectors is same as the number of components of the vector.

Problem 14.5. Check for the linear independence the polynomials $i + x + x^2, -(1+i) - 2x + 2ix^2, x - x^2$ over

- (a) \mathbb{C}
(b) \mathbb{R}

Solution:

Let $p_1 = i + x + x^2, p_2 = -(1+i) - 2x + 2ix^2, p_3 = x - x^2$

(a) Let $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$ such that

$$\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 = 0$$

Then

$$(i\alpha_1 - (1+i)\alpha_2) + (\alpha_1 - 2\alpha_2 + \alpha_3)x + (\alpha_1 + 2i\alpha_2 - \alpha_3)x^2 = 0$$

so that

$$\begin{aligned} i\alpha_1 - (1+i)\alpha_2 &= 0 \\ \alpha_1 - 2\alpha_2 + \alpha_3 &= 0 \\ \alpha_1 + 2i\alpha_2 - \alpha_3 &= 0 \end{aligned}$$

The corresponding coefficient matrix is

$$\begin{aligned} A &= \begin{pmatrix} i & -(1+i) & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{pmatrix} \\ &\sim \begin{pmatrix} i & -(1+i) & 0 \\ 0 & -(1+i) & 1 \\ 0 & (1+i) & -1 \end{pmatrix} \text{ (applying } R_2 \rightarrow R_2 + iR_1, R_3 \rightarrow R_3 + iR_1) \\ &\sim \begin{pmatrix} i & -(1+i) & 0 \\ 0 & -(1+i) & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ (applying } R_3 \rightarrow R_3 + R_2) \end{aligned}$$

Since the third column is not a pivot column so a non-zero solution exists. Thus the system is equivalent to

$$\begin{aligned} i\alpha_1 - (1+i)\alpha_2 &= 0 \\ -(1+i)\alpha_2 + \alpha_3 &= 0 \end{aligned}$$

Hence

$$\begin{aligned} \alpha_1 &= (1-i)\alpha_2 \\ \alpha_3 &= (1+i)\alpha_2 \end{aligned}$$

Hence the vectors p_1, p_2, p_3 are linearly dependent over \mathbb{C} . Taking $\alpha_2 = 1$, we get

$$\alpha_1 = (1-i), \alpha_3 = (1+i)$$

Thus $(1-i)p_1 + p_2 + (1+i)p_3 = 0$.

(ii) Since $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$ and not in \mathbb{R} ,

$$\therefore \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 = 0$$

has no solution in \mathbb{R} . Therefore p_1, p_2, p_3 are linearly independent over \mathbb{R} .

Problem 14.6. Let $S = \{(1, 2), (2, 1), (1, 0), (5, 3)\}$. Find a smallest subset T of S such that $\text{Span}(S) = \text{Span}(T)$.

Solution: Let $v_1 = (1, 2), v_2 = (2, 1), v_3 = (1, 0), v_4 = (5, 3)$.

Then $S = \{v_1, v_2, v_3, v_4\}$.

We remove all those vectors one by one from S which are linearly dependent on the remaining ones.

Step 1 First we check S for linear independence. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = 0 \tag{14.15}$$

The coefficient matrix of vector equation (14.15) is

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & 1 & 5 \\ 2 & 1 & 0 & 3 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 2 & 1 & 5 \\ 0 & -3 & -2 & -7 \end{pmatrix} \text{ (applying } R_2 \rightarrow R_2 - 2R_1). \end{aligned}$$

The first and second columns of A are the pivot columns. Hence $T = \{v_1, v_2\}$ is the smallest subset S such that $\text{Span}(S) = \text{Span}(T)$.

Problem 14.7. Let $S = \{v_1, v_2, v_3, v_4\}$ where $v_1 = \begin{pmatrix} -6 \\ 4 \\ -9 \\ 4 \end{pmatrix}$,

$$v_2 = \begin{pmatrix} 8 \\ -3 \\ 7 \\ -3 \end{pmatrix}, v_3 = \begin{pmatrix} -9 \\ 5 \\ -8 \\ 3 \end{pmatrix}, v_4 = \begin{pmatrix} 4 \\ 7 \\ -8 \\ 3 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 1 \\ -2 \\ 1 \end{pmatrix}.$$

Does $b \in \text{Span}(S)$? If yes, express b as a linear combination of elements of S .

Solution:

$$\text{Let } A = [v_1 \ v_2 \ v_3 \ v_4], \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Step 1 $b \in \text{Span}(S)$ if the matrix equation

$$AX = b \tag{14.16}$$

has a solution.

i.e. In $[A \dot{:} b]$, the augmented column does not have pivot.

Step 2 Reducing $[A \dot{:} b]$ to echelon form. We get

$$[A \dot{:} b] \sim \begin{pmatrix} \mathbf{-1} & \mathbf{1} & \mathbf{0} & \mathbf{2} & \vdots & \mathbf{-1} \\ \mathbf{0} & \mathbf{1} & \mathbf{5} & \mathbf{15} & \vdots & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \vdots & \mathbf{0} \end{pmatrix} = [B \dot{:} c] \quad (\text{say})$$

The pivots are in bold. Since the augmented column does not have a pivot, therefore (14.16) has a solution. Hence $b \in \text{Span}(S)$.

Step 3 To express b as a linear combination of elements of S we find a solution of Eq. 14.16. For this obtain the reduced echelon form of $[A \dot{:} b]$. Thus

$$[A \dot{:} b] \sim \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{-2} & \vdots & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{5} & \vdots & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \vdots & \mathbf{0} \end{pmatrix}$$

Hence a solution of $AX = b$ is given by

$$\begin{aligned} x_1 - 2x_4 &= 1 \\ x_2 + 5x_4 &= 1 \\ x_3 + 2x_4 &= 0 \\ 0 &= 0 \end{aligned}$$

This gives x_4 to be a free variable. Taking $x_4 = 0$, a solution is $x_1 = x_2 = 1, x_3 = 0$.

$$\begin{aligned} \therefore \quad b &= x_1v_1 + x_2v_2 + x_3v_3 + x_4v_4 \\ b &= v_1 + v_2 + 0v_3 + 0v_4. \end{aligned}$$

Problem 14.8. Let $S = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} \right\}$.

Show that $\text{Span}(S) = \mathbb{R}^4$.

Solution:

$$\text{Let } v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_4 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}, v_5 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}$$

and $A = [v_1 \ v_2 \ v_3 \ v_4 \ v_5]$ $X = (x_1, x_2, x_3, x_4)^t$.

$\text{Span}(S) = \mathbb{R}^4$ if $AX = b$ has a solution for every $b \in \mathbb{R}^4$. For this every row of A must have a pivot. Reducing A to echelon form we get

$$A \sim \begin{pmatrix} \mathbf{1} & 1 & 0 & 2 & 0 \\ 0 & \mathbf{1} & 1 & 1 & -1 \\ 0 & 0 & \mathbf{2} & 3 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 \end{pmatrix}$$

The pivots are in bold. Since every row has a pivot

$$\therefore \quad \text{Span}(S) = \mathbb{R}^4.$$

Problem 14.9. Let $S = \{(1, -1, 0, 2)^t, (3, -1, 2, 1)^t, (1, 0, 0, 1)^t\}$. Find the condition so that $(a, b, c, d)^t \in \text{Span}(S)$.

Solution:

Let $v_1 = (1, -1, 0, 2)^t, v_2 = (3, -1, 2, 1)^t, v_3 = (1, 0, 0, 1)^t, u = (a, b, c, d)^t$ and $A = [v_1 \ v_2 \ v_3]$.

$u \in \text{Span}(S)$ if $AX = u$ has a solution.

i.e., in $[A:u]$, the augmented column is not a pivot column.

Reducing $[A:u]$ to echelon form, we get

$$[A:u] \sim \begin{pmatrix} 1 & 3 & 1 & \vdots & a \\ 0 & 2 & 1 & \vdots & a+b \\ 0 & 0 & 1 & \vdots & a+b-c \\ 0 & 0 & 0 & \vdots & 10a-4b-3c-2d \end{pmatrix}$$

The augmented column is not a pivot column if

$$10a - 4b - 3c - 2d = 0$$

This is the required condition.

14.3 Exercise

- Find which of the following sets of vectors are linearly independent over \mathbb{R} :
 - $\{(1, 3, -3), (1, 4, -3), (2, 1, -1), (2, 0, -1)\}$.
 - $\{(1, 3, 1, 0), (-2, -6, 2, 1), (4, 2, 8, 3), (2, 4, 6, -3)\}$.
 - $\{(1, 1, 1, 1), (1, 2, 1, 2), (2, 1, 2, 1), (1, 2, 3, 4)\}$.
 - $\{(2, 0, 3), (2, 1, 3), (0, -4, 0)\}$.
- Determine whether the given set S is linearly independent or dependent over the given field.
 - $S = \{1 + i, 2 + 3i\}$ over \mathbb{C} .
 - S as in (i) over \mathbb{R} .
 - $S = \{1 + x^2, 2x, 1 + x + x^2\}$ over \mathbb{R} .
 - $S = \{(0, 0, 4, 2), (2, 0, 2, 0), (5, 2, -11, -6), (-1, 3, 11, 9)\}$ over \mathbb{R} .
 - $S = \{\sin x, \cos x\}$ over \mathbb{R} .
 - $S = \{e^x \sin x, e^x \cos x\}$ over \mathbb{R} .
 - If $A = \{1, 2, 3\}$, then $S = \{\{1\}, \{2\}, \{1, 2\}\}$ over \mathbb{Z}_2 , where $S \subseteq P(A)$.
 - $S = \{3i + (1 + i)x^2, 1 + ix + x^2, 2i + x + x^2\}$ over \mathbb{R} .
 - S as in (viii) over \mathbb{C} .
- For the following set S , does there exist a proper subset T of S such that $\text{Span}(T) = \text{Span}(S)$? Find a smallest such T .
 - $S = \{(0, 1, 2), (1, 2, 3), (2, 1, 4), (-1, -1, 2)\}$ in $\mathbb{R}^3(\mathbb{R})$.
 - $S = \{1 + x + 3x^3, 1 + x^2 + 2x^3, -2 - 2x - 6x^3, x, -x^2 + x^3\}$ in $P_3(\mathbb{R})$.
 - $S = \{\sin x, \cos x\}$ in the space of all real valued continuous functions.
 - $S = \{2 + \sqrt{3}, 2 - \sqrt{3}, -2 + \sqrt{3}, -2 - \sqrt{3}\}$ in $\mathbb{R}(\mathbb{Q})$.
 - $S = \{1 + \sqrt{2}, 1 - \sqrt{2}, 1 + i\sqrt{3}, 1 - i\sqrt{3}\}$ in $\mathbb{C}(\mathbb{Q})$.
 - $S = \{(2, 3), (1, 1), (0, 1), (1, 2)\}$ in $\mathbb{R}^2(\mathbb{R})$.
 - $S = \{(1, 2, 3, 4), (1, 2, 1, 2), (1, 1, 1, 1), (2, 1, 2, 1)\}$ in $\mathbb{R}^4(\mathbb{R})$.
- Give 3 examples of linearly dependent subsets S of $V(F)$ in each of the following cases:
 - $\mathbb{R}^2(\mathbb{R})$, S contains 2 elements.
 - $\mathbb{R}^3(\mathbb{R})$, S contains 2 elements.
 - $\mathbb{R}^3(\mathbb{R})$, S contains 3 elements.
 - $\mathbb{R}^4(\mathbb{R})$, S contains 3 elements.
 - $V(\mathbb{R})$, where $V = M_2(\mathbb{R})$, S contains 4 elements.
 - $V(\mathbb{C})$, where $V = M_2(\mathbb{C})$, S contains 4 elements.
- Give 3 examples of linearly independent subsets S of $V(F)$, in each of the following cases:
 - $V = \mathbb{R}^2$, $F = \mathbb{R}$, S contains 2 elements.
 - $V = \mathbb{R}^3$, $F = \mathbb{R}$, S contains 3 elements.
 - $V = \mathbb{P}_3(\mathbb{R})$, $F = \mathbb{R}$, S contains 4 elements.
 - $V = M_2(\mathbb{R})$, $F = \mathbb{R}$, S contains 4 elements.
 - $V = \mathcal{C}[0, 1]$, $F = \mathbb{R}$, S contains 2 elements.
 - $V = \mathcal{P}(1, 2, 3)$, $F = \mathbb{Z}_2$, S contains 3 elements.

6. $v_1 = (1, 2, 1, 0)^t, v_2 = (1, 1, -1, 0)^t, v_3 = (1, 1, 0, 0)^t, v_4 = (2, 3, 0, 0)^t$.
Prove that $(x_1, x_2, x_3, x_4)^t \in \mathbb{R}^4$ is a linear combination of v_1, v_2, v_3, v_4 if and only if $x_4 = 0$.
7. $u = (1, 3, 2)^t, v = (-2, 4, 3)^t$. Find the condition that $(x, y, z)^t$ lies in $\text{Span}(\{u, v\})$. Interpret the problem geometrically.
8. Which of the following polynomials are linear combinations of $\{p_1, p_2, p_3\}$ where $p_1 = x^2 + x + 1, p_2 = x^2 + 1, p_3 = x$?
- $x^2 + 2x + 1$.
 - $2x^2 - 3x + 3$.
 - $x^2 + 2x + 2$.
 - $3x^2 + 3x + 2$.
9. Prove that $1, 1 + x, (1 + x)^2$ span $\mathbb{P}_2(\mathbb{R})$.
10. (i) Do $x^3 + 2x + 1, x^2 - x + 2, x^3 + 2, -x^3 + x^2 - 5x + 2$ span P_3 ?
(ii) Does $S = \{2 - t, 1 - 2t, 2 - 4t, -1 + t - t^2, 1 + 2t + t^2\}$ span P_2 ?
11. If u_1, u_2, u_3 are linearly independent vectors, then prove that $u_1 + u_2, u_2 + u_3, u_3 + u_1$ are also linearly independent. What can you say about $u_1 - u_2, u_2 + u_3, u_3 + u_1$?
12. Prove the following:
- The union of two linearly dependent sets is linearly dependent.
 - The intersection of two linearly independent sets is linearly independent.
 - The intersection of a linearly independent set and a linearly dependent set is linearly independent.
 - The union of a linearly independent set and a linearly dependent set is linearly dependent.
13. If v_1, v_2, \dots, v_n are linearly independent vectors of a vector space $V(F)$ and $u = v_1 + v_2 + \dots + v_n$, then prove that the vectors $u - v_1, u - v_2, \dots, u - v_n$ are linearly independent.
14. Test whether the vector $v = (-1, -1, 7)^t$ belongs to the $\text{Span}(S)$, where $S = \{(1, 2, 6)^t, (-3, 1, -7)^t, (1, -4, 8)^t\}$. If $v \in \text{Span}(S)$, express v as a linear combination of elements of S .
15. The set $\mathcal{B} = \{v_1, v_2, v_3\}$ where, $v_1 = (-1, 1, 0), v_2 = (1, 2, -1), v_3 = (0, 1, 0)$ and $v = (-1, 8, -2)$. Show that $v \in \text{Span}(\mathcal{B})$. Also express v as a linear combination of elements of \mathcal{B} .
16. Given that $\mathcal{B} = \{p_1, p_2, p_3\}$ where, $p_1 = 2x^2 + x, p_2 = x^2 + 3, p_3 = x$ spans P_2 . Let $p = 8x^2 - 4x + 6$ and $q = 7x^2 - x + 9$. Express p and q as a linear combinations of elements of \mathcal{B} .
17. Let $S = \left\{ A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$
Show that
- S is linearly independent.
 - S spans $M_2(\mathbb{R})$
- Also express $P = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, as a linear combination of elements of S .

18. Show that the vectors $\{v_1, v_2, v_3\}$ are linearly dependent, where $v_1 = (0, 2, -4)^t$, $v_2 = (1, -2, -1)^t$, $v_3 = (1, -4, 3)^t$. Also express one of them in terms of the other.

14.4 Basis of Vector Space

We study the structure of a vector space V by determining a smallest subset of V that describes V completely. In this section we shall find this smallest subset of V .

Definition 14.2. (Basis): A subset S of a vector space V is said to be a basis for V if

- (i) S spans V .
(ii) S is linearly independent.

Remark 14.4.

- (i) If a subset of S forms a basis for a vector space V , then the vectors in S must be distinct and non-zero, otherwise S will be linearly dependent.
(ii) S may be finite or infinite.

Examples

- Let $e_1 = (1, 0)$, $e_2 = (0, 1)$. Then $S = \{e_1, e_2\}$ forms a basis for $\mathbb{R}^2(\mathbb{R})$.
- Let $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 1)$. Then $S = \{e_1, e_2, \dots, e_n\}$ forms a basis for $\mathbb{R}^n(\mathbb{R})$. S is called the standard basis for \mathbb{R}^n .
- $S = \{1, x, x^2, \dots, x^n\}$ is a basis for P_n , the set of all polynomials of degree n or less, over \mathbb{R} . S is called a standard basis for P_n .
- $S = \{\sin x, \cos x\}$ forms a basis of the vector space of all solutions of the equation $\frac{d^2y}{dx^2} + y = 0$.
- $\{1, x, x^2, \dots\}$ is a basis of $\mathbb{R}[x]$ the space of all polynomials over \mathbb{R} . This is the standard basis for $\mathbb{R}[x]$.
- Let $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.
Then $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ is a standard basis for $M_2(\mathbb{R})$.

The following questions come to our mind.

- Does a basis always exist?
- In case a basis exists, is it unique?

These questions will be answered later in this chapter.

The following theorem leads to a characterization of a basis.

Theorem 14.5. If $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ is a basis for a vector space $V(F)$, then every vector in V is expressible uniquely as a linear combination of elements of \mathcal{B} .

Proof: Since \mathcal{B} is a basis of V , therefore \mathcal{B} spans V . Thus every vector in V can be written as a linear combination of elements of \mathcal{B} .

Let $v \in V$ be expressible as a linear combination of elements of \mathcal{B} in two different ways. Hence there exists, $\alpha_i, \beta_i \in F, 1 \leq i \leq n$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \quad (14.17)$$

$$v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n \quad (14.18)$$

Subtracting (14.18) from (14.17) we get,

$$0 = (\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n$$

Since v_1, v_2, \dots, v_n are linearly independent

$$\therefore \alpha_i - \beta_i = 0, 1 \leq i \leq n$$

$$\text{i.e. } \alpha_i = \beta_i, 1 \leq i \leq n$$

Hence the expressions (14.17) and (14.18) are identical. Thus v is expressible uniquely as a linear combination of elements of \mathcal{B} . \square

The converse of the above theorem is also true.

Theorem 14.6. *If \mathcal{B} is a subset of a vector space $V(F)$ such that every element of V is expressible uniquely as a linear combination of \mathcal{B} , then \mathcal{B} is basis of V .*

Proof: Since every vector v of V can be expressed as a linear combination of elements of \mathcal{B} , therefore \mathcal{B} spans V .

We now prove that \mathcal{B} is linearly independent. Let S be any finite subset of \mathcal{B} . Let $S = \{v_1, v_2, \dots, v_n\}$.

Let $\alpha_1 \alpha_2 \dots \alpha_n \in F$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 \quad (14.19)$$

$$\text{Also } 0v_1 + 0v_2 + \cdots + 0v_n = 0 \quad (14.20)$$

Thus $0 \in V$ has two expressions as a linear combination of elements of \mathcal{B} , namely (14.19) and (14.20). By uniqueness of the expression, we must have

$$\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_n = 0$$

Thus S is linearly independent. However S is any finite subset of \mathcal{B} , therefore every finite subset of \mathcal{B} is linearly independent. This proves that

\mathcal{B} is linearly independent. Hence \mathcal{B} is a basis of V . \square

The above two theorems can be combined to give a characterization of a basis.

Theorem 14.7. *A subset \mathcal{B} of a vector space $V(F)$ is a basis of V if and only if every element of V can be expressed uniquely as a linear combination of elements of \mathcal{B} .*

Examples

1. Consider the vector space \mathbb{R}^3 . Let $v_1 = (1, 0, 1), v_2 = (0, 1, 1), v_3 = (1, 1, 2), v = (1, 2, 3) \in \mathbb{R}^3$, and

$$v = v_1 + 2v_2 + 0v_3$$

Also

$$v = 0v_1 + v_2 + v_3$$

Thus v has two different representations as a linear combination of elements of $\mathcal{B} = \{v_1, v_2, v_3\}$. Thus \mathcal{B} is not a basis of \mathbb{R}^3 , by Theorem 14.7.

2. Let $\mathcal{B} = \{v_1, v_2\}$ where $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$, and $v = (1, 2, 3) \in \mathbb{R}^3$. Then v can not be expressed as a linear combination of elements of \mathcal{B} . This is because, if $\alpha, \beta \in \mathbb{R}$

$$\alpha v_1 + \beta v_2 = (\alpha, \beta, 0)$$

so that the 3rd component is always zero. Thus \mathcal{B} is not basis of $\mathbb{R}^3(\mathbb{R})$.

The above theorem is very important as it gives us a very simple way to determine whether a given set is a basis or not. The following form of the above theorem is also useful: If \mathcal{B} is a subset of V , then \mathcal{B} is not a basis if any one of the following holds:

- (i) Some vector $v \in V$ is not a linear combination of elements of \mathcal{B} .
- (ii) Some vector $v \in V$ can be represented as a linear combination of elements of \mathcal{B} in two different ways.

14.5 Coordinates Relative to an Ordered Basis

An ordered basis for a vector space V is an ordered set of linearly-independent vectors which spans V .

Example 14.7. Let $e_1 = (1, 0)^t$, $e_2 = (0, 1)^t$. Then $\mathcal{B}_1 = \{e_1, e_2\}$, $\mathcal{B}_2 = \{e_2, e_1\}$ are two different ordered basis for \mathbb{R}^2 .

If $V = V(\mathbb{R})$ is vector space and $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ is an ordered basis of V , then for each $v \in V$, there exist unique $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$$

Hence with each vector $v \in V$, we can associate a vector $(\alpha_1, \alpha_2, \dots, \alpha_n)^t$ in \mathbb{R}^n . Conversely, with each $(c_1, c_2, \dots, c_n)^t \in \mathbb{R}^n$, we associate the vector $v = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$ of V . This leads to the following definition.

Definition 14.3. (coordinates of a vector relative to an ordered basis): Let $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ be an ordered basis of a vector space V . Let $v \in V$.

$$v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n, \quad \alpha_i \in F,$$

then the column vector $(\alpha_1, \alpha_2, \dots, \alpha_n)^t$ is called the coordinate vector of v relative to the ordered basis \mathcal{B} . It is denoted by $[v]_{\mathcal{B}}$. The entries of the coordinate vector are called the coordinates of v relative to the basis \mathcal{B} .

Example 14.8. Let $v = (2, -3)^t$. Then if $\mathcal{B}_1, \mathcal{B}_2$ are as defined in Example 14.7,

$$[v]_{\mathcal{B}_1} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \quad [v]_{\mathcal{B}_2} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

Example 14.9. If $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$. $\mathcal{B} = \{v_1, v_2\}$.
Let

$$v = c_1 v_1 + c_2 v_2 \tag{14.21}$$

The vector Equation (14.21)

$$\Leftrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$$

Hence $[v]_{\mathcal{B}} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$.

It is easily verified that for any ordered basis \mathcal{B} of a vector space V , for any $u, v \in V$

$$\begin{aligned} [u + v]_{\mathcal{B}} &= [u]_{\mathcal{B}} + [v]_{\mathcal{B}} \\ [\alpha u]_{\mathcal{B}} &= \alpha [u]_{\mathcal{B}}. \end{aligned}$$

and consequently

$$[\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n]_{\mathcal{B}} = \alpha_1 [v_1]_{\mathcal{B}} + \alpha_2 [v_2]_{\mathcal{B}} + \cdots + \alpha_n [v_n]_{\mathcal{B}}.$$

Hence the linear dependence relation of v_1, v_2, \dots, v_k is equivalent to the linear dependence relation of the vectors $[v_1]_{\mathcal{B}}, [v_2]_{\mathcal{B}}, \dots, [v_k]_{\mathcal{B}}$ of \mathbb{R}^n .

If $S = \{v_1, v_2, \dots, v_k\}$, when studying the linear independence, linear dependence or span of the set S , we study the same property for the corresponding subset $\{[v_1]_{\mathcal{B}}, [v_2]_{\mathcal{B}}, \dots, [v_k]_{\mathcal{B}}\}$ of \mathbb{R}^n .

The following theorem shows that we can extract a basis from any spanning set by weeding out the redundant vectors (in the sense of spanning).

Theorem 14.8. *Let $S = \{v_1, v_2, \dots, v_n\}$ be a set of non-zero vectors in a vector space V such that $\text{span}(S) = V$. Then, some subset T of S is a basis of V .*

Proof: We are looking for a linearly independent subset T of S such that $\text{Span}(T) = \text{Span}(S)$. Two cases arises :

Case 1 S is linearly independent. In this case $T = S$, is a basis of V .

Case 2 S is not linearly independent. Then some vector v_j is a linear combination of the preceding vectors, by Theorem 14.2, Section 14.1.

Let $S_1 = \text{Span}(S \setminus \{v_j\})$ then $\text{Span}(S_1) = \text{Span}(S)$ by Corollary 14.3.

If S_1 is linearly independent, then S_1 is the desired subset of S , which is a basis of V , otherwise we weed out some vector v_k from S_1 in the same manner as above.

Let $S_2 = S_1 \setminus \{v_k\}$, $1 \leq k \leq n, k \neq j$. Then

$$\text{Span}(S_2) = \text{Span}(S_1) = \text{Span}(S).$$

We continue this process. The process of weeding out a redundant vector can not continue for more than $(n-1)$ steps, as after $(n-1)$ steps we will be left with a single non-zero vector, which is always linearly independent.

Thus after k steps, for some $k, 1 \leq k \leq n-1$, we have a linearly independent subset S_k of S such that

$$\text{Span}(S_k) = \text{Span}(S).$$

□

The above theorem can also be stated as “every finitely generated vector space has a basis”. Moreover, the number of elements in a basis is at most equal to the number of elements in the spanning set.

Theorem 14.9. *If A is a given matrix and B is row equivalent to A , then the linear dependence relation, if any, between the columns of A is same as the linear dependence relation amongst the columns of B .*

Proof:

Since B is row equivalent to A , therefore the solution set of $AX = 0$ is same as that of $BX = 0$. Let $A = [c_1 c_2 \dots c_n]$, $B = [b_1 b_2 \dots b_n]$.

For any $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$

$$\alpha_1 c_1 + \alpha_2 c_2 + \dots + \alpha_n c_n = 0$$

$$\iff [c_1 \ c_2 \ \dots \ c_n] \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = 0$$

$$\iff Au = 0, \text{ where } u = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

$$\iff Bu = 0$$

$$\iff \alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n = 0$$

$$\text{Thus } \alpha_1 c_1 + \alpha_2 c_2 + \dots + \alpha_n c_n = 0$$

$$\iff \alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n = 0$$

Thus the linear dependence relationship, (if any) between columns of A is the same as the linear dependence relationship between the columns of B . \square

Theorem 14.10. *Let A be a $m \times n$ matrix. Then the pivot columns of A forms a basis for col A .*

Proof: Let $A = [c_1 \ c_2 \ \dots \ c_n]$. Let B be the reduced echelon form of A , and suppose that the number of pivot columns is k . Then $k \leq n$. Moreover the $(k+1)^{th}, \dots, n^{th}$ rows are zero rows. Observe that the j^{th} pivot column of B is

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \leftarrow j^{th} \text{ position}$$

Clearly, the k pivot columns of B are linearly independent. Any non-pivot column of B is of the form

$$b = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then

$$b = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \alpha_k \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hence b is a linear combination of the pivot columns of B .

Thus every non-pivot column of B is a linear combination of the pivot columns of B , so the pivot columns of B are the only linearly independent columns of B .

Hence the pivot columns of A form a basis of $\text{Col } A$. □

For extracting a basis for \mathbb{R}^m , from a given set $\{v_1, v_2, \dots, v_m\}$ of vectors of \mathbb{R}^m , the following steps can be performed.

Step 1 Let $A = [v_1 \ v_2 \ \dots \ v_m]$.

Step 2 Reduce A to echelon form and find the pivot columns.

Step 3 If every row of A has a pivot then the pivot columns form a basis for \mathbb{R}^m . If every row of A does not have a pivot then $v_1 \dots v_m$ do not span \mathbb{R}^m and we can not extract a basis from these vectors.

Example 14.10. Let $v_i \in \mathbb{R}^4$, $1 \leq i \leq 6$ and $A = [v_1 \ v_2 \ v_3 \ v_4 \ v_5 \ v_6]$

Let the echelon form of the matrix be

$$\begin{pmatrix} \blacksquare & \star & \star & \star & \star & \star \\ 0 & 0 & \blacksquare & \star & \star & \star \\ 0 & 0 & 0 & \blacksquare & \star & \star \\ 0 & 0 & 0 & 0 & \blacksquare & \star \end{pmatrix}$$

Then the 1st, 3rd, 4th and 5th columns have pivots. Also every row has a pivot. Thus $\{v_1, v_3, v_4, v_5\}$ is a basis of \mathbb{R}^4 .

Example 14.11. Let $v_i \in \mathbb{R}^4$, $1 \leq i \leq 4$ and $A = [v_1 \ v_2 \ v_3 \ v_4]$

Suppose the echelon form of A is

$$\begin{pmatrix} \blacksquare & \star & \star & \star \\ 0 & \blacksquare & \star & \star \\ 0 & 0 & \blacksquare & \star \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then the 1st three columns are the pivot columns. Every row does not have a pivot. Thus $\{v_1, v_2, v_3\}$ is not basis of \mathbb{R}^4 . It is only a linearly independent subset of \mathbb{R}^4 .

Let $V(F)$ be a vector space of dimension n . Given a subset $S = \{v_1, v_2, \dots, v_n\}$ of V to find a basis \mathcal{B}' of $\text{Span}(S)$ under certain conditions. Let \mathcal{B} be the standard basis for V .

Condition I To find a subset of S which forms a basis for $\text{span}(S)$

Let $A = [[v_1]_{\mathcal{B}} \ [v_2]_{\mathcal{B}} \ \dots \ [v_n]_{\mathcal{B}}]$. Then the vectors corresponding to the pivot columns of A gives \mathcal{B}' .

Condition II Given a linearly independent set of vectors u_1, u_2, \dots, u_k of $\text{span}(S)$, to find a basis containing u_1, u_2, \dots, u_k . Consider

$$A = ([u_1]_{\mathcal{B}} \ [u_2]_{\mathcal{B}} \ \dots \ [u_k]_{\mathcal{B}} \ [v_1]_{\mathcal{B}} \ [v_2]_{\mathcal{B}} \ \dots \ [v_n]_{\mathcal{B}})$$

Then the vectors corresponding to the pivot columns of A gives \mathcal{B}' . Since the vectors u_i 's are linearly independent therefore the first k columns of A will be amongst the pivot columns of A . Hence $\{u_1, u_2, \dots, u_k\} \subseteq \mathcal{B}'$.

Condition III The basis \mathcal{B}' should not contain any element from S .

We can proceed in two ways.

Step 1 Find a basis $\mathcal{B} = \{w_1, w_2, \dots, w_t\}$ of $\text{Span}(S)$ as in *Condition I*. Let $u = w_1 + w_2 + \dots + w_t \in \text{Span}(S)$. Show that $\{u - w_1, u - w_2, \dots, u - w_t\}$ is a required basis of $\text{span}(S)$. Note that it does not work for $t = 1, 2$.

Step 2 Let $A = [[v_1]_{\mathcal{B}} \ [v_2]_{\mathcal{B}} \ \dots \ [v_n]_{\mathcal{B}}]$. Reduce A^t to echelon form. If any row of the echelon form is same as that of A , then we can change it applying some row transformations. The non-zero rows of the echelon form of A^t written as column vectors is a basis for $\text{Span}(S)$.

Given below is a summary of the results of this section. Let $v_1, v_2, \dots, v_n \in \mathbb{R}^m$, $S = \{v_1, v_2, \dots, v_n\}$ and $A = [v_1 \ v_2 \ \dots \ v_n]$.

	Problem in Vector Space	Equivalent problem in system of linear equations	Solution
1.	Is S a basis of \mathbb{R}^m ?	Does $AX = 0$ has a unique solution?	Yes, if every column of A is a pivot column.
2.	To extract a basis from S		Pivot columns of A form a basis.
3.	Let $\mathcal{B} = \{u_1, u_2, \dots, u_k\}$ be an ordered basis of a vector space V . If $v \in V$, find $[v]_{\mathcal{B}}$	\mathcal{B}' is the standard basis. Let $w_i = [u_i]_{\mathcal{B}'}$, $1 \leq i \leq k$, $b = [v]_{\mathcal{B}'}$ and $B = [w_1 \ w_2 \ \dots \ w_k]$. The solution of $BX = b$ is $[v]_{\mathcal{B}}$.	If reduced echelon form of $[B:b]$ is $[I:c]$. Then $c = [v]_{\mathcal{B}}$.

Problem 14.10. Show that $\mathcal{B} = \{v_1, v_2, v_3\}$ is a basis of \mathbb{R}^3 , where

$$v_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}.$$

Find the coordinates of $(0, -1, 7)$ relative to \mathcal{B} .

Solution: Let $A = (v_1 \ v_2 \ v_3)$.

Step 1 We reduce A to echelon form as follows

$$\begin{aligned} A &= \begin{pmatrix} 2 & 0 & 2 \\ 2 & 1 & 1 \\ 2 & -1 & -3 \end{pmatrix} \\ &\sim \begin{pmatrix} 2 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & -1 & -5 \end{pmatrix} \text{ (applying } R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - R_1) \\ &\sim \begin{pmatrix} 2 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & -6 \end{pmatrix} \text{ (applying } R_3 \rightarrow R_3 + R_2) \\ &= B \end{aligned}$$

Since every column of B has a pivot, therefore each column of A is a pivot column. Thus v_1, v_2, v_3 is a basis for $\text{Col } A$.

Step 2 Since every row of A has a pivot therefore $AX = b$ has a solution for every $b \in \mathbb{R}^3$

$\Rightarrow b \in \text{Col } A$.

Hence $\text{Col } A = \mathbb{R}^3$

Step 3 By Step 1 and Step 2 $\{v_1, v_2, v_3\}$ is a basis of \mathbb{R}^3 .

Problem 14.11. Let $S = \{v_1, v_2, v_3\}$, where $v_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix}$,

$v_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$. Find a subset of S which forms a basis of $\text{Span}(S)$.

Solution: Let $A = [v_1 \ v_2 \ v_3]$.

We reduce A to echelon form.

$$\begin{aligned} A &= \begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & -1 \\ 2 & -4 & 2 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{pmatrix} \text{ (applying } R_2 \rightarrow R_2 + R_1, R_3 \rightarrow R_3 - 2R_1) \\ &\sim \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \text{ (applying } R_3 \rightarrow R_3 + 2R_2) \\ &\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \text{ (applying } R_1 \rightarrow R_1 + R_2) \end{aligned}$$

The 1st and 2nd columns are the pivot columns, therefore the corresponding vectors v_1 and v_2 are linearly independent and forms a basis for $\text{Col } A$.

\therefore But $\text{Col } A = \text{Span}(S)$. \therefore a basis of $\text{Span}(S)$ is $\{v_1, v_2\}$.

Problem 14.12. In the above problem, find a basis of $\text{Span}(S)$ containing v_3 .

Solution: In this case we keep v_3 in the first position, consider the vector equation. Let $B = [v_3 \ v_1 \ v_2]$.

Applying row operation we get

$$B \sim \begin{pmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -3/2 \\ 0 & 0 & 0 \end{pmatrix}$$

The 1st and 2nd columns are the pivot columns, therefore the corresponding vectors v_3 and v_1 are linearly independent and form a basis for $\text{Col } B$. But $\text{Col } B = \text{Span}(S)$

$\therefore \{v_3, v_1\}$ is a basis for $\text{Span}(S)$.

Problem 14.13. Find a basis for $\text{Col } A$ which are not from the columns of A , where

$$A = \begin{pmatrix} -2 & -2 & 0 & -6 \\ -1 & 0 & -2 & -1 \\ 1 & 1 & 0 & 3 \\ 2 & 1 & 1 & 5 \end{pmatrix}$$

Solution: Since we are applying row operations we write the columns of A as the rows of A^t , so that $\text{Col } A = \text{Row } A^t$. To find a basis for $\text{Col } A$, we find a basis for $\text{Row } A^t$. Reducing A^t to reduced echelon form,

$$A^t \sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

A basis for row A^t is $\{(1, 0, 1, 0), (0, 1, 0, -1/2)\}$.

Hence a basis for $\text{Col } A$ is $\{(1, 0, 1, 0)^t, (0, 1, 0, -1/2)^t\}$

Problem 14.14. Let $S = \{(0, 1, 2)^t, (-1, -1, 2)^t, (2, 1, 4)^t, (1, 2, 3)^t\}$. Find two different bases of $\text{Span}(S)$, one from the set S and the other not having any element from S .

Solution:

$$\text{Let } A = \begin{pmatrix} 0 & -1 & 2 & 1 \\ 1 & -1 & 1 & 2 \\ 2 & 2 & 4 & 3 \end{pmatrix}$$

Reducing A to echelon form, we have

$$A \sim \begin{pmatrix} 1 & -1 & 1 & 2 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 10 & 3 \end{pmatrix} = B(\text{say})$$

The 1st, 2nd and 3rd columns of A are the pivot columns.

Thus $\{(0, 1, 2)^t, (-1, -1, 2)^t, (2, 1, 4)^t\}$ is a basis for $\text{Span}(S)$ from the set S .

Let us now find a basis whose elements are not from S . reducing A^t to reduced echelon form, we get

$$A^t \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = B(\text{say})$$

The rows of B having the pivot elements, written as column vectors, gives a basis for $\text{Span}(S)$. It is $\{(1, 0, 0)^t, (0, 1, 0)^t, (0, 0, 1)^t\}$.

Problem 14.15. Prove that $1, 1 - x, 2 - 4x + x^2, 6 - 18x + 9x^2 - x^3$ forms a basis for P_3 . Find the coordinates of $3x^2 - 8x + 7$ relative to this basis.

Solution: Let $S = \{p_1, p_2, p_3, p_4\}$, where $p_1 = 1, p_2 = 1 - x, p_3 = 2 - 4x + x^2, p_4 = 6 - 18x + 9x^2 - x^3$ and $p = 3x^2 - 8x + 7$.

The standard basis for P_3 is $\{1, x, x^2, x^3\} = \mathcal{B}$. Then

$u_1 = [p_1]_{\mathcal{B}} = (1, 0, 0, 0)^t, u_2 = [p_2]_{\mathcal{B}} = (1, -1, 0, 0)^t, u_3 = [p_3]_{\mathcal{B}} = (2, -4, 1, 0)^t,$
 $u_4 = [p_4]_{\mathcal{B}} = (6, -18, 9, -1)^t, u = [p]_{\mathcal{B}} = (7, -8, 3, 0)^t.$

Let $A = [u_1, u_2, u_3, u_4]$.

We have to

- (i) Prove that the columns of A are linearly independent.
- (ii) Find the solution of

$$x_1 u_1 + x_2 u_2 + x_3 u_3 + x_4 u_4 = u$$

i.e. of $AX = u$, where $X = (x_1, x_2, x_3, x_4)^t$.

Reduce the augmented matrix $[A : u]$ to echelon form.

$$[A : u] = \begin{pmatrix} \mathbf{1} & 1 & 2 & 6 & \vdots & 7 \\ 0 & \mathbf{-1} & -4 & -18 & \vdots & -8 \\ 0 & 0 & \mathbf{1} & 9 & \vdots & 3 \\ 0 & 0 & 0 & \mathbf{-1} & \vdots & 0 \end{pmatrix}$$

which is already in echelon form. The pivots are in bold. Since every column of A is a pivot column, therefore, the columns of A are linearly independent. Since every row has a pivot element, \therefore the columns of A span P_3 . Hence S is a basis for P_3 , $o(S) = 4 = \dim P_3$.

Reduce $[A : u]$ to reduced echelon form to obtain the required linear combination.

$$[A : u] \sim \begin{pmatrix} 1 & 0 & 0 & 0 & \vdots & 5 \\ 0 & 1 & 0 & 0 & \vdots & -4 \\ 0 & 0 & 1 & 0 & \vdots & 3 \\ 0 & 0 & 0 & 1 & \vdots & 0 \end{pmatrix}$$

Hence solution is $x_1 = 5, x_2 = -4, x_3 = 3, x_4 = 0$.

$\therefore p = 5p_1 - 4p_2 + 3p_3 + 0p_4 = 5p_1 - 4p_2 + 3p_3.$

$[p]_{\mathcal{B}} = (5, -4, 3, 0)^t$

Problem 14.16. For what values of α, β do the vectors v_1, v_2 form a basis, for $\text{span}(\{v_1, v_2\})$, where $v_1 = (1, 1, 1)^t, v_2 = (1, \alpha, \beta)^t$.

Solution: Let $S = \{v_1, v_2\}$ and $A = (v_1 \ v_2)$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & \alpha \\ 1 & \beta \end{pmatrix}$$

S is basis if every column of A is a pivot column.

$$A \sim \begin{pmatrix} 1 & 1 \\ 0 & \alpha - 1 \\ 0 & \beta - 1 \end{pmatrix} = B(\text{say})$$

In B , 2nd column is a pivot column if $\alpha - 1 \neq 0$, $\beta - 1$ may take any value. Thus $\alpha \neq 1$, β can take any value. Thus A has two pivot columns if $\alpha \neq 1$, β can take any value.

14.6 Exercise

- Show that $\mathcal{B} = \{v_1, v_2, v_3\}$ forms a basis for \mathbb{R}^3 , where
 - $v_1 = (1, 2, 0), v_2 = (1, 0, 1), v_3 = (0, 1, 1)$.
 - $v_1 = (-1, 2, 0), v_2 = (3, -5, 2), v_3 = (4, -7, 3)$.
 - $v_1 = (1, -3, 2), v_2 = (2, 4, 1), v_3 = (1, 1, 1)$.
- Let $\mathcal{B} = \{v_1, v_2, v_3\}$, where $v_1 = (1, -3), v_2 = (2, -8), v_3 = (-3, 7)$.
 - Show that \mathcal{B} spans \mathbb{R}^2 .
 - Express $(1, 1)$ as a linear combination of elements of \mathcal{B} in two different ways.
 - Is \mathcal{B} a basis for \mathbb{R}^2 ? Justify your answer.
- Let $S = \{(1, 2), (2, 4), (4, 5)\}$. Show that S generates \mathbb{R}^2 . Find all subsets of S which generate \mathbb{R}^2 .
- Given a basis $\mathcal{B} = \{(1, 0, 0), (0, 1, 0), (0, 1, 1)\}$ for \mathbb{R}^3 , replace a suitable element of \mathcal{B} by the vector $(1, 2, 2)$ to get another basis for \mathbb{R}^3 .
- Find two bases for $\mathbb{C}^4(\mathbb{C})$ such that the only vectors common to both the bases are $(1, 1, 0, 0)^t$ and $(0, 0, 1, 1)^t$.
- Let $v_1 = (4, -3, 7)^t, v_2 = (1, 9, -2)^t, v_3 = (7, 11, 6)^t$ and let $W = \text{Span}\{v_1, v_2, v_3\}$. Find three bases for W .
- Let $v_1 = (7, 4, -9, -5)^t, v_2 = (4, -7, 2, 5)^t, v_3 = (1, -5, 3, 4)^t$. Find a basis for $W = \text{Span}\{v_1, v_2, v_3\}$ such that in each element the number of 0s preceding the first non-zero entry is different.
- Let $\mathcal{B} = \{v_1, v_2, v_3\}$, where $v_1 = (1, -3, 2), v_2 = (2, 4, 1), v_3 = (1, 1, 1)$.
 - Prove that \mathcal{B} is a basis of \mathbb{R}^3 .
 - Find the coordinates of v_1, v_2, v_3 relative to \mathcal{B} .
 - Find the coordinates of the vectors $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ relative to \mathcal{B} .
- Show that the set $\mathcal{B} = \{v_1, v_2, v_3\}$ forms a basis for \mathbb{R}^3 . Also find the coordinates of v relative to the ordered basis \mathcal{B} .
 - $v_1 = (1, 0, 1), v_2 = (2, 0, 3), v_3 = (2, 1, 3), v = (-1, 1, -2)$.
 - $v_1 = (1, 1, 0), v_2 = (1, 0, 1), v_3 = (1, 2, 3), v = (1, -5, -10)$.
 - $v_1 = (1, 0, 1), v_2 = (1, 1, 0), v_3 = (1, 2, 3), v = (1, -5, -10)$.
 - $v_1 = (1, 5, 3), v_2 = (1, 0, -1), v_3 = (1, 0, 0), v = (2, 5, 5)$.
- Let $v_1 = (-1, 2, 0), v_2 = (3, -5, 2), v_3 = (4, -7, 3)$. If $\mathcal{B} = \{v_1, v_2, v_3\}$ is an ordered basis for V , find $[u]_{\mathcal{B}}$, where $u = (0, 1, -5)$.

11. Let $v_1 = (-1, 2, 0)$, $v_2 = (3, -5, 2)$, $v_3 = (4, -7, 3)$. If $\mathcal{B} = \{v_1, v_2, v_3\}$ is an ordered basis for V , find $[u]_{\mathcal{B}}$, where
- $u = (1, 0, 0)$.
 - $u = (0, 1, 0)$.
 - $u = (0, 0, 1)$.
12. What are the coordinates of v_1, v_2, v_3 relative to the ordered basis
- $\{v_1, v_2, v_3\}$.
 - $\{v_2, v_1, v_3\}$.
 - $\{v_1 + v_2, v_2 + v_3, v_3 + v_1\}$.
 - $\{v_3, v_3 + v_2, v_3 + v_2 + v_1\}$.
13. Given a basis $\{v_1, v_2, v_3\}$ for \mathbb{R}^3 , where $v_1 = (1, 2, 0)$, $v_2 = (0, 1, 1)$, $v_3 = (1, 0, 1)$, find $[u]_{\mathcal{B}}$, where $u = (1, 2, 3)$ and a ordered basis is given by
- $\{v_1, v_2, v_3\}$.
 - $\{v_2, v_1, v_3\}$.
 - $\{v_1, v_3, v_2\}$.
 - $\{v_3, v_2, v_1\}$.
14. Find a basis for each of the following vector spaces:
- $\mathbb{C}(\mathbb{R})$.
 - $\mathbb{C}^2(\mathbb{C})$.
 - $\mathbb{C}^2(\mathbb{R})$.
 - The vector space of all 2×2 complex diagonal matrices, over \mathbb{R} .
 - The vector space of all 2×2 complex diagonal matrices, over \mathbb{C} .
 - $\mathbb{C}(\mathbb{C})$.
 - $V(\mathbb{Q})$ where $V = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.
15. Does the set $S = \{3, x^2 - x + 3, 3x^3 + x^2 - x + 5, 3x^3 + x^2 + 6\}$ form a basis for $P_3(\mathbb{R})$. Give another basis for $P_3(\mathbb{R})$. What is the number of elements in each of the two bases?
16. In $P_2(\mathbb{R})$, obtain a $[p]_{\mathcal{B}}$, where $\mathcal{B} = \{1, 1 + x, (1 + x)^2\}$ and $p = 2 + 3x - x^2$.
17. Let $p_1(x) = 2 + x^2$, $p_2(x) = 1 + 2x$, $p_3(x) = 1 + x + x^2$. Consider the bases $\mathcal{B} = \{p_1, p_2, p_3\}$ for P_2 .
- Find $[p]_{\mathcal{B}}$, where $p(x) = 3 - x + 4x^2$
 - If $[p]_{\mathcal{B}}$ is $(4, -5, 1)^t$, find $p(x)$.
18. In the vector space $\mathbb{C}^3(\mathbb{C})$, find $[u]_{\mathcal{B}}$, where $u = (3 + 4i, 6i, 3 + 7i)^t$ and $\mathcal{B} = \{(1, 0, 0), (1, 1, 1), (1, 1, 0)\}$.
19. W_1 and W_2 are two subspaces of \mathbb{R}^4 with basis $\mathcal{B}_1 = \{e_1, e_2\}$ and $\mathcal{B}_2 = \{e_2, e_3, e_4\}$ respectively.
- What is the form of any element of $W_1 \cap W_2$? Give a basis for $W_1 \cap W_2$.
 - What is the form of any element of $W_1 + W_2$? Give a basis for $W_1 + W_2$.
 - What are $\dim W_1$, $\dim W_2$, $\dim(W_1 + W_2)$ and $\dim(W_1 \cap W_2)$? What relationship do you see amongst them?

20. Show that $\{A_1, A_2, A_3, A_4\}$, where

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

forms a basis for the vector space $V(\mathbb{R})$, where $V = M_2(\mathbb{R})$.

21. Find all values of a for which $(a^2, 0, 1), (0, a, 2), (1, 0, 1)$ is a basis for \mathbb{R}^3 .

22. Show that if $\{v_1, v_2, \dots, v_n\}$ is a basis for a vector space V and $0 \neq \alpha \in F$, then $\{\alpha v_1, v_2, \dots, v_n\}$ is also a basis for V .

23. Show that if $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ is a basis for a vector space V , then show that $S = \{w_1, w_2, \dots, w_n\}$, where $w_i = v_i + v_{i+1} + \dots + v_n, 1 \leq i \leq n$ is also a basis for V .

14.7 Dimension

In the previous section, we have seen that a finitely generated vector space can have more than one basis. There is one thing common about the different bases. It is the number of elements in each of them. This leads us to believe that different bases have the same number of elements. That this is true, will be proved subsequently.

Definition 14.4. (Finite dimensional vector space): A vector space $V(F)$ is said to be finite dimensional if it can be generated by a finite subset of V .

One of the main results of vector spaces is “In a finite dimensional vector space, any two bases have the same number of elements”.

Before proving this, we prove a few results.

Theorem 14.11. If a vector space $V(F)$ has a basis consisting of n vectors, then any set S consisting of m vectors with $m > n$, must be linearly dependent.

Proof: Let B be a basis for V consisting of n vectors. Let $S = \{u_1, u_2, \dots, u_m\}$ be a subset of V where $m > n$. Let $w_i = [u_i]_B, 1 \leq i \leq m$ are m vectors of \mathbb{R}^n .

If $A = [w_1 \ w_2 \ \dots \ w_m]$, be the matrix whose columns are w_1, w_2, \dots, w_m . Then A is $n \times m$ matrix, so A can have at most n pivots (as A has n rows). Thus, A can have at most n pivot columns so that the m columns of A must be linearly dependent.

\therefore There exist scalars $\alpha_1, \alpha_2, \dots, \alpha_m$ not all zero such that

$$\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_m w_m = 0$$

But, $[\alpha_1 u_1 + \dots + \alpha_m u_m]_B = \alpha_1 [u_1]_B + \dots + \alpha_m [u_m]_B = \alpha_1 w_1 + \dots + \alpha_m w_m$

Hence, $[\alpha_1 u_1 + \dots + \alpha_m u_m]_B = 0$ so that $\alpha_1 u_1 + \dots + \alpha_m u_m = 0$ Since α_i s are not all zero, therefore u_1, u_2, \dots, u_m are linearly dependent. \square

Corollary 14.12. If a vector space $V(F)$ has a basis consisting of n vectors, and S is a linearly independent set containing m elements, then $m \leq n$.

Definition 14.5. In a vector space $V(F)$, if S is a linearly independent subset of V and every superset of S is linearly dependent, then S is called a maximal linearly independent subset of V .

Corollary 14.13. *If B is a basis for V , then B is a maximal linearly independent subset of V .*

Thus, we have that the number of linearly independent elements in a vector space is not more than the number of elements in a basis.

Theorem 14.14. *In a finite dimensional vector space, any two bases have the same number of elements.*

Proof: Let $V(F)$ be a finite dimensional vector space. Therefore V has a finite basis. Let B and B' be two bases of V containing m and n elements, respectively.

Since B is a basis and B' is a linearly independent set, therefore by the above corollary

$$n \leq m \dots (1)$$

Similarly, interchanging the roles of B and B' in the above argument we get

$$m \leq n \dots (2)$$

(1) and (2) $\Rightarrow n = m$. Hence proved. \square

We are now in a position to assign a proper name to the number of elements in a finite dimensional vector space.

Definition 14.6. (Dimension):

In a finite dimensional vector space $V(F)$, the number of elements in a basis is called the dimension of V and is denoted by $\dim_F V$ or simply $\dim V$. The dimension of the zero space is defined to be zero.

If V is not finite dimensional, then V is said to be infinite dimensional, and we write $\dim V = \infty$

Example 14.12. Consider $\mathbb{R}^3(\mathbb{R})$ $v_1 = (1, 0, 0), v_2 = (0, 1, 0), v_3 = (0, 0, 1)$
Then $\{v_1, v_2, v_3\}$ is a basis of \mathbb{R}^3 . $\therefore \dim \mathbb{R}^3 = 3$

Example 14.13. A basis of $\mathbb{C}(\mathbb{R})$ is $\{1, i\}$

$$\therefore \dim_{\mathbb{R}} \mathbb{C} = 2$$

Example 14.14. A basis of $\mathbb{C}(\mathbb{C})$ is $\{1\}$

$$\therefore \dim_{\mathbb{C}} \mathbb{C} = 1$$

Example 14.15. A basis of $P_n(\mathbb{R})$ is $\{1, x, \dots, x^n\}$

$$\therefore \dim_{\mathbb{R}} P_n = n + 1$$

Example 14.16. A basis of all 2×3 matrices with real entries ($M_{2 \times 3}$) over \mathbb{R} is

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

Hence, $\dim_{\mathbb{R}} M_{2 \times 3} = 6$.

Theorem 14.15. *Let $V(F)$ be a finite dimensional vector space. Then, every linearly independent subset of V can be extended to form a basis of V .*

Proof: Let $\dim V = n$ and let $S = \{v_1, v_2, \dots, v_k\}$ be a linearly independent subset of V . If $\text{Span}(S) = V$, then S is the required basis of V . If $\text{Span}(S) \neq V$, then there exists $u_1 \in V$ such that $u_1 \notin \text{Span}(S)$. Let $S_1 = S \cup \{u_1\} = \{v_1, v_2, \dots, v_k, u_1\}$. Then S_1 is a linearly independent subset of V , because no vector of S_1 can be expressed as a linear combination of the preceding ones. If $\text{Span}(S_1) = V$, then S_1 is the required basis of V . Else, choose $u_2 \in V, u_2 \notin \text{Span}(S_1)$. After $n-k$ steps, we reach at a set S_k containing n linearly independent elements. We assert that $\text{Span}(S_k) = V$. On the contrary, supposed $\text{Span}(S_k) \neq V$. Then there exists $u_{k+1} \in V, u_{k+1} \notin \text{Span}(S_k)$. Now, $S_{k+1} = S_k \cup \{u_{k+1}\}$ is a linearly independent set containing $n+1$ elements. This contradicts Corollary 14.12. Hence our assumption is wrong, so that $\text{Span}(S_k) = V$. Hence S_k is the required basis. \square

Theorem 14.16. *If W is a subspace of a finite dimensional vector space $V(F)$, then W is also finite dimensional, and*

$$\dim_F W \leq \dim_F V$$

Proof: Let $\dim_F V = n$. Let W be a subspace of V . Since $\dim V = n$, \therefore any set of $n+1$ vectors in V are linearly dependent. In particular, any set of $n+1$ vectors in W are linearly dependent. Thus, we can find a largest set of linearly independent vectors in W , say w_1, w_2, \dots, w_m . Then $m \leq n$. Let $S = \{w_1, w_2, \dots, w_m\}$. If $w \in W$ then $S_1 = S \cup \{w\}$ is a linearly dependent set. Thus, there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_m, \alpha$ not all zero such that

$$\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_m w_m + \alpha w = 0$$

If $\alpha = 0$, by the linear independence of S , we get $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$, which is a contradiction. Hence, $\alpha \neq 0$ and

$$w = -\alpha^{-1}(\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_m w_m)$$

Thus, $\text{Span}(S) = W$. Thus, W is finite dimensional over F .

$$\dim_F W = m \leq n = \dim_F V$$

$$\text{i.e. } \dim_F W \leq \dim_F V \quad \square$$

A set \mathcal{B} of vectors in a vector space V is a basis for V if

1. \mathcal{B} spans V , and
2. \mathcal{B} is linearly independent.

If the number of vectors in the set is same as the dimension of V then we need to check only one of the two conditions. This is proved in the following theorem.

Theorem 14.17. *Let $V(F)$ be an n dimensional vector space. Then*

1. Any set of n linearly independent vectors of V is a basis.
2. Any set of n vectors in V which spans V is a basis.

Proof:

1. Let $S = \{v_1, v_2, \dots, v_n\}$ be a set of n linearly independent vectors. By theorem 14.15 S can be extended to a basis of V , say S' . Then $o(S') = \dim V = n$. Now $S \subset S'$, S' is finite and $o(S) = o(S')$, $\therefore S = S'$. Hence, S is a basis for V .
2. Let $T = \{u_1, u_2, \dots, u_n\}$ be a set of n vectors such that $\text{Span}(T) = V$. Then, T has a subset S which is a basis for V . Since $\dim V = n$, therefore S has n elements. Now, $S \subseteq T$ and $o(S) = o(T)$. $\therefore S = T$. Hence T is a basis for V . \square

Example 14.17. Extend $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -1 \\ 1 \end{pmatrix} \right\}$ to a basis of \mathbb{R}^4 .

$$\text{Let } u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 2 \\ -1 \\ 1 \end{pmatrix}.$$

Clearly, $S = \{u_1, u_2\}$ is a linearly independent set.

$$\text{Span } S = \{\alpha u_1 + \beta u_2 \mid \alpha, \beta \in \mathbb{R}\} = \{(\alpha, 2\beta, -\beta, \beta) \mid \alpha, \beta \in \mathbb{R}\}.$$

Choosing $\alpha = 1, \beta = 0$, we see that $(1, 0, 0, 0) \in \text{Span}(S)$

Observe that $u_3 = (1, 0, 0, 1) \notin \text{Span}(S)$.

Let $S_1 = S \cup \{u_3\}$. Then S_1 is a linearly independent set. $\text{Span}(S_1) = \{\alpha u_1 + \beta u_2 + \gamma u_3 \mid \alpha, \beta, \gamma \in \mathbb{R}\} = \{(\alpha + \gamma, 2\beta, -\beta, \beta + \gamma) \mid \alpha, \beta, \gamma \in \mathbb{R}\}$ For $\alpha = \beta = \gamma = 1$, $(2, 2, -1, 2) \in \text{Span}(S_1) \therefore u_4 = (2, 2, 0, 2) \notin \text{Span}(S_1)$

Let $S_2 = S_1 \cup \{u_4\}$.

Then S_2 is a linearly independent set.

$$\text{Let } A = [u_1 \quad u_2 \quad u_3 \quad u_4] = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 0 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

$$\text{Reducing } A \text{ to Echelon form, } A \sim \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Since every column of A has a pivot, therefore

1. The 4 columns of A are linearly independent.
2. The columns of A span \mathbb{R}^4 .

Thus, $B = \{u_1, u_2, u_3, u_4\}$ is a basis of \mathbb{R}^4 .

An alternative method of extending a given linearly independent subset S of a vector space V , to a basis is the following:

1. Let B be a basis of V . You can take the standard basis.
2. $S \cup B$ spans V .
3. Extract a basis from $S \cup B$, by retaining the set S .

Example 14.18 illustrates this procedure.

Example 14.18. Let $v_1 = \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$. Extend $\{v_1, v_2\}$ to a basis

of V by finding a subset of $\{v_1, v_2, e_1, e_2, e_3, e_4\}$ containing v_1, v_2 which forms a basis of V .

Let $S = \{v_1, v_2\}$. Let $T = \{v_1, v_2, e_1, e_2, e_3, e_4\}$ We shall find a subset B of T such that

1. $S \subseteq B$.
2. B is a basis of V .

To do this, proceed as follows:

$$\begin{aligned} \text{Let } A = (v_1 \ v_2 \ e_1 \ e_2 \ e_3 \ e_4) &= \begin{pmatrix} -2 & -1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \\ &\sim \begin{pmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & -1 & 1 & 2 \end{pmatrix} = B \text{ (say)} \end{aligned}$$

B is an echelon form of A . The first 4 columns of B are the pivot columns, so that the first 4 columns of A are the pivot columns of A . Since the pivot columns are linearly independent, $\therefore v_1, v_2, e_1, e_2$ are linearly independent. Since $\dim V = 4$, therefore $\{v_1, v_2, e_1, e_2\}$ forms a basis for V .

Note

While writing the matrix A , the elements of the given linearly independent set must be written first. This way, they will remain in the required basis.

Theorem 14.18. If A is any $m \times n$ matrix, and $r =$ number of basic variables in the equation $AX = 0$ then

1. $\dim(\text{Col}A) =$ number of basic variables.
2. $\dim(\text{Nul}A) =$ number of free variables.
3. $\dim(\text{Col}A) + \dim(\text{Nul}A) = n$.

Proof: Let $A = [v_1 \ v_2 \ \dots \ v_n]$, where $v_i \in \mathbb{R}^m$, $1 \leq i \leq n$.

1. We know that, number of basic variables + number of free variables = total number of variables. Since there are r basic variables, therefore, there are r pivot columns in A . The pivot columns of A form a basis of $\text{Col}A$.

$$\therefore \qquad \qquad \qquad \dim \text{Col}A = r$$

2. Without any loss of generality, we can assume that the first r columns are the pivot columns as it amounts to renaming the variables only. Then, the Echelon form of A is

$$B = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & b_{1 \ r+1} & b_{1 \ r+2} & \dots & b_{1 \ n} \\ 0 & 1 & 0 & \dots & 0 & b_{2 \ r+1} & b_{2 \ r+2} & \dots & b_{2 \ n} \\ 0 & 0 & 1 & \dots & 0 & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & 1 & b_{r \ r+1} & b_{r \ r+2} & \dots & b_{r \ n} \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Any solution of $AX = 0$ is the same as that of $BX = 0$, so that $NulA = NulB$.

Any solution of $BX = 0$ is

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = x_{r+1} \begin{pmatrix} -b_{1 \ r+1} \\ -b_{2 \ r+1} \\ \vdots \\ -b_{r \ r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{r+2} \begin{pmatrix} -b_{1 \ r+2} \\ -b_{2 \ r+2} \\ \vdots \\ -b_{r \ r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} -b_{1 \ n} \\ -b_{2 \ n} \\ \vdots \\ -b_{r \ n} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= x_{r+1}u_{r+1} + x_{r+2}u_{r+2} + \dots + x_nu_n \text{ (say)}$$

Thus, $S = \{u_{r+1}, \dots, u_n\}$ spans $NulB$ and $\therefore NulA$. If $\alpha_{r+1}u_{r+1} + \dots + \alpha_nu_n = 0$, for α_i 's in F then on comparing the last $n - r$ components, we get

$$\alpha_{r+1} = \dots = \alpha_n = 0$$

$\therefore u_{r+1}, \dots, u_n$ are linearly independent. Hence, S is a basis of $NulA$.

$$\therefore \dim NulA = n - r = \text{number of free variable}$$

3. Since $r + (n - r) = n \quad \therefore \dim ColA + \dim NulA = \text{total number of variables}$
 $= \text{number of columns of } A$
 $= n$ □

Example 14.19. Let $A = \begin{pmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & -5 & -7 & 3 \\ 1 & -2 & 0 & 0 & -5 & -4 \end{pmatrix}$. The echelon

form of A is $B = \begin{pmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -1 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Number of basic variables = 4.

Number of free variables = 2

Let us find a basis for $ColA$ and its dimension. The 1st, 2nd, 4th and 5th

columns are the pivot columns of A . Thus, a basis of $ColA$ is

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \\ -3 \\ -2 \end{pmatrix}, \begin{pmatrix} 7 \\ 10 \\ 1 \\ -5 \\ 0 \end{pmatrix}, \begin{pmatrix} 9 \\ 13 \\ 1 \\ -7 \\ -5 \end{pmatrix} \right\}$$

$$\dim ColA = 4$$

$\dim NulA =$ number of free variables

$$= 2$$

$\therefore \dim ColA + \dim NulA = 4 + 2 = 6 =$ number of columns of A .

Let us now find a basis of $NulA = \{X \in \mathbb{R}^6 | AX = 0\}$.

$$\begin{aligned} \text{If } X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \in NulA, \text{ then } AX = 0 &\iff BX = 0 \\ \iff x_1 + x_2 - 3x_3 + 7x_4 + 9x_5 - 9x_6 = 0 \\ x_2 - x_3 + 3x_4 + 4x_5 - 3x_6 = 0 \\ x_4 = 0 \\ x_5 + 2x_6 = 0 \end{aligned}$$

Taking x_3, x_6 as the free variables,

we get $x_1 = 2x_3 + 16x_6$

$$x_2 = x_3 + 11x_6$$

$$x_4 = 0$$

$$x_5 = -2x_6$$

\therefore

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = x_3 \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} 16 \\ 11 \\ 0 \\ 0 \\ -2 \\ 1 \end{pmatrix} \dots (1)$$

$$\text{Let } u_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 16 \\ 11 \\ 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}. \text{ Then } X \in NulA \text{ iff } X = x_3 u_1 + x_6 u_2$$

where x_3, x_6 are arbitrary. Thus $X \in Span\{u_1, u_2\}$. Since u_1, u_2 are linearly independent, we get that $\{u_1, u_2\}$ is a basis for $NulA$.

This solution is the same as in (1). Basis of $NulA$ is $\{u_1, u_2\}$.

Note that if we find the solution by using the reduced echelon form of A instead of echelon form, then the calculations are simplified. So, it is advisable to use the reduced echelon form if a basis for $NulA$ is needed. Calculations using both the forms have been done.

Steps to find a basis of Col A, dim(ColA), NulA, dim(NulA)

Let A be a given matrix.

Step 1 Reduce A to Echelon form. Find the pivot columns, number of free variables and number of basic variables.

Step 2 A basis for Col A are the pivot columns of A.

dim(ColA) = number of pivot columns = number of basic variables

Step 3 dim(NulA) can be found without finding a basis for NulA.

dim(NulA) = number of free variables.

To find a basis for NulA, reduce A to reduced echelon form. Express a solution of $AX = 0$ in terms of the free variables. Give values to the free variables as follows: One free variable is assigned the value 1, all others are assigned the value zero. The vectors obtained in this manner form a basis for NulA.

14.8 Rank of a Matrix

Given a system of linear equations, some of the equations are linear combination of others. Thus, there are only a certain number of essentially different equations and the others are linear combination of them. Also, there are some variables which are basic variables. We will study the relationship between the number of essentially different equations and the number of basic variables.

Theorem 14.19. *Two row equivalent matrices have the same row space.*

Proof: Let A and B be two row equivalent matrices. Let $A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}$

and $B = \begin{pmatrix} R'_1 \\ R'_2 \\ \vdots \\ R'_m \end{pmatrix}$. Since B is obtained from A by applying row operations,

therefore the rows of B are linear combinations of rows of A. Hence $R'_j \in \text{Span}(R_1, \dots, R_m)$, $1 \leq j \leq m$, so that $\text{Span}(R'_1, \dots, R'_m) \subset \text{Span}(R_1, \dots, R_m)$

$$\therefore \text{Row}B \subseteq \text{Row}A \quad \dots (1)$$

Since A can be obtained from B by applying row operations, so interchanging the roles of A and B, we get

$$\text{Row}A \subseteq \text{Row}B \quad \dots (2)$$

From (1) and (2), we get

$$\text{Row}A = \text{Row}B$$

Hence, proved. □

Corollary 14.20. *If B is an echelon form of A, then the non-zero rows of B form a basis for RowA.*

Proof: Since A and B are row equivalent,

$$\therefore \text{Row}A = \text{Row}B \quad \dots(1)$$

B is in echelon form, therefore the non-zero rows of B span $\text{Row}B$. Since B is in echelon form, therefore the leading non-zero elements of the rows are in different columns. Hence, the non-zero rows are linearly independent. Thus, the non-zero rows of B form a basis for $\text{Row}B$. By (1), we get that the non-zero rows of B form a basis for $\text{Row}A$. \square

Definition 14.7. (Row rank): The dimension of $\text{Row}A$ is called the row rank of the matrix A .

Definition 14.8. (Column rank): The dimension of $\text{Col}A$ is called the column rank of the matrix A .

Theorem 14.21. The column rank and row rank of a matrix are equal.

Proof: Let A be an $m \times n$ matrix. Let column rank of $A = r$. Then r is the number of pivot columns of A (By Theorem 14.10). Let B be an echelon form of A . Since no two pivots in B lies in the same row or in the same column, Number of pivots of $B =$ Number of pivot columns of A , and Number of pivots of $B =$ Number of non-zero rows of B .

So, Number of pivot columns of $A =$ Number of non-zero rows of B .

This gives that Column rank of $A =$ Row rank of A (By Theorem 14.18) \square

In view of the above theorem, the common number (column rank of A or the row rank of A) associated with a given matrix is called its rank.

Definition 14.9. (Rank of a matrix): The rank of a matrix is the dimension of the column space of A (or the row space of A).

Remark 14.5. 1. Let B be an echelon form of a given matrix A . Then a basis for $\text{Row}A$ is the non-zero rows of B . It is important to note that the pivot columns of B do not form a basis for $\text{Col}A$. Actually, the pivot columns of A form a basis for $\text{Col}A$.

2. Observe that if A is an $m \times n$ matrix then $\text{Row}A$ is a subspace of \mathbb{R}^m , $\text{Col}A$ is a subspace of \mathbb{R}^n . In general $m \neq n$. Thus, $\text{Row}A$ and $\text{Col}A$ are subspaces of different vector spaces but they are of the same dimension.

Example 14.20. Let $A = \begin{pmatrix} 2 & -3 & 6 & 2 & 5 \\ -2 & 3 & -3 & -3 & -4 \\ 4 & -6 & 9 & 5 & 9 \\ -2 & 3 & 3 & -4 & 1 \end{pmatrix}$ Then, an echelon

form of A is

$B = \begin{pmatrix} \mathbf{2} & -3 & 6 & 2 & 5 \\ 0 & 0 & \mathbf{3} & -1 & 1 \\ 0 & 0 & 0 & \mathbf{1} & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ The pivots have been encircled. Let us find bases for $\text{Col}A$ and $\text{Row}A$.

Basis for ColA

The basis for ColA are the pivot columns of A. These are the 1st, 3rd and 4th columns of A.

So a basis for Col A

$$\left\{ \begin{pmatrix} 2 \\ -2 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} 6 \\ -3 \\ 9 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 5 \\ -4 \end{pmatrix} \right\}$$

Column rank of A = dimColA = 3

Basis for RowA

A basis for RowA are the non-zero rows of B. It is $\{ (2 \ -3 \ 6 \ 2 \ 5), (0 \ 0 \ 3 \ -1 \ 1), (0 \ 0 \ 0 \ 1 \ 3) \}$

row rank of A = dimRowA
= 3

Thus, column rank of A = row rank of A = 3.

Hence Rank A = 3

Definition 14.10. (Nullity A): If A is any matrix, then the dimension of NulA is called the nullity of A.

Theorem 14.22. (Rank nullity theorem): If A is any matrix, then Rank A + Nullity A = number of columns of A.

Proof: Let A be an $m \times n$ matrix. Then by Theorem 14.18 $\dim \text{Col}A + \dim \text{Nul}A = n$

i.e. RankA + NullityA = number of columns of A. □

Steps to find RankA, basis for RowA and dim RowA.

Step 1 Reduce A to Echelon Form. Call it B.

Step 2 The non-zero rows of B form a basis for RowA.
 $\dim \text{Row}A = \text{number of non-zero rows of } B.$

Step 3 RankA = dim rowA
= number of non-zero rows of B.

Given below is a summary of the results of this section.

Let $S = \{v_1, \dots, v_k\}$ be a linearly independent set of a vector space V. Let $\mathcal{B} = \{u_1, \dots, u_n\}$ be some basis of V.

Let $B = [v_1 \dots v_k \ u_1 \dots u_n]$ and A is any $m \times n$ matrix.

S. No.	Problem in vector spaces	Method of solving
1	Extend S to a basis of V	Pivot columns of B
2	$\dim \text{Col} A$	Number of pivot columns of A
3	Basis for $\text{Col} A$	Pivot columns of A
4	$\dim \text{Nul} A$	Number of non-pivot columns of A
5	Basis for $\text{Nul} A$	Express general solution of $AX = 0$ as a linear combination of vectors with free variables as coefficients. These vectors form a basis for $\text{Nul} A$
6	$\dim \text{Row} A$	Number of non-zero rows in an echelon form of A
7	Basis for $\text{Row} A$	Non-zero rows in an echelon form of A
8	$\text{Rank} A$	Number of pivot columns of A
9	$\text{Nullity} A$	Number of non-pivot columns of A

Problem 14.17. Let $S = \{u_1, u_2, u_3\}$ and $W = \text{span}(S)$, where $u_1 = (7, -4, -2, 9)^t$, $u_2 = (-4, 5, -1, -7)^t$, $u_3 = (-9, 4, 4, -7)^t$. Find $\dim W$. Does $b \in W$, where $b = (-9, 7, 4, 8)^t$. If yes, express b in terms of u_1, u_2, u_3 .

Solution: Let $A = [u_1 \ u_2 \ u_3]$
 $W = \text{Span} S = \{AX | X \in \mathbb{R}^3\}$, $b \in W$ if there exists some $X \in \mathbb{R}^3$
such that $AX = b \dots (1)$
i.e. if the equation $AX = b$ has a solution. We reduce $[A|b]$ to echelon form.

$$\begin{aligned}
 [A|b] &= \begin{pmatrix} 7 & -4 & -9 & \vdots & -9 \\ -4 & 5 & 4 & \vdots & 7 \\ -2 & -1 & 4 & \vdots & 4 \\ 9 & -7 & -7 & \vdots & 8 \end{pmatrix} \\
 &\sim \begin{pmatrix} -2 & -1 & 4 & \vdots & 4 \\ 0 & \textcircled{1} & 0 & \vdots & 3 \\ 0 & 0 & \mathbf{2} & \vdots & 11 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix} \\
 &= [A|B] \text{ (say)}
 \end{aligned}$$

Since every column of B has a pivot element, therefore u_1, u_2, u_3 are linearly independent. Hence, $\dim W = \dim \text{Col} A = 3$. Since augmented column does not have a pivot, $\therefore (1)$ has a solution. Thus b is a linear combination of u_1, u_2, u_3 i.e. $b \in W$. To find the linear combination, we find the solution of (1). For this, reduce $[A|b]$ to reduced echelon form. Then

$$[A|b] \sim \begin{pmatrix} 1 & 0 & 0 & 15/2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 11/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus, solution is

$$x_1 = 15/2, x_2 = 3, x_3 = 11/2$$

$$\text{Hence } b = \frac{15}{2}u_1 + 3u_2 + \frac{11}{2}u_3$$

Problem 14.18. Let W be the subspace of \mathbb{R}^4 generated by $S = \{v_1, v_2, v_3\}$

$$\text{where } v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ 3 \\ -1 \\ 0 \end{pmatrix}$$

1. Find a subset T of S which forms a basis of W .
2. Extend T to form a basis of \mathbb{R}^4

Solution:

1. Let $A = [v_1 \ v_2 \ v_3]$
We reduce A to Echelon form.

$$\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Applying row operations, we get

$$A \sim \begin{pmatrix} \textcircled{1} & 0 & 2 \\ 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The columns of A corresponding to the first two pivot columns, namely v_1, v_2 are linearly independent.

Thus, $T = \{v_1 \ v_2\}$ is a basis of W .

2. We extend T to form a basis of \mathbb{R}^4 .
 $\text{Span}(T) = \{\alpha v_1 + \beta v_2 | \alpha, \beta \in \mathbb{R}\} = \{(\alpha, \alpha + \beta, -\beta, 0) | \alpha, \beta \in \mathbb{R}\}$ Observe

$$\text{that } u_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in \text{Span } T.$$

Let $\mathcal{B}' = T \cup \{u_1\} = \{v_1, v_2, u_1\}$. Then \mathcal{B}' is a set linearly independent vectors. $\text{Span}(\mathcal{B}') = \{\alpha v_1 + \beta v_2 + \gamma v_3 | \alpha, \beta, \gamma \in \mathbb{R}\}$
 $= \{(\alpha + \gamma, \alpha + \beta + 2\gamma, -\beta - \gamma, \gamma) | \alpha, \beta, \gamma \in \mathbb{R}\}$.

Further it is easy to see that $u_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \notin \text{Span}(\mathcal{B}')$, and that v_1, v_2, u_1, u_2

are linearly independent vectors in \mathbb{R}^4 . Hence $\{v_1, v_2, u_1, u_2\}$ is a basis of \mathbb{R}^4 . Thus T is extended to form a basis of \mathbb{R}^4 .

Problem 14.19. Find linearly independent rows of matrix A given by

$$A = \begin{pmatrix} 1 & 1 & 0 & 3 \\ -2 & 0 & -2 & -4 \\ 4 & 4 & 0 & 12 \\ 1 & 0 & 1 & 2 \end{pmatrix}$$

Solution: Let $A = \begin{pmatrix} 1 & 1 & 0 & 3 \\ -2 & 0 & -2 & -4 \\ 4 & 4 & 0 & 12 \\ 1 & 0 & 1 & 2 \end{pmatrix}$.

The linearly independent rows of A are the linearly independent columns of A^t . To find these columns, we reduce A^t to Echelon form.

$$A^t \sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The first two columns of A^t are linearly independent, so that the 1st two rows of A are linearly independent.

Problem 14.20. If A is an $m \times n$ matrix, prove that $\text{Rank} A^t = \text{Rank} A$.

Solution: $\text{Rank} A^t = \text{Column Rank } A^t = \text{Row Rank } A = \text{column rank } A = \text{Rank } A$

Hence, $\text{Rank} A^t = \text{Rank } A$

Problem 14.21. Verify the Rank Nullity theorem for the matrix

$$A = \begin{pmatrix} 3 & 2 & 10 & 1 & 3 \\ 2 & -2 & 0 & 0 & 4 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 2 & 3 \\ 3 & 0 & 6 & 0 & 3 \end{pmatrix}$$

Solution: Step 1 Transform A to reduced echelon form.

$$A \sim \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B \text{ (say)}$$

The pivots have been encircled. Number of free variables = 2.

Step 2 $\text{Rank } A = \text{number of non-zero rows of } B = 3$. $\text{Nullity } A = \text{dimNul } A = \text{number of free variable} = \text{number of non-pivot columns of } B = 2$. $\therefore \text{Rank } A + \text{Nullity } A = 3 + 2 = 5 = \text{number of columns in } A$. Hence verified.

Problem 14.22. Find two different bases of \mathbb{R}^3 , which contain the vectors $(1, 1, 0), (1, -1, 2)$.

Solution: Let $v_1 = (1, 1, 0), v_2 = (1, -1, 2)$. It can be easily seen that $\{v_1, v_2\}$ is linearly independent over \mathbb{R} . Let $S = \{v_1, v_2\}$, $\text{Span}(S) = \{\alpha v_1 + \beta v_2 | \alpha, \beta \in \mathbb{R}\} = \{(\alpha + \beta, \alpha - \beta, 2\beta) | \alpha, \beta \in \mathbb{R}\}$. Taking $\alpha = \beta = 1$, we get

$$\alpha + \beta = 2, \alpha - \beta = 0, 2\beta = 2$$

Thus, $(2, 0, 1), (2, 0, 0) \notin \text{Span}(S)$. So that $S_1 = S \cup \{(2, 0, 1)\}, S_2 = S \cup \{(2, 0, 0)\}$ are linearly independent sets. Hence, S_1, S_2 are two bases of \mathbb{R}^3 containing v_1, v_2 .

14.9 Exercise

- Let $S = \{(0, 0, 0, 3), (1, 1, 0, 0), (0, 1, -1, 0)\}$. Find a basis for \mathbb{R}^4 containing maximum possible elements of S .
- Let $S = \{(1, 2, -1, 0), (1, 1, 0, 3), (0, -1, 1, 3)\}$. Find a basis for \mathbb{R}^4 containing maximum possible elements of S .
- Can the set $S = \{(1, 0, 0, 0), (0, 1, 0, 0), (1, -1, 0, 0)\}$ be extended to form a basis for \mathbb{R}^4 ?
- Extend the set S to form 2 different bases of \mathbb{R}^3 where

$$S = \{(0, 0, 1), (0, 1, 1)\}$$

- Prove that P_5 cannot be spanned by any set containing five elements of P_5 .
- Find the dimension of the subspace spanned by the vectors $(3, 1, 1, 5)^t, (4, -2, -4, -1)^t, (-6, 3, 4, 1)^t$ and $(1, -2, -1, -2)^t$.
- Extend the subset S to form a basis of V by adding appropriate elements of the standard basis.
 - $S = \{(1, 3, 2, 1)^t, (2, -1, -2, -1)^t, (-1, 2, 3, 1)^t\}; V = \mathbb{R}^4$.
 - $S = \{1 + x + x^2 + x^3, 1 + 2x + 3x^2 + 4x^3, 1 + 3x + 6x^2 + 10x^3\}; V = P_3$.
 - $S = \{(1, 2, 3)^t, (0, 2, 5)^t\}; V = \mathbb{R}^3$.

- Find the dimensions of the following subspaces of $M_{2 \times 3}(\mathbb{C})$ over \mathbb{C} :

$$(i) W_1 = \left\{ \begin{pmatrix} a & b & c \\ d & 0 & 0 \end{pmatrix} : a = b + c, c = b + d \right\}.$$

$$(ii) W_2 = \left\{ \begin{pmatrix} a & b & c \\ d & e & 0 \end{pmatrix} : a + c = d + e \right\}.$$

$$(iii) W_3 = \left\{ \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} : a + b = d + e, c = f \right\}.$$

- Let $W_1 = \{(x_1, x_2, 0, 0)^t | x_1, x_2 \in \mathbb{R}\}$, and $W_2 = \{(0, x_2, x_3, x_4)^t | x_2, x_3, x_4 \in \mathbb{R}\}$ be subspaces of \mathbb{R}^4 . Find $\dim(W_1 \cap W_2), \dim(W_1 + W_2)$.
- Prove or disprove that "Every subspace of an infinite dimensional vector space infinite dimensional"?
- Give an example of an infinite dimensional subspace of an infinite dimensional vector space.

- Let $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : 2x + y + z = 0, x - y - z = 0 \right\}$. Find a basis for W .

13. If $W = \text{Span} \{(3, 8, -3)^t, (1, -2, 5)^t, (2, 3, 1)^t\}$, find $\dim W$. Is $W = \mathbb{R}^3$? If not,
- give an element of \mathbb{R}^3 which is not in W .
 - extend this basis of W to a basis of \mathbb{R}^3 .
14. Find a basis for the subspace

$$W = \begin{pmatrix} 2a + 3b \\ a \\ b \end{pmatrix} : a, b \in \mathbb{R}$$

Hence, find $\dim W$.

15. Find a basis for the subspace

$$W = \begin{pmatrix} a - b + 2c \\ 2a + 3b + 4c \\ b \end{pmatrix} : a, b, c \in \mathbb{R}.$$

Hence, find $\dim W$.

16. Do the vectors p_1, p_2, p_3, p_4 span P_3 , where $p_1 = 1 + 4x, p_2 = -1 + x + 3x^2 + x^3, p_3 = 2 + x^2, p_4 = -3 + 2x + 4x^2 + 2x^3$
17. The echelon form of a matrix A is given. Find $\dim \text{Col}A$, $\dim \text{Nul}A$, $\dim \text{Row}A$ and $\text{Rank } A$.

(i) $\begin{pmatrix} \blacksquare & * & * & 0 \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{pmatrix}$

(ii) $\begin{pmatrix} \blacksquare & * & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

(iii) $\begin{pmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{pmatrix}$

(iv) $\begin{pmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{pmatrix}$

(v) $\begin{pmatrix} 0 & * \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

18. The echelon form of a matrix A is given. Find $\dim \text{Col}A$, $\dim \text{Nul}A$, $\dim \text{Row}A$ and $\text{Rank } A$.

(i) $\begin{pmatrix} 2 & 3 & 4 & -5 & 0 \\ 0 & 0 & -3 & 6 & 0 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

$$(ii) \begin{pmatrix} 0 & 2 & 0 & 5 & 0 & -2 \\ 0 & 0 & 0 & -3 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 5 & 1 & 0 & -9 & 2 \\ 0 & 2 & 6 & 0 & 3 \\ 0 & 0 & -3 & 2 & 0 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(v) \begin{pmatrix} 1 & -1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

19. In the above problem, find a basis for *Row A*, *Nul A* and *Col A*.

20. For the following matrices find a basis for *Row A*, *Col A* and *Nul A*.

$$(i) \begin{pmatrix} 0 & 2 & 1 \\ 3 & 1 & 4 \\ 2 & 4 & 6 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 0 & 1 & 2 & 4 & -1 \\ 1 & 3 & -1 & 0 & 5 \\ 2 & 0 & 4 & 1 & 3 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 1 & -1 & 2 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 1 & 1 & 2 & 0 \\ 2 & 2 & 4 & 0 \\ 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & -1 \\ 1 & 2 & 3 & -1 \end{pmatrix}$$

$$(v) \begin{pmatrix} 1 & 0 & 1 & 2 & 0 \\ 2 & 1 & -1 & 5 & 3 \\ -2 & 0 & 0 & -4 & 2 \\ 1 & 2 & -1 & 4 & 2 \end{pmatrix}$$

$$(vi) \begin{pmatrix} 1 & 2 & 5 & -2 & 7 \\ 2 & 3 & -2 & 4 & 1 \\ 5 & 1 & 0 & 2 & 1 \end{pmatrix}$$

21. If A is an $m \times n$ matrix, then prove that $Col A = \mathbb{R}^m$ if and only if the number of linearly independent columns of A is m .

22. If A is an $m \times n$ matrix, prove that
- A, A^t have the same rank.
 - $\dim \text{Col}A + \dim \text{Nul}A^t = \text{Number of rows in } A$.
23. Consider the linear system of equations $AX = 0$. Then, which of the following are true:
- $\dim \text{Row}A = \text{Number of free variables}$.
 - $\dim \text{Col}A = \text{Number of basic variables}$.
 - $\dim \text{Nul}A = \text{Number of free variables}$.
24. Let A be an $m \times n$ matrix. Then prove or disprove the following:
- $\text{Row}A, \text{Nul}A$ are subspaces of \mathbb{R}^n
 - $\text{Col}A$ and $\text{Nul}A$ are subspaces of \mathbb{R}^m
 - $\text{Row}A = \text{Col}A$
 - $\dim \text{Row}A = \dim \text{Col}A$
25. Let A and B be row equivalent matrices. Which of the following statements are true? Correct all the false ones.
- $\text{Col}A = \text{Col}B$
 - $\text{Row}A = \text{Row}B$
 - A basis for $\text{Col}A$ is also a basis for $\text{Col}B$
 - A basis for $\text{Row}A$ is also a basis for $\text{Row}B$
26. Can a finite dimensional vector space be generated by an infinite subset? Justify your answer.
27. Give an example in each of the following ones:
- A vector space over \mathbb{C} of dimension 1.
 - A vector space which is not finite dimensional.
 - A subspace of dimension 5 in P_5 .
 - A subspace of dimension 2 in \mathbb{R}^3 .
28. Can \mathbb{R}^4 have a linearly independent subset containing 5 elements. Justify your answer.
29. Verify the Rank Nullity theorem for the following matrices:

$$(i) \begin{pmatrix} 0 & 0 & 2 & 3 & 4 \\ 4 & 2 & -11 & 11 & 11 \\ 0 & 2 & 5 & -1 & 5 \\ 2 & 0 & -6 & 9 & 7 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 1 & 3 & 2 & 0 & 0 & 1 \\ 2 & 1 & -5 & 1 & 2 & 0 \\ 3 & 2 & 5 & 1 & -2 & 1 \\ 5 & 8 & 9 & 1 & -2 & 2 \\ 9 & 9 & 4 & 2 & 0 & 2 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 0 & 2 & 5 & -1 & 5 \\ 0 & 2 & 7 & 2 & 9 \\ 4 & 2 & -9 & 14 & 15 \\ 2 & 2 & -1 & 8 & 12 \end{pmatrix}$$

30. Find a basis for $Row A$ (a) consisting of vectors that are not the rows of A , and (b) consisting of vectors that are rows of A , where A is given by

$$(i) A = \begin{pmatrix} 1 & 6 & 3 & 8 \\ 2 & 4 & 6 & -1 \\ 3 & 10 & 9 & 7 \\ 4 & 16 & 12 & 15 \end{pmatrix}.$$

$$(ii) A = \begin{pmatrix} 1 & 3 & 4 & 7 \\ 2 & 4 & 5 & 8 \\ 3 & 1 & 2 & 3 \end{pmatrix}$$

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

31. Compute the row rank and the column rank of A by giving a basis for $Row A$ and $Col A$.

$$(i) A = \begin{pmatrix} 2 & -1 & -8 & -4 & 0 \\ 3 & 1 & -5 & -2 & 1 \\ 4 & 7 & 4 & 4 & 4 \end{pmatrix}.$$

$$(ii) A = \begin{pmatrix} -2 & -1 & -3 & -1 \\ 1 & 2 & 3 & -1 \\ 0 & 1 & 1 & -1 \end{pmatrix}.$$

$$(iii) A = \begin{pmatrix} 3 & 4 & -1 & -6 \\ 2 & 3 & 2 & -3 \\ 2 & 1 & -14 & -9 \\ 1 & 3 & 13 & 3 \end{pmatrix}.$$

32. Show that if $\dim V = n$, then no set of $n - 1$ vectors in V can span V .

14.10 Solved Problems

Problem 14.23. If W is a subspace of an n -dimensional vector space V such that $\dim W = \dim V$, prove that $W = V$.

Solution: $\dim V = n. \therefore \dim W = n$

Let B be a basis of W . Then $\text{span} B = W \dots (1)$

Then B is a linearly independent subset of V . Since $W \subset V, \therefore B$ is linearly independent of V . Also, B has n elements and $\dim V = n. \therefore B$ is a basis for V .
 $\therefore \text{Span } B = V$

Hence, $V = W$, using (1).

Problem 14.24. Prove that $\mathbb{R}[x]$ is an infinite dimensional vector space.

Solution: Let, if possible, $\mathbb{R}[x]$ be finite dimensional, say of dimension n . Now P_{n+1} is a subspace of $\mathbb{R}[x]$ and $\dim P_{n+1} = n + 2$. Since $n + 2 > n$, $\therefore \dim P_{n+1} > \dim \mathbb{R}[x]$... (1)

But, $\dim P_{n+1} \leq \dim \mathbb{R}[x]$, as P_{n+1} is a subspace of $\mathbb{R}[x]$. This contradicts (1). Hence, our assumption is wrong. So, $\mathbb{R}[x]$ is an infinite dimensional vector space.

Problem 14.25. Prove that $\mathbb{R}_{\mathbb{Q}}$ is not a finite dimensional vector space.

Solution: Let, if possible, \mathbb{R} be a finite dimensional vector space over \mathbb{Q} . Let $B = \{v_1, v_2, \dots, v_n\}$ be an ordered basis of \mathbb{R} . For any $v \in \mathbb{R}$,

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n, \alpha_i \in \mathbb{Q}, i = 1(1)n$$

so that

$$[v]_B = (\alpha_1 \alpha_2 \dots \alpha_n)^t$$

Hence, $[v]_B \in \mathbb{Q}^n$.

\mathbb{Q} is countable $\Rightarrow \mathbb{Q}^n$ is countable $\Rightarrow \{[v]_B : v \in \mathbb{R} \text{ is countable}\} \Rightarrow \mathbb{R}$ is countable. This contradicts the fact that \mathbb{R} is an uncountable set. Hence our assumption is wrong. So, \mathbb{R} is not finite dimensional over \mathbb{Q} . Hence \mathbb{R} is infinite dimensional over \mathbb{Q} .

Problem 14.26. If $\{v_1, v_2, \dots, v_n\}$ is a linearly independent subset of \mathbb{R}^n and A is any singular matrix, then prove that Av_1, Av_2, \dots, Av_n is linearly dependent.

Solution: Let $B = \{v_1, v_2, \dots, v_n\}$. Since B is a linearly independent subset of \mathbb{R}^n , and $\dim \mathbb{R}^n = n$. $\therefore B$ is a basis for \mathbb{R}^n . A is singular. \therefore there exists a non-zero vector $u \in \mathbb{R}^n$ such that $Au = 0$... (1).

Since B is a basis of \mathbb{R}^n $\therefore u = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n$... (2).

Since $u \neq 0$, \therefore at least one of the β_i 's is non-zero. From (1) and (2), $A(\beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n) = 0 \Rightarrow \beta_1 Av_1 + \beta_2 Av_2 + \dots + \beta_n Av_n = 0$. $\therefore \beta_i$'s are not all zero, $\therefore \{Av_1, Av_2, \dots, Av_n\}$ is linearly dependent.

14.11 Supplementary Exercises

1. State whether the following are true or false. Justify the false ones.

- (i) ϕ is a linearly independent set.
- (ii) $\mathcal{P}(A)$, the set of all subsets of A is a vector space.
- (iii) ϕ is a linearly independent set.
- (iv) $\{0\}$ is a linearly independent set.
- (v) $\{0\}$ has no basis.
- (vi) If V is spanned by a set containing n elements, and T is a set containing m elements, with $m > n$, then T must be linearly dependent.
- (vii) If $\dim V = n$ and B is a basis for V with m elements, then m must be less than n .
- (viii) The union of two linearly independent sets is linearly independent.
- (ix) The intersection of two linearly independent sets is linearly dependent.

- (x) If V is spanned by a subset having m elements, then $\dim V \leq m$
 - (xi) If $V = \text{span}(S)$, then every basis of V must be a subset of S .
 - (xii) If V is spanned by an infinite number of vectors then V is infinite dimensional.
 - (xiii) If $\dim V = n$, then for every $m, 1 \leq m \leq n, V$ has a subspace of dimension m .
 - (xiv) If $V = \text{span}(S)$ then some subset of S is a basis of V .
 - (xv) If W is a subspace of V and B is a basis of W , then every basis of V contains B .
 - (xvi) If $V = \text{span}(S)$, then S is a basis of V .
 - (xvii) If $V = \text{span}(S)$, and T is a linearly independent subset of S , then T is a basis of V .
 - (xviii) If V has a linearly independent subset containing m elements then $\dim V \leq m$.
 - (xix) If A is a 3×5 matrix then the possible dimensions of $\text{Col}A$ are 4 or 5.
 - (xx) If A is a 5×3 matrix then $\dim \text{Row}A$ is at most 3.
 - (xxi) If A is a 5×6 matrix then maximum possible $\dim \text{Nul}A$ is 1.
 - (xxii) For a 5×6 matrix, the maximum possible value of $\text{Rank} A$ is 6.
 - (xxiii) If A is a non-zero matrix, it is possible that A has no pivot columns.
 - (xxiv) If A is a 5×6 matrix having 4 pivot columns, then $\dim \text{Nul}A = 1$
 - (xxv) If A is a 5×6 matrix and $\dim \text{Nul}A = 2$, then $\dim \text{Col}A = 3$.
 - (xxvi) If A is a 2×3 matrix then A can have 3 pivot columns.
2. Find the dimension and a basis of the subspaces of \mathbb{R}^4 spanned by the vectors v_1, v_2, v_3, v_4, v_5 given by:
 - (a) $v_1 = (-1, 0, 3, -2)^t, v_2 = (0, 1, 2, -3)^t, v_3 = (3, 4, -1, -6)^t,$
 $v_4 = (-1, 3, 8, -7)^t, v_5 = (2, 1, -6, 9)^t$
 - (b) $v_1 = (1, 2, 1, 3)^t, v_2 = (2, 4, 2, 6)^t, v_3 = (5, 5, 0, 5)^t, v_4 = (11, 15, 4, 19)^t,$
 $v_5 = (-3, 2, 5, -2)^t$
 - (c) $v_1 = (-6, 4, -9, 4)^t, v_2 = (8, -3, 7, -3)^t, v_3 = (-9, 5, -8, 3)^t, v_4 =$
 $(4, 7, -8, 3)^t, v_5 = (2, 1, -2, 1)^t$
 3. Give an example of a linearly dependent set of vectors in \mathbb{C}^2 .
 4. Show that $u = (1 + i, 2i)^t, v = (1, 1 + i)^t \in \mathbb{C}^2$ are linearly dependent over \mathbb{C} are linearly independent over \mathbb{R} .
 5. Find a basis for the subspace of P_3 consisting of all vectors of the form $ax^3 + bx^2 + cx + d$ where $a = b$ and $c = d$.
 6. Find a basis for P_3 that includes the polynomials $p_1 = 1 + x, p_2 = x - x^2$.
 7. Consider the subset of the vector space of all real-valued functions

$$S = \cos^2 t, \sin^2 t, \cos 2t$$

Find a basis for the subspace $W = \text{span}(S)$. What is $\dim W$?

8. Find a basis for the given plane $2x - 5y + 7z = 0$.
9. In the vector space $M_{2 \times 2}(\mathbb{R})$,
 let $\beta = \left(\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \right)$
 and $v = \begin{pmatrix} 1 & 3 \\ -2 & 2 \end{pmatrix}$. Find the coordinate vector $[v]_\beta$.
10. In the vector space P_2 , find the coordinates of $P = 4t^2 - 2t + 3$ relative to basis $\beta = t^2 - t + 1, t + 1, t^2 + 1$.
11. Find the rank and nullity of the matrices given below:
- (i) $\begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 1 & -4 & -5 \\ 7 & 8 & -5 & -1 \\ 10 & 14 & -2 & 8 \end{pmatrix}$
- (ii) $\begin{pmatrix} 1 & -2 & 7 & 0 \\ 1 & -1 & 4 & 0 \\ 3 & 2 & -3 & 5 \\ 2 & 1 & -1 & 3 \end{pmatrix}$
12. Prove that the polynomials $1, 2x, -2 + 4x^2, -12x + 8x^3$ forms a basis of P_3 . Also, find the coordinates of $7 - 12t - 8t^2 + 12t^3$ relative to this basis.
13. Let $S = \{u, v, w\}$ be a subset of \mathbb{R}^3 . If every proper subset of S is linearly independent, then is S linearly independent. Justify your answer. Also, interpret your answer geometrically.
14. Give an example of subspaces W_1, W_2 of a finite dimensional vector space V such that $V = W_1 \oplus W_2$.
15. Let V be a finite dimensional vector space and W_1 a subspace of V . Prove that there exists a subspace W_2 of V such that $V = W_1 \oplus W_2$.
16. If A is an $m \times n$ matrix, then prove that $Col A = \mathbb{R}^m$ if and only if A has m linearly independent rows.
17. Find a basis for the following subspace W of a vector space $V(\mathbb{R})$. Also, find $\dim W$.
- (i) $V = \mathbb{R}^3, W = \{(x_1, x_2, x_3) | x_1 + 2x_2 + 3x_3 = 0, 2x_1 - x_2 = 0\}$
- (ii) $V = \mathbb{R}^n, W = \{(x_1, x_2, \dots, x_n) | x_1 + x_2 + \dots + x_n = 0\}$
- (iii) $V = P_n, W = \{a_0 + a_1x + \dots + a_nx^n | a_0 + a_1 + \dots + a_n = 0\}$
18. Let W be the subspace of \mathbb{R}^4 generated by $S = \{(1, -2, 5, -3), (2, 3, 1, -4), (3, 8, -3, -5)\}$.
- (i) Find a subset T of S which forms a basis of W .
- (ii) Extend T to form a basis of \mathbb{R}^4 .
19. Give an example of a non-zero subspace of a vector space V such that it contains no element of a basis of V .

20. If $W = \{f \in C(\mathbb{R}) \mid f(1/2) = 0\}$, then prove that W is not finitely generated.
21. V is a finite dimensional vector space and W is a subspace of V of dimension m . If $v \in V \sim W$ and $W_1 = W \cup v$, then find $\dim \text{span}(W_1)$.
22. If W_1, W_2 are two subspaces of a finite dimensional vector space V , then prove that

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$$

23. If $v_i \in \mathbb{R}^m$, $V = \text{Span}\{v_1, v_2, \dots, v_n\}$ and $A = [v_1 \ v_2 \ \dots \ v_n]$, then prove or disprove the following:

- A basis for $\text{Col}A$ is a basis for V .
- $\dim V = \text{Rank}A$.
- A basis for $\text{Row}A$ is also a basis for V .
- $\dim \text{Row}A = \dim V$.

24. Consider the system of linear equations $AX = b$ where A is an $m \times n$ matrix. Comment on the following statements:

- If all the rows of A are linearly independent then $AX = b$ has a solution for every $b \in \mathbb{R}^m$.
- If $AX = b$ has a solution for every $b \in \mathbb{R}^m$, then there are no free variables.
- If all the columns of A are linearly independent then $AX = b$ has a solution for every $b \in \mathbb{R}^m$.
- $\text{Col}A = \mathbb{R}^m$ is equivalent to saying that $AX = b$ has a solution for every $b \in \mathbb{R}^m$.

25. Prove that \mathbb{R} is a vector space over \mathcal{Q} . Is \mathbb{R} finite dimensional over \mathcal{Q} ?

26. Give examples of two non-zero 3×3 matrices A and B such that

- $\text{Rank}(A + B) = \text{Rank}(A) + \text{Rank}(B)$.
- $AB \neq 0$ and $\text{Rank}(AB) < \min(\text{Rank}(A), \text{Rank}(B))$.

27. Give an example of two non-zero 2×3 matrices of A and B such that $A + B \neq 0$ and $\text{Rank}(A + B) < \text{Rank}(A) + \text{Rank}(B)$.

28. (a) Use coordinate vectors to test the linear independence of the following sets in P_n , for suitable n .

- $3 - t + 4t^2, 2 - 5t^2, 8 - 2t + 7t^2$
- $2 + t^2, 1 + 2t, 1 + t + t^2$
- $1 + t, 2 + t^2, t + 2t^2 + t^3, t - t^2$

- (b) Use coordinate vectors to express $1 + 2t - 6t^2 + 2t^3$ as a linear combination of polynomials in (iii)

29. Suppose $S = \{v_1, v_2, \dots, v_m\}$ is a linearly independent set in \mathbb{R}^n and A is a singular matrix. Is $\{Av_1, Av_2, \dots, Av_m\}$ also linearly independent. Justify your answer.

14.12 Answers to Exercises

Exercise - 14.3

1. (ii).
2. (ii), (v), (vi), (viii) are linearly independent. Others are linearly dependent.
3.
 - (i) $\{(1, 2, 3), (2, 1, 4), (-1, -1, 2)\}$.
 - (ii) $\{1 + x + 3x^3, 1 + x^2 + 2x^3, -x^2 + x^3\}$.
 - (iii) Does not exist.
 - (iv) $\{2 + \sqrt{3}, 2 - \sqrt{3}\}$ or $\{2 + \sqrt{3}, -2 + \sqrt{3}\}$ etc.
 - (v) Does not exist.
 - (vi) T consists of any 2 vectors from S .
 - (vii) Any one of the last three vectors can be removed.
7. $x - 7y + 10z = 0$. Find the condition that points $(1, 3, 2), (-2, 4, 3), (x, y, z)$ are coplanar.
10. (i) No (ii) Yes.
11. Linearly dependent.
14. $v = v_1 + v_2 + v_3$.
15. $v = 3v_1 + 2v_2 + v_3$.
16. $p = 3p_1 + 2p_2 - 7p_3, q = 2p_1 + 3p_2 - 3p_3$.
17. $P = 2A - 2B + C - D$.
- 18) $v_2 = v_1 + v_3$.

Exercise - 14.6

2. (iii) No.
4. Take $v_1 = (1, 2, 2), v_2 = (1, 0, 0), v_3 = (0, 1, 0), v_4 = (0, 1, 1)$. Then

$$[v_1 \ v_2 \ v_3 \ v_4] \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Thus $\{v_1, v_2, v_3\}$ or $\{v_1, v_2, v_4\}$ is a basis.

5. $\mathcal{B}_1 = \{(1, 1, 0, 0)^t, (0, 0, 1, 1)^t, (1, 0, 0, 0)^t, (0, 0, 1, 0)^t\}$
 $\mathcal{B}_2 = \{(1, 1, 0, 0)^t, (0, 0, 1, 1)^t, (0, 1, 0, 0)^t, (0, 0, 0, 1)^t\}$.
6. $\{v_1, v_2\}, \{v_2, v_3\}, \{v_1, v_3\}$
7. $\mathcal{B} = \{(1, -5, 3, 4)^t, (0, 39, -20, -33)^t, (0, 0, 1, 0)^t\}$.
8. (ii) $(1, 0, 0), (0, 1, 0), (0, 1, 0)$
 (iii) $(\frac{3}{2}, \frac{5}{2}, \frac{-11}{2})^t, (\frac{-1}{2}, \frac{-1}{2}, \frac{3}{2})^t, (-1, -2, 5)^t$

9. (i) $(1, -2, 1)$.
 (ii) $(3, 2, -4)$.
 (iii) $(2, 3, -4)$.
 (iv) $(1, -2, 3)$.
10. $(-4, 8, -7)$.
11. (i) $(1, 6, -4)$.
 (ii) $(1, 3, -2)$.
 (iii) $(1, -1, 1)$.
12. (i) $(1, 0, 0), (0, 1, 0), (0, 0, 1)$.
 (ii) $(0, 1, 0), (1, 0, 0), (0, 0, 1)$.
 (iii) $(1/2, -1/2, 1/2), (1/2, 1/2, -1/2), (-1/2, 1/2, 1/2)$.
 (iv) $(0, -1, 1), (-1, 1, 0), (1, 0, 0)$.
13. (i) $(0, 2, 1)^t$.
 (ii) $(2, 0, 1)^t$.
 (iii) $(0, 1, 2)^t$.
 (iv) $(1, 2, 0)^t$.
14. (i) $\{1, i\}$.
 (ii) $\{(1, 0), (0, 1)\}$.
 (iii) $\{(1, 0), (0, 1), (i, 0), (0, i)\}$.
 (v) $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & i \end{pmatrix} \right\}$
 (v) $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$
 (vi) $\{1\}$
 (vii) $\{1, \sqrt{2}\}$.
15. Yes
16. *Hint:* $v_1 = 1, v_2 = 1 + x, v_3 = 1 + x^2$. If $A = [[v_1]_{\mathcal{B}}, [v_2]_{\mathcal{B}}, [v_3]_{\mathcal{B}}], b = [p]_r$, where r is the standard basis for \mathbb{P}_2 . If solution of $AX = b$ is $[-2, 5, -1]^t$, then $p = -2v_1 + 5v_2 - v_3$.
17. (i) $[p]_{\mathcal{B}} = (1, -2, 3)^t$.
 (ii) $p(x) = 4 - 9x + 5x^2$.
18. (i) $[u]_{\mathcal{B}} = (3 - 2i, 3 + 7i, -3 - i)^t$.
19. (i) $\{\alpha e_2 | \alpha \in \mathbb{R}\}, \{e_2\}$.
 (ii) $\{\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 + \alpha_4 e_4 | \alpha_i \in \mathbb{R} \ 1 \leq i \leq 4\}, e_1, e_2, e_3, e_4$.
 (iii) $2, 3, 4, 1. \dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$.
21. $a \neq 0, 1, -1$

Supplementary Exercises

1. (i) True
- (ii) False
- (iii) True
- (iv) False
- (v) False
- (vi) True
- (vii) False
- (viii) True
- (ix) False
- (x) True
- (xi) False
- (xii) False
- (xiii) True
- (xiv) True
- (xv) False
- (xvi) False
- (xvii) False
- (xviii) False
- (xix) False
- (xx) False
- (xxi) False
- (xxii) False
- (xxiii) False
- (xxiv) False
- (xxv) False
- (xxvi) False

8. *Hint:* We need basis for $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid 2x - 5y + 7z = 0 \right\}$

9. $\begin{pmatrix} -1 \\ 2 \\ -2 \\ 4 \end{pmatrix}$

10. $[p]_{\beta} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$

11. (a) Rank = 2, Nullity = 2.
- (b) Rank = 3, Nullity = 1.

13. No. $u = (1, 1, 1); v = (1, 0, 1); w = (0, 1, 0)$. As $u - v - w = 0$, S is linearly dependent

14. Consider $V = \mathbb{R}^2, W_1 = \text{span}((1, 0)), W_2 = \text{span}((0, 1))$. Then $V = W_1 \oplus W_2$.

15. Let B_1 be a basis for W_1 . Extend it to form a basis B for V . Then $B \sim B_1$ is a basis for a subspace W_2 such that $V = W_1 \oplus W_2$.

17. One possible answer is as follows:

(a) $(3, 6, -5), \dim W = 1$

(b) $(1, 0, \dots, -1), (0, 1, \dots, -1), \dots, (0, \dots, 1, -1), \dim W = n - 1$

(c) $1 - x, 1 - x^2, \dots, 1 - x^n, \dim W = n$

19. $V = \mathbb{R}^2, W = \text{span}[(1, 1)^t]; e_1, e_2 \notin W$.

20. *Hint:* Let $W_n = \{p(x) \in P_n(x) \mid p(1/2) = 0\}$. A basis for W_n is $(x - \frac{1}{2}), (x - \frac{1}{2})^2, \dots, (x - \frac{1}{2})^n$. Then W_n is a subspace of V and $\dim W_n = n$. If $\dim V = m$, then $\dim W_{m+1} = m + 1 < \dim V$ is a contradiction.

21. $m + 1$

23. (a) True

(b) True

(c) False

(d) True.

24. (a) True.

(b) False. For a 3×5 matrix, there are 2 free variables.

(c) False. For a 3×2 matrix, only 2 pivots. So one row does not have a pivot.

(d) True.

25. No

26. (a) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$$\text{Rank}(A) = \text{Rank}(B) = 1, \text{Rank}(A + B) = 2$$

(b) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$\text{Rank}(A) = \text{Rank}(B) = 2, \text{Rank}(AB) = 1$$

27.

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

Chapter 15

Linear Transformation

In chapter 12 we discussed the linear transformations of \mathbb{R}^2 and \mathbb{R}^3 . We now extend this concept for a general vector space.

15.1 Definitions and Examples

Definition 15.1. Let V and W be vector spaces over the same field F . A mapping

$$T : V \rightarrow W$$

is called a linear transformation from V into W if it preserves vector addition and scalar multiplication, i.e.

$$\begin{aligned} T(u + v) &= T(u) + T(v) \\ T(\alpha u) &= \alpha T(u), \forall u, v \in V, \alpha \in F \end{aligned}$$

Example 15.1. Let $V(F)$ and $W(F)$ be vector spaces. Then $T : V \rightarrow W$ defined by $T(v) = O_W \quad \forall v \in V$ where O_W is the zero of W , is a linear transformation. It is called zero transformation as each element is mapped to zero. The transformation $I : V \rightarrow V$ defined by $I(v) = v \quad \forall v \in V$ is also a linear transformation. It is called identity transformation. These two are the trivial transformations on any vector space.

Example 15.2. Let $M_{2 \times 3}(\mathbb{C})$ be the space of all 2×3 matrices with complex entries. Define

$$f : M_{2 \times 3}(\mathbb{C}) \rightarrow M_{3 \times 2}(\mathbb{C}) \text{ by } f(A) = A^t$$

Let $A, B \in M_{2 \times 3}(\mathbb{C}), \alpha \in \mathbb{C}$

$$\begin{aligned} \text{Then } f(A + B) &= (A + B)^t \\ &= A^t + B^t \\ &= f(A) + f(B) \\ f(\alpha A) &= (\alpha A)^t \\ &= \alpha A^t \\ &= \alpha f(A) \end{aligned}$$

Hence f is a linear transformation.

Example 15.3. Let $V = \mathbb{C}(\mathbb{R})$. Define

$$T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R}) \text{ by } T(a + ib) = a - ib$$

Then T is a linear transformation, because if $u, v \in \mathbb{C}, u = a + ib, v = c + id$, then

$$\begin{aligned} T(u + v) &= T((a + c) + i(b + d)) \\ &= a + c - i(b + d) \\ &= a - ib + c - id \\ &= T(u) + T(v) \\ \text{For } \alpha \in \mathbb{R}, T(\alpha u) &= T(\alpha a + i\alpha b) \\ &= \alpha a - i\alpha b \\ &= \alpha(a - ib) \\ &= \alpha T(u) \end{aligned}$$

Hence T is a linear transformation.

Example 15.4. Let V be the set of all differentiable functions from \mathbb{R} into \mathbb{R} . Then V is a vector space over \mathbb{R} . Define

$$\mathbb{T} : V \rightarrow V$$

by $T(f(x)) = \int_0^x f(t)dt$. Then T is a linear transformation, as linearity of integration is one of its fundamental properties.

Example 15.5. Let $\mathbb{C}[x]$ be the set of all polynomial in x whose coefficients are complex numbers. Define

$$T : \mathbb{C}[x] \rightarrow \mathbb{C}[x]$$

by $T(ax^2 + bx + c) = (a + b + c)x$

Let $p_1(x), p_2(x) \in \mathbb{C}[x]$ and $\alpha \in \mathbb{C}$. Let

$$\begin{aligned} p_1(x) &= a_1x^2 + b_1x + c_1 \\ p_2(x) &= a_2x^2 + b_2x + c_2 \\ T(p_1(x) + p_2(x)) &= T((a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2)) \\ &= (a_1 + a_2 + b_1 + b_2 + c_1 + c_2)x \\ &= (a_1 + b_1 + c_1)x + (a_2 + b_2 + c_2)x \\ &= T(p_1(x)) + T(p_2(x)) \\ T(\alpha p_1(x)) &= T(\alpha a_1x^2 + \alpha b_1x + \alpha c_1) \\ &= (\alpha a_1 + \alpha b_1 + \alpha c_1)x \\ &= \alpha(a_1 + b_1 + c_1)x \\ &= \alpha T(p_1(x)) \end{aligned}$$

Hence T is a linear transformation.

Example 15.6. Consider the vector space \mathbb{C} over \mathbb{C} .

Define $T : \mathbb{C} \rightarrow \mathbb{C}$ by $T(\alpha + i\beta) = \alpha - i\beta$. Take $u = 1 + i \in \mathbb{C}, i \in \mathbb{C}$

$$\begin{aligned} T(iu) &= T(i - 1) = -i - 1 \\ iT(u) &= i(1 - i) = i + 1 = 1 + i \end{aligned}$$

Hence $T(iu) \neq iT(u)$ so that T is not a linear transformation.

The following properties are immediate consequences of the definition.

Theorem 15.1. *Let $T : V \rightarrow W$ be a linear transformation. Then*

- (i) $T(0_V) = 0_W$
- (ii) $T(-u) = -T(u), \quad \forall u \in V$

Proof: Since T is a linear transformation, $\therefore T(\alpha u) = \alpha T(u), \forall u \in V, \alpha \in F$.

- (i) Let $\alpha = 0, u \in V$
Then $T(0u) = 0T(u)$
 $\Rightarrow T(0_V) = 0_W$.

- (ii) Let $u \in V$, take $\alpha = -1$. As T is a linear transformation.
 $\therefore T(-u) = (-1)T(u)$
 $\Rightarrow T(-u) = -T(u).$ □

Remark 15.1. *If $T(0) \neq 0$ then we can say that T is not a linear transformation.*

Theorem 15.2. *A linear transformation $T : V \rightarrow W$ is one-to-one if and only if $T(v) = 0 \implies v = 0$.*

Proof: Suppose that T is one-to-one. Let $v \in V$ be such that

$$T(v) = 0$$

Since $T(0) = 0$ by theorem 15.1

$$\therefore T(v) = T(0)$$

since T is one-to-one

$$\therefore v = 0.$$

Hence the condition holds.

Conversely, let the condition holds. Let $v_1, v_2 \in V$ such that $T(v_1) = T(v_2)$

$$\text{then } T(v_1) - T(v_2) = 0$$

$$\implies T(v_1) + T(-v_2) = 0$$

$$\implies T(v_1 + (-v_2)) = 0$$

$$\implies T(v_1 - v_2) = 0$$

$$\implies v_1 - v_2 = 0 \quad (\text{using the given condition})$$

$$\implies v_1 = v_2$$

Hence T is one-to-one. □

We now give an equivalent condition for a mapping to be a linear transformation.

Theorem 15.3. *Let V and W be vector spaces over the same field F . A mapping*

$$T : V \rightarrow W$$

is a linear transformation if and only if $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) \quad \forall \alpha, \beta \in F, u, v \in V$

Proof: Let T be a linear transformation. Let $u, v \in V, \alpha, \beta \in F$

$$\begin{aligned} T(\alpha u + \beta v) &= T(\alpha u) + T(\beta v) \\ &= \alpha T(u) + \beta T(v) \end{aligned}$$

Hence the given condition holds. Conversely, let the given condition holds, i.e.

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v), \quad \forall u, v \in V, \alpha, \beta \in F \quad \dots(1)$$

Let $u, v \in V, \alpha \in F$

$$\begin{aligned} T(u + v) &= T(1 \cdot u + 1 \cdot v) \\ &= 1 \cdot T(u) + 1 \cdot T(v) \quad \text{using(1)} \\ &= T(u) + T(v) \end{aligned}$$

$$\begin{aligned} T(\alpha u) &= T(\alpha u + 0v) \\ &= \alpha T(u) + 0T(v) \quad \text{using(1)} \\ &= \alpha T(u) + 0_w \\ &= \alpha T(u) \end{aligned}$$

Hence T is a linear transformation. \square

In fact, we have another equivalent condition for a linear transformation.

Theorem 15.4. *Let V and W be vector spaces over the same field F . A mapping $T : V \rightarrow W$ is a linear transformation if and only if, $\forall u, v \in V, \alpha \in F$*

$$T(\alpha u + v) = \alpha T(u) + T(v)$$

Proof: Left to the reader. \square

Theorem 15.5. *Let V be a finite dimensional vector space over a field F and let $\mathcal{B} = \{v_1, v_2, v_3, \dots, v_n\}$ be an ordered basis for V . Let W be a vector space over the same field F and $\{w_1, w_2, w_3, \dots, w_n\}$ be any vectors in W . Then there is precisely one linear transformation T from V into W such that $Tv_i = w_i, \forall i, 1 \leq i \leq n$.*

Proof: *Existence*

To prove that there exists some such linear transformation, we proceed as follows:

Let $v \in V$. Since \mathcal{B} is an ordered basis for V , \therefore there exists unique $\alpha_1, \alpha_2, \dots, \alpha_n \in F$, such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$$

Then

$$\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_n w_n \in W$$

We define

$$T(v) = \alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_n w_n$$

Then T is a well-defined rule which associates to each $v \in V$ a vector $T(v) \in W$.

Clearly, $T(v_i) = w_i, i = 1, 2, \dots, n$. We prove that T is a linear transformation.

Let $u_1, u_2 \in V, \alpha, \beta \in F$. Then,

$$\begin{aligned} u_1 &= \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n \quad \text{for some } \alpha_1, \alpha_2, \dots, \alpha_n \in F \\ u_2 &= \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n \quad \text{for some } \beta_1, \beta_2, \dots, \beta_n \in F \\ \alpha u_1 + \beta u_2 &= \alpha(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) + \beta(\beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n) \\ &= (\alpha\alpha_1 + \beta\beta_1)v_1 + (\alpha\alpha_2 + \beta\beta_2)v_2 + \dots + (\alpha\alpha_n + \beta\beta_n)v_n \\ \therefore T(\alpha u_1 + \beta u_2) &= (\alpha\alpha_1 + \beta\beta_1)w_1 + (\alpha\alpha_2 + \beta\beta_2)w_2 + \dots + (\alpha\alpha_n + \beta\beta_n)w_n \\ &= \alpha(\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_n w_n) + \beta(\beta_1 w_1 + \beta_2 w_2 + \dots + \beta_n w_n) \\ &= \alpha T(u_1) + \beta T(u_2) \end{aligned}$$

Hence T is a linear transformation.

Uniqueness

Let T_1 be a linear transformation from V into W such that

$$T_1(v_i) = w_i, \forall i, 1 \leq i \leq n$$

We must prove that $T_1(v) = T(v)$ for all $v \in V$.

Let $v \in V$. Then there exists $\alpha_1, \alpha_2, \dots, \alpha_n \in F$ such that

$$\begin{aligned} v &= \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n \\ T_1(v) &= T_1(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) \\ &= \alpha_1 T_1 v_1 + \alpha_2 T_1 v_2 + \dots + \alpha_n T_1 v_n \quad \because T_1 \text{ is a linear transformation} \\ &= \alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_n w_n \\ &= T(v) \end{aligned}$$

Hence $T_1(v) = T(v) \quad \forall v \in V$, so that $T_1 = T$.

This proves that T is unique. \square

The above theorem tells us that there are many linear transformations from a finite dimensional vector space to another vector space. Moreover, it also gives us a way to find these linear transformations. If a mapping is defined on an ordered basis, then it can be extended linearly to the whole space.

Example 15.7. Let $V = \mathbb{R}^3$ and $W = \mathbb{R}(x)$. Then V and W are vector spaces over \mathbb{R} . $\mathcal{B} = \{e_1, e_2, e_3\}$ is an ordered basis for V , where $e_1 = (1, 0, 0)^t$, $e_2 = (0, 1, 0)^t$, $e_3 = (0, 0, 1)^t$. Define $T_1 : V \rightarrow W$, by

$$\begin{aligned} T_1(e_1) &= 1 \\ T_1(e_2) &= x \\ T_1(e_3) &= x^2 \end{aligned}$$

Then by Theorem 15.5, T_1 can be extended to \mathbb{R}^3 such that T_1 is a linear transformation.

Then $T_1(a, b, c) = a + bx + cx^2$.

Consider $x, x^3, x^5 \in W$. Define $T_2 : V \rightarrow W$, by

$$\begin{aligned} T_2(e_1) &= x \\ T_2(e_2) &= x^3 \\ T_2(e_3) &= x^5 \end{aligned}$$

Then T_2 can be extended to \mathbb{R}^3 such that T_2 is a linear transformation, Then $T_2(a, b, c) = ax + bx^3 + cx^5$.

We can thus define many more linear transformations on V .

Theorem 15.6 proves that there is a one-to-one correspondence between an n -dimensional vector space V and \mathbb{R}^n .

Theorem 15.6. (coordinate transformation) Let $V(\mathbb{R})$ be an n -dimensional vector space and \mathcal{B} be an ordered basis for V . Let $T : V \rightarrow \mathbb{R}^n$ be defined by

$$T(v) = (v)_{\mathcal{B}}$$

Show that T is a one-to-one linear transformation of V onto \mathbb{R}^n .

Proof: Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for V . For any $v \in V$, there exists unique scalars $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$ such that $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$.

$$\text{Then } [v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n.$$

Let $T : V \rightarrow \mathbb{R}^n$ be defined by $T(v) = [v]_{\mathcal{B}}$

Step 1 T is a linear transformation.

Let $u, v \in V$, $\alpha \in \mathbb{R}$. Then

$$[u]_{\mathcal{B}} = (\alpha_1, \dots, \alpha_n)^t$$

$$[v]_{\mathcal{B}} = (\beta_1, \dots, \beta_n)^t$$

$$\begin{aligned} \text{Then } \alpha u + v &= \alpha(\alpha_1 v_1 + \dots + \alpha_n v_n) \\ &\quad + (\beta_1 v_1 + \dots + \beta_n v_n) \\ &= (\alpha\alpha_1 + \beta_1)v_1 + (\alpha\alpha_2 + \beta_2)v_2 + \dots + (\alpha\alpha_n + \beta_n)v_n \end{aligned}$$

$$\begin{aligned} \therefore [\alpha u + v]_{\mathcal{B}} &= (\alpha\alpha_1 + \beta_1, \alpha\alpha_2 + \beta_2, \dots, \alpha\alpha_n + \beta_n)^t \\ &= (\alpha\alpha_1, \dots, \alpha\alpha_n)^t + (\beta_1, \dots, \beta_n)^t \\ &= \alpha(\alpha_1, \dots, \alpha_n)^t + (\beta_1, \dots, \beta_n)^t \\ &= \alpha[u]_{\mathcal{B}} + [v]_{\mathcal{B}} \end{aligned}$$

$$\therefore [\alpha u + v]_{\mathcal{B}} = \alpha[u]_{\mathcal{B}} + [v]_{\mathcal{B}}$$

$$\begin{aligned} \text{Now } T(\alpha u + v) &= [\alpha u + v]_{\mathcal{B}} \\ &= \alpha[u]_{\mathcal{B}} + [v]_{\mathcal{B}} \\ &= \alpha T u + T v \end{aligned}$$

Hence T is a linear transformation.

Step 2 T is one-to-one.

Let $v = \alpha_1 v_1 + \dots + \alpha_n v_n \in V$ such that $Tv = 0$. Then $[v]_{\mathcal{B}} = 0$

$$\begin{aligned} \Rightarrow (\alpha_1, \dots, \alpha_n)^t &= 0 = (0, 0, \dots, 0)^t \\ \Rightarrow \alpha_i &= 0, \quad \forall i = 1, 2, \dots, n \\ \Rightarrow v &= 0 \end{aligned}$$

Hence $Tv = 0 \Rightarrow v = 0$, so that T is one-to-one.

Step 3 T is onto.

Let $x = (\alpha_1, \alpha_2, \dots, \alpha_n)^t \in \mathbb{R}^n$. Define $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$. Then $v \in V$ and $T(v) = [v]_{\mathcal{B}} = x$. Thus for every $x \in \mathbb{R}^n$, there exists some $v \in V$ such that $Tv = x$. Hence T is onto.

Step 4 Steps 1–3 prove that T is a one-to-one linear transformation from V onto \mathbb{R}^n .

This transformation $T : V \rightarrow \mathbb{R}^n$ is called coordinate transformation of a vector space V . \square

Problem 15.1. Let $T : V(F) \rightarrow W(F)$ be a linear transformation. Prove that T is one-to-one if and only if T maps every linearly independent set of vectors of V to a linearly independent set of vectors in W .

Solution: Let T be one-to-one and let $S = \{v_1, v_2, \dots, v_k\}$ be a linearly independent subset of V . Consider S_1 be the image of S under T .

Then $S_1 = \{Tv_1, Tv_2, \dots, Tv_k\}$

Let $\alpha_1, \alpha_2, \dots, \alpha_k \in F$, such that

$$\alpha_1 Tv_1 + \alpha_2 Tv_2 + \dots + \alpha_k Tv_k = 0 \quad (15.1)$$

Since T is a linear transformation \therefore (15.1)

$$\Rightarrow T(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k) = 0 = T(0).$$

$$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0 \text{ as } T \text{ is one-to-one}$$

$$\alpha_1 = \alpha_2 = \dots = \alpha_k = 0 \text{ as } S \text{ is linearly independent}$$

Hence S_1 is linearly independent.

Thus T maps every linearly independent set to a linearly-independent set.

Conversely, let the condition hold. Let, if possible, T not be one-to-one. Then there exist $v_1, v_2 \in V$ such that $v_1 \neq v_2$ but $T(v_1) = T(v_2)$. Consider $w = v_1 - v_2$.

Then $w \neq 0$ but $T(w) = 0$. Thus $\{w\}$ is a linearly-independent set but $\{T(w)\} = \{0\}$ is a linearly dependent set which contradicts the hypothesis. Thus our assumption is wrong, so that T is one-to-one.

Corollary 15.7. *Let V and W be finite dimensional vector spaces such that $\dim V = \dim W$. Let $T : V \rightarrow W$ be a linear transformation. Then T is one-one if and only if T maps a basis of V onto a basis of W .*

Proof: Let $\dim V = \dim W = n$ (say)

Let T be one-to-one and let $B = \{v_1, v_2, \dots, v_n\}$ be a basis of V .

Let $B' = \{Tv_1, Tv_2, \dots, Tv_n\}$. Then, by Problem 15.1, B' is a linearly-independent set. Since B' contains n linearly independent elements and $n = \dim W$, $\therefore B'$ is a basis of W . Hence a basis of V is mapped to a basis of W .

Conversely, suppose that a basis is mapped to a basis. Let, if possible, T be not one-to-one. Thus, there exists $0 \neq v \in V$ such that $Tv = 0$. Now $\{v\}$ is a linearly-independent subset of V , so it can be extended to a basis for V , say $\{v_1, v_2, \dots, v_n\}$. By the given hypothesis $\{Tv, Tv_2, \dots, Tv_n\}$ is a basis for W , i.e. $\{0, Tv_2, \dots, Tv_n\}$ is a basis for W . This is a contradiction, as a linearly-independent set cannot contain the zero vector. Hence our assumption is wrong, so that T is one-to-one. \square

Problem 15.2. *Define $T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ by $T(z) = \text{Im } z$. Is T a linear transformation.*

Solution: Let $z_1, z_2 \in \mathbb{C}$, $\alpha \in \mathbb{R}$. Then,

$$\begin{aligned} T(\alpha z_1 + z_2) &= \text{Im}(\alpha z_1 + z_2) \\ &= \text{Im}(\alpha z_1) + \text{Im}(z_2) \\ &= \alpha \text{Im}(z_1) + \text{Im}(z_2) \quad \because \alpha \in \mathbb{R} \\ &= \alpha T(z_1) + T(z_2) \end{aligned}$$

Hence T is a linear transformation.

Problem 15.3. *Define $T : \mathbb{C}(\mathbb{C}) \rightarrow \mathbb{C}(\mathbb{C})$ by $T(z) = \text{Im } z$. Is T a linear transformation?*

Solution: Intuitively, we feel that $Im(iz) \neq iIm(z) \quad \forall z \in \mathbb{C}$.

Thus we take an example to prove this.

Let $v = i \in \mathbb{C}$, $\alpha = i \in \mathbb{C}$. Then,

$$\begin{aligned} T(\alpha v) &= T(-1) = 0 \\ \alpha T(v) &= i Im v = i(1) = i \\ \therefore T(\alpha v) &\neq \alpha T(v) \end{aligned}$$

Hence T is not a linear transformation.

Remark 15.2. The above problems shows that whether a mapping is a linear transformation depends also upon the field over which the vector space is taken.

Problem 15.4. Let $V = M_2(\mathbb{C})$. Define $T : V(\mathbb{C}) \rightarrow \mathbb{C}$ by

$$T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = a + b + c + d - 1$$

Is T a linear transformation?

Solution:

$$\begin{aligned} T\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) &= 0 + 0 + 0 + 0 - 1 \\ &= -1 \\ &\neq 0 \end{aligned}$$

Since $T(0) \neq 0$

$\therefore T$ is not a linear transformation.

Problem 15.5. Is there a linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that

$$\begin{aligned} T((2, 1)) &= (0, 3) \\ T((1, 2)) &= (3, 0) \\ T((1, 1)) &= (1, 2) \end{aligned}$$

Solution: Suppose T is a linear transformation, such that

$v_1 = (2, 1), v_2 = (1, 2), v_3 = (1, 1)$, and $T(v_1) = (0, 3), T(v_2) = (3, 0), T(v_3) = (1, 2)$

$$\begin{aligned} v_1 + v_2 &= (3, 3), \quad 3v_3 = (3, 3) \\ T(v_1 + v_2) &= T(v_1) + T(v_2) \\ &= (0, 3) + (3, 0) \\ \therefore T((3, 3)) &= (3, 3) \\ \text{Also } T(3v_3) &= 3T(v_3) \\ &= 3(1, 2) \\ \therefore T((3, 3)) &= (3, 6) \end{aligned}$$

\therefore Our assumption is wrong, so that no such linear transformation exists.

Problem 15.6. Let \mathbb{P}_1 be the vector space of polynomials in t of degree 1 over the field of real numbers \mathbb{R} .

$$T : \mathbb{P}_1 \rightarrow \mathbb{P}_1$$

such that

$$T(1+t) = t$$

$$T(1-t) = 1$$

Find $T((2-3t))$.

Solution: It can be easily seen that $2-3t = -\frac{1}{2}(1+t) + \frac{5}{2}(1-t)$. Since T is a linear transformation

$$\begin{aligned} \therefore T((2-3t)) &= -\frac{1}{2}T(1+t) + \frac{5}{2}T(1-t) \\ &= -\frac{1}{2}t + \frac{5}{2} \\ &= \frac{5-t}{2} \end{aligned}$$

Aliter: We illustrate another way of solving the problem. $\mathcal{B} = \{1, t\}$ is a basis of \mathbb{P}_1 . We first find $T(1)$ and $T(t)$. Then $T(p)$ for $p \in \mathbb{P}_1$ can be obtained.

Now

$$\begin{aligned} 1 &= \frac{1}{2}(1+t) + \frac{1}{2}(1-t) \\ \therefore T(1) &= \frac{1}{2}T((1+t)) + \frac{1}{2}T((1-t)) \\ &= \frac{t+1}{2} \\ t &= \frac{1}{2}(1+t) - \frac{1}{2}(1-t) \\ \therefore T(t) &= \frac{t-1}{2} \end{aligned}$$

$$\text{Now } T(2-3t) = 2T(1) - 3T(t) = \frac{5-t}{2}.$$

Problem 15.7. Find a linear transformation $T_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ such that

$$T_1(1, 2)^t = (1, 2, 3)^t$$

$$T_1(3, 4)^t = (4, 5, 6)^t$$

Solution:

$$\begin{aligned} \text{Let } \mathcal{B} &= \{(1, 2)^t, (3, 4)^t\} \\ &= \{v_1, v_2\} \text{ (say)}. \end{aligned}$$

Step 1 Express e_1, e_2 where $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ as a linear combination of v_1 and v_2 . If $\alpha_1 v_1 + \alpha_2 v_2 = e_1$ and $\beta_1 v_1 + \beta_2 v_2 = e_2$ where $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$

then $X_1 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$, $X_2 = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ are the solution of $AX = e_1$ and $AX = e_2$ respectively, where $A = [v_1 \ v_2]$. Then

$$[A \dot{\vdots} e_1 \dot{\vdots} e_2] = \begin{pmatrix} 1 & 3 & \vdots & 1 & \vdots & 0 \\ 2 & 4 & \vdots & 0 & \vdots & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \vdots & -2 & \vdots & \frac{3}{2} \\ 0 & 1 & \vdots & 1 & \vdots & \frac{-1}{2} \end{pmatrix} \text{ by ap-}$$

plying E -row operations.

$$\text{Thus } e_1 = -2v_1 + v_2, e_2 = \frac{3}{2}v_1 - \frac{1}{2}v_2.$$

Step 2 Let us first find $T_1(e_1), T_1(e_2)$, where $e_1 = (1 \ 0)^t, e_2 = (0 \ 1)^t$.

$$\begin{aligned} e_1 &= -2v_1 + v_2 \\ e_2 &= \frac{3}{2}v_1 - \frac{1}{2}v_2 \\ \therefore T_1e_1 &= (2, 1, 0)^t \\ T_1e_2 &= \frac{1}{2}(-1, 1, 3)^t \end{aligned}$$

If $v = (x_1 \ x_2)^t \in \mathbb{R}^2$ then

$$\begin{aligned} v &= x_1e_1 + x_2e_2 \\ T_1v &= x_1T_1e_1 + x_2T_1e_2 \\ &= (2x_1 - \frac{1}{2}x_2, x_1 + \frac{1}{2}x_2, \frac{3}{2}x_2)^t \end{aligned}$$

15.2 Exercise

1. Are the mappings $T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ linear transformations?

- (i) $T(z) = \operatorname{Re}(z)$
- (ii) $T(z) = |\operatorname{Re}(z)|$
- (iii) $T(z) = \bar{z}$
- (iv) $T(z) = |z|$

2. Are the mappings

$$T : \mathbb{C}(\mathbb{C}) \rightarrow \mathbb{C}(\mathbb{C})$$

linear transformations?

- (i) $T(z) = \operatorname{Re}(z)$
- (ii) $T(z) = \bar{z}$
- (iii) $T(z) = 2z$

3. Let $V = M_2(\mathbb{C})$. Prove that $T : V(\mathbb{C}) \rightarrow (\mathbb{C})$, defined by $T(A) = \operatorname{trace} A$, is a linear transformation

4. Let $V = M_2(\mathbb{R})$. Is $T : V(\mathbb{R}) \rightarrow \mathbb{R}$ defined by $T(A) = \det A$ a linear transformation? Justify your answer.

5. Let $T : \mathbb{P}_2(\mathbb{R}) \rightarrow \mathbb{P}_1(\mathbb{R})$ be defined below. Test whether T is a linear transformation.

- (i) $T(at^2 + bt + c) = at + b$
- (ii) $T(at^2 + bt + c) = at + b + 1$
- (iii) $T(at^2 + bt + c) = ct + b$
- (iv) $T(at^2 + bt + c) = (b + c)t + a$

6. Let $V = M_2(\mathbb{C})$ be a vector space over \mathbb{C} . $T : V \rightarrow V$ be defined below. Test whether T is a linear transformation.

$$(i) \ T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a-b & 0 \\ 0 & c-d \end{pmatrix}$$

$$(ii) \quad T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \bar{a} & Re\ b \\ Im\ b & -\bar{d} \end{pmatrix}$$

$$(iii) \quad T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+i & b+i \\ c-i & d-i \end{pmatrix}$$

$$(iv) \quad T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+ib & 0 \\ c+id & 0 \end{pmatrix}$$

7. Prove that the mapping $T : e[0, 1] \rightarrow \mathbb{R}$ defined by

$$T(f) = \int_0^1 f(t)dt$$

is a linear transformation.

8. Let $V = M_2(\mathbb{R})$ and $B \in V$ be fixed. Define $T : V(\mathbb{R}) \rightarrow V(\mathbb{R})$ by

$$T(A) = BA - AB$$

Prove that T is a linear transformation.

9. Let $V = M_n(\mathbb{C})$. Define $T : V(\mathbb{C}) \rightarrow V(\mathbb{C})$ by

$$T(A) = A^\theta$$

Is T is a linear transformation?

10. Let $V(F)$ be a vector space and $b \in V$. Define $T : V \rightarrow V$ by

$$T(v) = v + b.$$

Find b such that T is a linear transformation.

11. Let $V = M_{2 \times 3}(\mathbb{C})$. Define $T : V(\mathbb{C}) \rightarrow V(\mathbb{C})$ by

$$T(A) = \begin{pmatrix} 1 & i \\ 0 & -i \end{pmatrix} A.$$

Prove that T is a linear transformation. Also find $T(B)$ where

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & -1 \end{pmatrix}.$$

12. If $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a linear transformation such that $T(1, 2) = (1, -2)$, $T(-1, 2) = (2, 0)$. Find $T((1, 0))$, $T((0, 1))$, $T((1, -2))$.

13. If $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a linear transformation such that

$$T(1, 0) = (a, b)$$

$$T(0, 1) = (c, d)$$

find $T((x_1, x_2))$.

14. Does there exist a linear transformation $T : \mathbb{P}_1 \rightarrow \mathbb{P}_1$ such that

$$T(1 - t) = 1, \quad T(1 + t) = t?$$

If yes, find it.

15. Find a linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $T(1, 2) = (3, 0)$, $T(-1, 2) = (0, 1)$.
16. Does there exist a linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $T(1, 2) = (3, 4)$, $T(-1, 2) = (5, 0)$, $T(1, -2) = (0, 0)$?
17. Give examples of 2009 linear transformations from \mathbb{R}^2 into \mathbb{R}^3 . How many linear transformations can be defined?
18. Prove that the image of a subspace W of a vector space V under a linear transformation $T : V \rightarrow W$ is a subspace of W .
19. Let $T : V \rightarrow W$ be a linear transformation. If $U = \text{Span}\{v_1, v_2, \dots, v_m\}$, prove that $T(U) = \text{Span}\{T(v_1), T(v_2), \dots, T(v_m)\}$.
20. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by $T((x \ y \ z)) = ((x + y \ y + z \ z + x))$. If $U = \text{Span}((1 \ 2 \ 3), (2 \ 3 \ 0))$. Find a basis for $T(U)$.
21. Let $T : V \rightarrow W$ be a linear transformation.
- If $\{v_1, v_2, \dots, v_m\}$ is a linearly independent subset of vectors in V , then can we say $\{T(v_1), T(v_2), \dots, T(v_m)\}$ is always linearly independent? Justify your answer.
 - If $\{T(v_1), T(v_2), \dots, T(v_m)\}$ is a linearly-independent subset of W , then is $\{v_1, v_2, \dots, v_m\}$ always linearly independent in V ? Prove or disprove.
22. Let $\{v_1, v_2, \dots, v_{10}\}$ be a basis of V and $W = \text{span}\{v_1, v_2, \dots, v_5\}$. Let $T : V \rightarrow V$ be a linear transformation such that

$$T(v_i) = v, \quad i = 1, 2, \dots, 5$$

for some fixed element v of V . Prove that $T(W) = \text{Span}(\{v\})$.

15.3 Range and Kernel

Definition 15.2. (Range): Let $T : V \rightarrow W$ be a linear transformation. The range of T is defined as the set of images of elements of V . We shall denote the range of T by $\text{Rng}(T)$. Thus

$$\text{Rng}(T) = \{T(v) \mid v \in V\}$$

Let us find the range of linear transformations defined in examples 15.1 to 15.5

Example 15.8. In Example 15.1, for the zero transformation O

$$\begin{aligned} \text{Rng}(O) &= \{O(v) \mid v \in V\} \\ &= \{0\} \end{aligned}$$

For the identity transformation

$$\begin{aligned} \text{Rng}(I) &= \{I(v) \mid v \in V\} \\ &= \{v \mid v \in V\} \\ &= V \end{aligned}$$

Example 15.9. Considering Example 15.2,

$$\begin{aligned} \text{Rng}(f) &= \{f(A) \mid A \in M_{2 \times 3}(\mathbb{C})\} \\ &= \{A^t \mid A \in M_{2 \times 3}(\mathbb{C})\} \\ &= M_{3 \times 2}(\mathbb{C}) \end{aligned}$$

Example 15.10. Considering Example 15.3,

$$\begin{aligned} \text{Rng}(f) &= \{T(a + ib) \mid a, b \in \mathbb{R}\} \\ &= \{a - ib \mid a, b \in \mathbb{R}\} \\ &= \mathbb{C} \end{aligned}$$

Example 15.11. Considering as in Example 15.4,

$$\begin{aligned} \text{Rng}(T) &= \{T(f) \mid f \in V\} \\ &= V \end{aligned}$$

$\therefore \text{Rng}(T) \subseteq V$. Conversely, let $f \in V$. If $\frac{d}{dx}f(x) = g(x)$, then

$$\begin{aligned} T(g(x)) &= \int_0^x g(t) dt \\ &= f(x) \end{aligned}$$

This implies that $f \in \text{Rng}(T)$. Hence $V \subseteq \text{Rng}(T) \subseteq V$.
Thus $\text{Rng}(T) = V$.

Example 15.12. Considering Example 15.5,

$$\begin{aligned} \text{Rng}(T) &= \{T(ax^2 + bx + c) \mid a, b, c \in \mathbb{C}\} \\ &= \{(a + b + c)x \mid a, b, c \in \mathbb{C}\} \\ &= \{kx \mid k \in \mathbb{C}\} \end{aligned}$$

Theorem 15.8. Let $T : V \rightarrow W$ be a linear transformation. Then $\text{Rng}(T)$ is a subspace of W .

Proof: $\text{Rng}(T) = \{T(v) \mid v \in V\}$. Since $T(0_V) = 0_W \therefore 0_W \in \text{Rng}(T)$. Hence $\text{Rng}(T) \neq \phi$.

Let $u, v \in \text{Rng}(T)$, $\alpha \in F$, then there exists $u_1, v_1 \in V$ such that $T(u_1) =$

$u, T(v_1) = v.$

Now

$$\begin{aligned}\alpha u + v &= \alpha T(u_1) + T(v_1) \\ &= T(\alpha u_1 + v_1) \quad \because T \text{ is a linear transformation} \\ &\in \text{Rng}(T)\end{aligned}$$

$\therefore \alpha u + v \in \text{Rng}(T).$ Hence $\text{Rng}(T)$ is a subspace of W . □

Definition 15.3. (Rank): The rank of a linear transformation T is the dimension of $\text{Rng } T$. It is denoted by $\text{Rank } T$.

Definition 15.4. (Kernel) Let $T : V \rightarrow W$ be a linear transformation. The kernel of T is defined as the set of all elements of V which are mapped to the zero element of W . It is denoted by $\text{Ker}(T)$ or $N(T)$. Thus

$$\text{Ker}(T) = \{v \in V \mid T(v) = 0_W\}$$

Let us find the kernel of the linear transformations defined in Examples 15.1 to 15.5,

Example 15.13. In Example 15.1, for the zero transformation,

$$\begin{aligned}\text{Ker}(T) &= \{v \in V \mid T v = 0_W\} \\ &= V\end{aligned}$$

For the identity transformation

$$\begin{aligned}\text{Ker}(I) &= \{v \in V \mid I(v) = 0_V\} \\ &= \{v \in V \mid v = 0_V\} \\ &= \{0\}\end{aligned}$$

Example 15.14. Consider Example 15.2

$$\begin{aligned}\text{Ker}(f) &= \{A \in M_{2 \times 3}(\mathbb{C}) \mid f(A) = 0\} \\ &= \{A \in M_{2 \times 3}(\mathbb{C}) \mid A^t = 0\} \\ &= \{O_{2 \times 3}\}\end{aligned}$$

Example 15.15. Consider Example 15.3

$$\begin{aligned}\text{Ker}(f) &= \{a + ib \in \mathbb{C} \mid T(a + ib) = 0\} \\ &= \{a + ib \in \mathbb{C} \mid a - ib = 0\} \\ &= \{0\}\end{aligned}$$

Example 15.16. In Example 15.4,

$$\begin{aligned}\text{Ker}(T_1) &= \{f \mid T_1(f) = 0\} \\ T_1 f = 0 &\iff \int_0^x f(x) dx = 0 \\ &\iff f(x) = 0 \\ \therefore \text{Ker}(T_1) &= 0\end{aligned}$$

Example 15.17. In Example 15.5,

$$\begin{aligned} \text{Ker}(T) &= \{ax^2 + bx + c \in P_2(\mathbb{C}) \mid T(ax^2 + bx + c) = 0\} \\ &= \{ax^2 + bx + c \in P_2(\mathbb{C}) \mid (a + b + c)x = 0\} \\ &= \{ax^2 + bx + c \in P_2(\mathbb{C}) \mid a + b + c = 0\} \end{aligned}$$

Theorem 15.9. Let $T : V \rightarrow W$ be a linear transformation. Then the kernel of T is a subspace of V .

Proof: $\text{Ker } T = \{v \in V \mid T(v) = 0\}$. $\therefore T(0) = 0, \therefore 0 \in \text{Ker } T$. Hence $\text{Ker } T \neq \emptyset$. Let $u, v \in \text{Ker } T, \alpha \in F$. Then $T(u) = T(v) = 0$.

$$\begin{aligned} T(\alpha u + v) &= \alpha T(u) + T(v) \\ &= \alpha 0 + 0 \\ &= 0 \end{aligned}$$

Hence $\alpha u + v \in \text{Ker}(T)$. Thus $\text{Ker}(T)$ is a subspace of V . \square

Definition 15.5. (Nullity): The nullity of a linear transformation T is the dimension of $\text{Ker}(T)$. It is denoted by $\text{Nullity } T$.

Theorem 15.10 gives a condition for a linear transformation to be one-to-one.

Theorem 15.10. Let $T : V \rightarrow W$ be a linear transformation. Then the following conditions are equivalent:

- (i) T is one-to-one
- (ii) $\text{Ker } T = \{0\}$.

Proof: (i) \Rightarrow (ii)

Let T be one-to-one. Since $T(0) = 0 \in \text{Ker } T$. $\therefore 0 \in \text{Ker}(T) \Rightarrow \{0\} \subseteq \text{Ker } T$

Let $v \in \text{Ker } T$. $\therefore Tv = 0 \Rightarrow Tv = T0 \Rightarrow v = 0 \quad \therefore T$ is one-to-one.

Hence $\text{Ker}(T) \subseteq \{0\}$. $\therefore \text{Ker } T = \{0\}$.

(ii) \Rightarrow (i)

Let $v_1, v_2 \in V$ be such that $Tv_1 = Tv_2$. $\therefore T(v_1 - v_2) = 0$

$\Rightarrow v_1 - v_2 \in \text{Ker } T = \{0\}$

$\Rightarrow v_1 - v_2 = 0$

$\Rightarrow v_1 = v_2$

$\Rightarrow T$ is one-to-one \square

Problem 15.8. Let A be a $m \times n$ matrix and $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ a linear transformation defined by $T(v) = Av$. Prove that

- (i) $\text{Ker } T = \text{Null}(A)$.
- (ii) $\text{Rng } T = \text{Col}(A)$.

Solution:

(i)

$$\begin{aligned} \text{Ker } T &= \{v \in \mathbb{R}^n \mid Tv = 0\} \\ &= \{v \in \mathbb{R}^n \mid Av = 0\} \\ &= \text{Null}(A) \end{aligned}$$

(ii) Let C_1, C_2, \dots, C_n be the columns of A .

$$\begin{aligned} \text{Rng } T &= \{Tv \mid v \in \mathbb{R}^n\} \\ &= \{Av \mid v \in \mathbb{R}^n\} \\ &= \left\{ (C_1 \ C_2 \ \cdots \ C_n) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \mid \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R} \right\} \\ &= \{(\alpha_1 C_1 + \alpha_2 C_2 + \cdots + \alpha_n C_n) \mid \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}\} \\ &= \text{Span of } \{C_1, C_2, \dots, C_n\} \\ &= \text{Col}(A). \end{aligned}$$

Problem 15.9. Let $T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ defined by

$$T(z) = \text{Re}(z)$$

be a linear transformation.

Let $v_1 = 2 + 3i, v_2 = 2i, v_3 = -4, v_4 = 0$

- (i) Do $v_1, v_2, v_3, v_4 \in \text{Ker } T$?
 (ii) Find $\text{Ker } T$.
 (iii) Do $v_1, v_2, v_3, v_4 \in \text{Rng } T$?
 (iv) Find $\text{Rng } T$.

Solution:

- (i)
- $$\begin{aligned} T(v_1) &= \text{Re}(v_1) = 2 \neq 0 \quad \therefore v_1 \notin \text{Ker } T \\ T(v_2) &= \text{Re}(v_2) = 0 \quad \therefore v_2 \in \text{Ker } T \\ T(v_3) &= \text{Re}(-4) = -4 \neq 0 \quad \therefore v_3 \notin \text{Ker } T \\ T(v_4) &= 0 \quad \therefore v_4 \in \text{Ker } T \end{aligned}$$

Hence $v_2, v_4 \in \text{Ker } T$.

- (ii)
- $$\begin{aligned} \text{Ker } T &= \{v \in V \mid Tv = 0\} \\ &= \{v \in V \mid \text{Re}(v) = 0\} = \{x + iy \in \mathbb{C} \mid x = 0\} \\ &= \{iy \mid y \in \mathbb{R}\}. \end{aligned}$$

$\therefore \text{Ker } T = \{iy \mid y \in \mathbb{R}\}$.

(iii) $v \in \text{Rng}(T)$ if $\exists w \in V$ such that

$$\begin{aligned} T(w) &= v \\ \text{i.e. } \text{Re}(w) &= v \Rightarrow v \text{ is real} \\ \therefore v_1 &\notin \text{Rng } T, v_2 \notin \text{Rng } T \\ \because T(-4) &= -4 = v_3 \\ \therefore v_3 &\in \text{Rng } T \\ T(0) &= 0, \quad \therefore v_4 \in \text{Rng } T \\ \text{Hence } v_3, v_4 &\in \text{Rng } T. \end{aligned}$$

- (iv) As shown in (iii), $u \in \text{Rng}(T) \Rightarrow u$ is real, i.e. $v \in \mathbb{R}$
 Let $v \in \mathbb{R}$. Then
 $\text{Re}(v) = v$
 $\Rightarrow T(v) = v$
 $\Rightarrow v \in \text{Rng}(T)$. Hence $\mathbb{R} = \text{Rng}(T)$.

Problem 15.10. Let $T : \mathbb{P}_2(\mathbb{R}) \rightarrow \mathbb{P}_3(\mathbb{R})$ be a linear transformation defined by
 $T(ax^2 + bx + c) = (a - b)x^3 + (b - c)x^2 + (c - a)x + (c - a)$, where $a, b, c \in \mathbb{R}$

Find

- (i) $\text{Ker } T$.
 (ii) A basis for $\text{Ker } T$.
 (iii) $\text{Rng } T$.
 (iv) A basis for $\text{Rng } T$.

Solution:

- (i) Let $f = ax^2 + bx + c \in \mathbb{P}_2$. Then $f \in \text{Ker } T$

$$\begin{aligned} \Leftrightarrow & T(f) = 0 \\ \Leftrightarrow & (a - b)x^3 + (b - c)x^2 + (c - a)x + (c - a) = 0 \\ \Leftrightarrow & a - b = 0, b - c = 0, c - a = 0 \\ \Leftrightarrow & a = b = c \end{aligned}$$

$$\therefore \text{Ker } T = \{ax^2 + ax + a \mid a \in \mathbb{R}\} = \{a(x^2 + x + 1) \mid a \in \mathbb{R}\}.$$

- (ii) From (i), we get $\text{Ker } T = \text{span}\{x^2 + x + 1\}$. Thus basis for $\text{Ker } T$ is $\{x^2 + x + 1\}$.

- (iii) Let $g \in \text{Rng } T = \text{Span}\{T(1), T(x), T(x^2)\}$.

Let $g = px^3 + qx^2 + rx + s$. Then there exists $f \in \mathbb{P}_2(\mathbb{R})$ such that $Tf = g$.
 Let $f = ax^2 + bx + c$. Then $(a - b)x^3 + (b - c)x^2 + (c - a)x + (c - a) = px^3 + qx^2 + rx + s$

$$\begin{aligned} \therefore a - b &= p \\ b - c &= q \\ c - a &= r \\ c - a &= s. \end{aligned}$$

Thus the equations

$$\begin{aligned} a - b &= p \\ b - c &= q \\ -a &+ c = r \\ -a &+ c = s \end{aligned}$$

must have a solution.

Then the augmented matrix is

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & p \\ 0 & 1 & -1 & q \\ -1 & 0 & 1 & r \\ -1 & 0 & 1 & s \end{array} \right) = (A \mid d)$$

Reducing to echelon form

$$(A | d) \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & p \\ 0 & 1 & 1 & q \\ 0 & 0 & 0 & p+q+r \\ 0 & 0 & 0 & r-s \end{array} \right)$$

For the system of equation to be consistent, the augmented column must not be a pivot column. So, we must have

$$\begin{aligned} r - s &= 0 \\ p + q + r &= 0 \\ \text{i.e. } s &= r, p + q + r = 0 \end{aligned}$$

Thus $\text{Rng } T = \{px^3 + qx^2 - (p+q)x - (p+q) | p, q \in \mathbb{R}\}$

(iv) As in (iii), any element in $\text{Rng } T$ is of the form

$$\begin{aligned} &px^3 + qx^2 - (p+q)x - (p+q); p, q \in \mathbb{R} \\ \text{i.e. } &p(x^3 - x - 1) + q(x^2 - x - 1); p, q \in \mathbb{R} \\ \therefore \text{Rng } T &= \text{Span}(\{x^3 - x - 1, x^2 - x - 1\}) \end{aligned}$$

It is easy to verify that $x^3 - x - 1, x^2 - x - 1$ are linearly independent.
 $\therefore \{x^3 - x - 1, x^2 - x - 1\}$ is a basis for $\text{Rng } T$.

Problem 15.11. Let $V = M_2(\mathbb{R})$ and let $T : V(\mathbb{R}) \rightarrow V(\mathbb{R})$ be a linear transformation defined by $T(A) = A \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} A$

Find

- (i) $\text{Ker } T$.
- (ii) Basis for $\text{Ker } T$.
- (iii) $\text{Rng } T$.
- (iv) Basis for $\text{Rng } T$.

Solution: Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V(\mathbb{R})$, then

$$\begin{aligned} T(A) &= A \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} A \\ &= \begin{pmatrix} (b-c) & a-d \\ -(a-d) & -(b-c) \end{pmatrix} \end{aligned}$$

(i) Let

$$\begin{aligned} &\begin{pmatrix} a & b \\ c & d \end{pmatrix} &&= A \in \text{Ker } T \\ \Rightarrow &T(A) &&= 0 \\ \Rightarrow &\begin{pmatrix} (b-c) & a-d \\ -(a-d) & -(b-c) \end{pmatrix} &&= 0 \\ \Rightarrow &b-c=0, a-d &&= 0 \\ \Rightarrow &b=c, a=d && \end{aligned}$$

$$\therefore \text{Ker } T = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$$

(ii) From (i), any element A of $\text{Ker } T$ is $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$, $a, b \in \mathbb{R}$

$$\text{i.e. } a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\text{Hence } \text{Ker } T = \text{Span} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$\text{Span} \{A_1, A_2\}$ (say). Also A_1, A_2 are linearly independent. \therefore Basis for

$$\text{Ker } T \text{ is } \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

(iii) To find $\text{Rng } T$

If $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \text{Rng } T$, then there exists $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = A \in V$ such that

$$T(A) = B. \therefore \begin{pmatrix} b-c & a-d \\ -(a-d) & -(b-c) \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

Thus the equations

$$\begin{aligned} b-c &= p \\ a-d &= q \\ -(a-d) &= r \\ -(b-c) &= s \end{aligned}$$

must have a solution for a, b, c, d

$$\text{Thus } \begin{pmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \\ s \end{pmatrix}$$

must have a solution.

The echelon form of the augmented matrix is

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & q \\ 0 & 1 & -1 & 0 & p \\ 0 & 0 & 0 & 0 & r+q \\ 0 & 0 & 0 & 0 & p+s \end{array} \right)$$

For a solution to exist, the augmented column must not be a pivot column,

$$\begin{aligned} \therefore p+s &= 0 \\ r+q &= 0 \end{aligned}$$

$$\therefore s = -p, r = -q. \text{Rng } T = \left\{ \begin{pmatrix} p & q \\ -q & -p \end{pmatrix} \mid p, q \in \mathbb{R} \right\}$$

(iv) Any element in $\text{Rng } T$ is of the form $\begin{pmatrix} p & q \\ -q & -p \end{pmatrix}$

i.e. $p \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + q \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Hence

$$\begin{aligned} \text{Rng } T &= \text{Span} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\} \\ &= \text{Span} \{B_1, B_2\} \quad (\text{say}) \end{aligned}$$

Clearly B_1, B_2 are linearly independent. $\therefore \{B_1, B_2\}$ is a basis for $\text{Rng } T$.

Problem 15.12. Let $V = M_2(\mathbb{C})$ be a vector space over \mathbb{R} . Let $T : V \rightarrow V$ be a linear transformation defined by $T \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) = \begin{pmatrix} \bar{a} & \text{Re } b \\ \text{Im } c & -\bar{d} \end{pmatrix}$.

Find

- (i) $\text{Ker } T$.
- (ii) Basis for $\text{Ker } T$.
- (iii) $\text{Rng } T$.
- (iv) Basis for $\text{Rng } T$.
- (v) Nullity T .
- (vi) Rank T .

Solution:

(i) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $T(A) = \begin{pmatrix} \bar{a} & \text{Re } b \\ \text{Im } c & -\bar{d} \end{pmatrix}$

$$A \in \text{Ker } T$$

$$\Leftrightarrow T(A) = 0$$

$$\Leftrightarrow \begin{pmatrix} \bar{a} & \text{Re } b \\ \text{Im } c & -\bar{d} \end{pmatrix} = 0$$

$$\Leftrightarrow \bar{a} = 0, \text{Re}(b) = 0, \text{Im}(c) = 0, -\bar{d} = 0$$

$$\Leftrightarrow a = d = 0, b = ix, c = y \text{ for some } x, y \in \mathbb{R}$$

$$\Leftrightarrow A = \begin{pmatrix} 0 & ix \\ y & 0 \end{pmatrix}, x, y \in \mathbb{R}$$

$$\text{Ker } T = \left\{ \begin{pmatrix} 0 & ix \\ y & 0 \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

(ii) Any element of $\text{Ker } T$ is of the form $\begin{pmatrix} 0 & ix \\ y & 0 \end{pmatrix}$, where $x, y \in \mathbb{R}$

$$= x \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, x, y \in \mathbb{R}$$

$$\text{Ker } T = \text{Span} \left\{ \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$$

Also, $\begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ are linearly independent over \mathbb{R} . \therefore A basis for

$$\text{Ker } T \text{ is } \left\{ \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}.$$

(iii) If $B = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \in \text{Rng } T$,

then there exists $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in V$ such that $T(A) = B$

$$\therefore \begin{pmatrix} \bar{a} & Re\ b \\ Im\ c & -\bar{d} \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

$$\Rightarrow \bar{a} = x, Re(b) = y, Im(c) = z, -\bar{d} = w$$

$$\Rightarrow a = \bar{x}, d = -\bar{w}, y, z \text{ is real, } z \text{ is real.}$$

Thus we must have x, w can be any complex numbers, whereas y, z are real.

$$\therefore Rng\ T = \left\{ \begin{pmatrix} x & y \\ z & w \end{pmatrix} \mid x, w \in \mathbb{C}, y, z \in \mathbb{R} \right\}.$$

(iv) From (iii) any element in the $Rng\ T$ is

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix}, x, w \in \mathbb{C}, y, z \in \mathbb{R}$$

$$= \begin{pmatrix} a + ib & y \\ z & c + id \end{pmatrix}, a, b, c, d, y, z \in \mathbb{R}$$

$$a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} i & 0 \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + z \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & i \end{pmatrix}$$

Thus $Rng\ T =$

$$Span \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & i \end{pmatrix} \right\}$$

$Span \{v_1, v_2, v_3, v_4, v_5, v_6\}$ (say)

Also if $\alpha_i \in \mathbb{R}, 1 \leq i \leq 6$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 + \alpha_5 v_5 + \alpha_6 v_6 = 0$$

$$\Rightarrow \begin{pmatrix} \alpha_1 + i\alpha_5 & \alpha_2 \\ \alpha_3 & \alpha_4 + i\alpha_6 \end{pmatrix} = 0$$

$$\alpha_1 + i\alpha_5 = \alpha_2 = \alpha_3 = \alpha_4 + i\alpha_6 = 0$$

$$\alpha_i = 0, \forall i, 1 \leq i \leq 6$$

Hence, $\{v_1, v_2, \dots, v_6\}$ is linearly independent.

$\therefore \{v_1, v_2, \dots, v_6\}$ is a basis for $Rng\ T$.

(v) From (ii), basis for $Ker\ T$ has 2 elements.

$$\therefore dim\ Ker\ T = 2$$

$$Nullity\ T = dim\ Ker\ T = 2$$

$$Nullity\ T = 2$$

(vi) From (iv) a basis for $Rng\ T$ has 6 elements

$$\therefore dim\ Rng\ T = 6$$

$$Rank\ T = dim\ Rng\ T = 6.$$

15.4 Exercise

1. For the following linear transformations, find $Ker\ T$ and $Rng\ T$.

(i) $T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ defined by $T(z) = \bar{z}$

(ii) $T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ defined by $T(z) = 2z$

(iii) $T : \mathbb{P}_2(\mathbb{R}) \rightarrow \mathbb{P}_1(\mathbb{R})$ defined by $T(at^2 + bt + c) = ct + b$

2. Find $Ker\ T$ and $Rng\ T$ for Q7 Exercise 15.2.

3. Let $V = M_2(\mathbb{C})$. Let $T : V(\mathbb{C}) \rightarrow \mathbb{C}$ be a linear transformation defined by $T(A) = trace\ A$

- (i) Do $\begin{pmatrix} 1+2i & -4+i \\ 3-i & 1-2i \end{pmatrix}, \begin{pmatrix} 1+i & 1-i \\ -1+i & -1-i \end{pmatrix} \in \text{Ker } T$
- (ii) Does $2+3i, -2+3i \in \text{Rng } T$
- (iii) Is T one-to-one?
- (iv) Is T onto?
- Justify your answer.

4. Find $\text{Ker } T, \text{Rng } T$ for Q8 Exercise 15.2

5. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation defined by $T(x_1, x_2) = (x_1, 0)$.
- (i) Find $\text{Ker } T$.
- (ii) Find a basis for $\text{Ker } T$.
- (iii) Find $\text{Rng } T$.
- (iv) Find a basis for $\text{Rng } T$.

6. Let $T : \mathbb{P}_2(\mathbb{R}) \rightarrow \mathbb{P}_1(\mathbb{R})$ be a linear transformation defined by $T(\alpha x^2 + \beta x + \gamma) = (\alpha - \beta)t + (\gamma - \alpha)$
- (i) Does $1 + x + x^2 \in \text{Ker } T$?
- (ii) Does $3x + 8 \in \text{Rng } T$?
- (iii) Describe $\text{Ker } T$.
- (iv) Describe $\text{Rng } T$.

7. Let $T : \mathbb{P}_2(\mathbb{R}) \rightarrow \mathbb{P}_2(\mathbb{R})$ be a linear transformation defined by $T(\alpha x^2 + \beta x + \gamma) = (\alpha + \beta)x^2 + (\alpha + \beta + \gamma)$
- (i) Does $2x^2 - x - 1 \in \text{Ker } T$?
- (ii) Does $2x^2 - 2x \in \text{Ker } T$.
- (iii) Find a basis for $\text{Ker } T$.
- (iv) Find a basis for $\text{Rng } T$.

8. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a linear transformation defined by $T((x_1, x_2, x_3)) = (x_1 + x_2, x_2 + x_3)$. Find
- (i) $\text{Ker } T$.
- (ii) Basis for $\text{Ker } T$.
- (iii) $\text{Rng } T$.
- (iv) Basis for $\text{Rng } T$.

9. Let $V(\mathbb{R})$ be the space of all differentiable functions from \mathbb{R} into \mathbb{R} . Let $T : V(\mathbb{R}) \rightarrow V(\mathbb{R})$ be a linear transformation defined by

$$T(f(x)) = \frac{d}{dx}(f(x))$$

- (i) Does $2x + 3 \in \text{Ker } T$?
- (ii) Does $2x + 3 \in \text{Rng } T$?
- (iii) Find $\text{Ker } T$.
- (iv) Find $\text{Rng } T$.
10. Let $V = M_2(\mathbb{R})$. Let $T : V(\mathbb{R}) \rightarrow V(\mathbb{R})$ be a linear transformation defined by $T(A) = BA - AB$, where $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Find a basis for
- (i) $\text{Ker } T$.
- (ii) $\text{Rng } T$.

11. Let $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ be a linear transformation defined by

$$T(X) = AX, \text{ where } A = \begin{pmatrix} 1 & 2 & 3 & -1 \\ 1 & 1 & -1 & 1 \\ 2 & 3 & 2 & 0 \\ 1 & 1 & -1 & 1 \end{pmatrix}$$

Find

- (i) $\text{Ker } T$.
 (ii) $\text{Rng } T$.
12. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation defined by $T(v) = Av$, where $A = m \times n$. Prove that
 (i) $\text{Ker } T = \text{Null}(A)$.
 (ii) $\text{Rng } T = \text{Col}(A)$.
13. Let $T : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ be a linear transformation defined by

$$T((x_1, x_2, x_3, x_4)^t) = \begin{pmatrix} x_1 - x_2 + 2x_3 \\ x_1 + x_3 + x_4 \end{pmatrix}$$

Find

- (i) Nullity T .
 (ii) Rank T .
 (iii) Verify that Nullity $T + \text{Rank } T = 4$.
14. Let $T : (\mathbb{R}^3) \rightarrow (\mathbb{R}^3)$ be a linear transformation defined by

$$T \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = \begin{pmatrix} x_1 + x_2 - x_3 \\ x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$$

Find

- (i) Basis for $\text{Ker } T$.
 (ii) Nullity T .
 (iii) Basis for $\text{Rng } T$.
 (iv) Rank T
15. Let $T : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ be a mapping defined by

$$T(r(\cos \theta + i \sin \theta)) = r(\cos(\theta + \alpha) + i \sin(\theta + \alpha))$$

- (i) Prove that T is a linear transformation where α is a fixed real number.
 (ii) Is T onto?
 (iii) Is T injective?
 (iv) Interpret the transformation geometrically.
16. If $B = \{v_1, v_2, \dots, v_n\}$ is a basis for V and $T : V \rightarrow W$ is a linear transformation such that $T(v_i) = 0, \forall i = 1, 2, \dots, n$, then prove that T must be the zero transformation.

15.5 Answers to Exercises

Exercise - 15.2

1. (i) Yes
(ii) No
(iii) Yes
(iv) No
2. (i) No
(ii) No
(iii) Yes
3. Yes
4. No
5. (i) Yes
(ii) No
(iii) Yes
(iv) Yes
6. (i) Yes
(ii) No
(iii) No
(iv) Yes
7. Yes
9. No
10. $b = 0$
12. $(-\frac{1}{2} \ -1), (\frac{3}{4} \ -\frac{1}{2}), (-2 \ 0)$
13. $(x_1a + x_2c, x_1b + x_2d)$
14. Yes; $T(\alpha t + \beta) = \frac{\alpha+\beta}{2}t - \frac{\alpha-\beta}{2}$
15. $T(x_1, x_2) = (\frac{3}{2}x_1 + \frac{3}{4}x_2, -\frac{1}{2}x_1 + \frac{1}{4}x_2)$
16. No
17. $T(x_1, x_2) = (\alpha x_1, \alpha x_2, 0), 1 \leq \alpha \leq 2009$. *Infinitely many.*
20. $\{(3,5,4), (5,3,2)\}$
21. (i) No. Zero transformation
(ii) Prove it.
23. *Hint:* Define $T : V \rightarrow W$ $T(v) = 0, \forall v \in V$ Take $0 \neq v \in V$. Then v is l.i. but $T(v)$ is l.d.

Exercise - 15.4

1. (i) $\{0\}, \mathbb{C}$
(ii) $\{0\}, \mathbb{C}$
(iii) $\{at^2 \mid a \in \mathbb{R}\}, P_1(\mathbb{R})$
2. Zero function, \mathbb{R} .
3. (i) No, Yes
(ii) Yes, Yes
(iii) No
(iv) Yes
5. (i) $(0, x_2)^t : x_2 \in \mathbb{R}; \{(0, 1)^t\}$
(ii) $\{(x_1, 0)^t : x_1 \in \mathbb{R}\}$
(iii) $\{(1, 0)^t\}$
6. (i) Yes
(ii) Yes
(iii) $\text{Span}\{1 + t + t^2\}$
(iv) $P_1(\mathbb{R})$
7. (i) No
(ii) Yes
(iii) $x^2 - x$
(iv) $\{x^2, 1\}$
8. (i) $\{(x_1, -x_1, x_1) \mid x_1 \in \mathbb{R}\}$
(ii) $\{(1, -1, 1)\}$
(iii) \mathbb{R}^2
(iv) $\{(1, 0), (0, 1)\}$
9. (i) No
(ii) Yes
(iii) Constant functions
(iv) $V(\mathbb{R})$.

Chapter 16

Change of Basis

16.1 Coordinate Mapping

Let V be an n dimensional vector space and $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ be an ordered basis. Then, every $v \in V$ can be uniquely expressed as a linear combination of elements of \mathcal{B} , so that there exists unique elements $\alpha_1, \alpha_2, \dots, \alpha_n \in F$ such that $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$.

Unless otherwise stated, V will denote a vector space over a field F .

This defines a mapping $T : V \rightarrow \mathbb{R}^n$ which maps $v \rightarrow (\alpha_1, \alpha_2, \dots, \alpha_n)^t$. It is easy to verify that T is a one-to-one linear transformation of V onto \mathbb{R}^n . This mapping is called the coordinate mapping and we denote it by $[\]_{\mathcal{B}}$, we write $[v]_{\mathcal{B}} = (\alpha_1, \alpha_2, \dots, \alpha_n)^t$.

Example 16.1. Consider $V = \mathbb{P}_3$, over \mathbb{R} . Then $\dim V = 4$, $\mathcal{B} = \{1, x, x^2, x^3\}$ is an ordered basis for V . Let $p = 2 - 3x^2 + 4x^3$. Then $[p]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 0 \\ -3 \\ 4 \end{pmatrix}$. If

$q = a_0 + a_1x + a_2x^2 + a_3x^3$, then $[q]_{\mathcal{B}} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$. $\mathcal{B}' = \{x^3, x^2, x, 1\}$ is another

ordered basis for V , and $[p]_{\mathcal{B}'} = \begin{pmatrix} 4 \\ -3 \\ 0 \\ 2 \end{pmatrix}$ also $[q]_{\mathcal{B}'} = \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix}$.

This illustration shows that the mapping heavily depends upon the ordered basis taken.

Example 16.2. Let $V = M_{2 \times 2}(C)$ and $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, where $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. If $A = \begin{pmatrix} -2 & 3 \\ -4 & 16 \end{pmatrix}$, then $A = -2E_{11} + 3E_{12} - 4E_{21} + 16E_{22}$. If $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$, then $B = b_1E_{11} + b_2E_{12} + b_3E_{21} + b_4E_{22}$.

Example 16.3. Let $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ be a basis for \mathbb{R}^3 , where $v_1 = \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}$,

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}. \text{ If } v = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}, \text{ compute } [v]_{\mathcal{B}_1}.$$

To find $[v]_{\mathcal{B}_1}$, we need to compute the scalars $\alpha_1, \alpha_2, \alpha_3$ such that $v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$. Thus we need to solve the system whose augmented matrix is

$$\left(\begin{array}{ccc|c} v_1 & v_2 & v_3 & v \end{array} \right). \text{ i.e. } \begin{pmatrix} 3 & 0 & 0 & 2 \\ -1 & 1 & 0 & -1 \\ -2 & 3 & 2 & 3 \end{pmatrix}$$

Transforming it to the reduced echelon form, we get $\begin{pmatrix} 1 & 0 & 0 & 2/3 \\ 0 & 1 & 0 & -1/3 \\ 0 & 0 & 1 & 8/3 \end{pmatrix}$

so that $\alpha_1 = 2/3, \alpha_2 = -1/3, \alpha_3 = 8/3$ hence $[v]_{\mathcal{B}_1} = 1/3 \begin{pmatrix} 2 \\ -1 \\ 8 \end{pmatrix}$.

16.2 Change of Basis

Let \mathcal{B} and \mathcal{B}' be two different ordered basis of V , then $[v]_{\mathcal{B}}, [v]_{\mathcal{B}'}$ are the coordinate vectors of $v \in V$ relative to $\mathcal{B}, \mathcal{B}'$ respectively. The question is: Is there a relationship between $[v]_{\mathcal{B}}$ and $[v]_{\mathcal{B}'}$? That is to say that if $[v]_{\mathcal{B}}$ is known, can we find $[v]_{\mathcal{B}'}$ and vice versa?

Theorem 16.1. Let V be an n dimensional vector space with two ordered basis \mathcal{B}_1 and \mathcal{B}_2 , then there exist a unique matrix P such that for any $v \in V$ $[v]_{\mathcal{B}_2} = P[v]_{\mathcal{B}_1}$.

Proof: Let $\mathcal{B}_1 = \{v_1, v_2, v_3, \dots, v_n\}$ be an ordered basis of V . Let $v \in V$ and if $[v]_{\mathcal{B}_1} = (\alpha_1, \alpha_2, \dots, \alpha_n)^t$ then

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$$

$$\therefore [v]_{\mathcal{B}_2} = [\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n]_{\mathcal{B}_2}$$

$$= \alpha_1 [v_1]_{\mathcal{B}_2} + \alpha_2 [v_2]_{\mathcal{B}_2} + \dots + \alpha_n [v_n]_{\mathcal{B}_2}$$

$$\left(\begin{array}{cccc} [v_1]_{\mathcal{B}_2} & [v_2]_{\mathcal{B}_2} & \dots & [v_n]_{\mathcal{B}_2} \end{array} \right) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

$$= P[v]_{\mathcal{B}_1}, \text{ where } j\text{th column of } P \text{ is } [v_j]_{\mathcal{B}_2}$$

Hence $[v]_{\mathcal{B}_2} = P[v]_{\mathcal{B}_1}$

The coordinate vector with respect to basis \mathcal{B}_2 is obtained by multiplying P by the coordinate vector with respect to the basis \mathcal{B}_1 . P is called the transition matrix from \mathcal{B}_1 to \mathcal{B}_2 .

Uniqueness

Let Q be any other matrix such $[v]_{\mathcal{B}_2} = Q[v]_{\mathcal{B}_1}$.

Take $v = v_1$ then $[v_1]_{\mathcal{B}_1} = (1 \ 0 \ 0 \dots 0)^t$.

$$\begin{aligned} \therefore [v_1]_{\mathcal{B}_2} &= Q[v_1]_{\mathcal{B}_1} \\ &= Q \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \end{aligned}$$

= First column of Q .

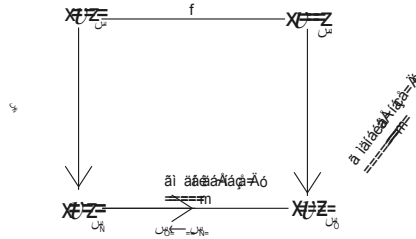
similarly we get, $[v_i]_{\mathcal{B}_2} = i$ th column of Q , $i = 1, 2, \dots, n$.

$$\therefore Q = \begin{pmatrix} [v_1]_{\mathcal{B}_2} & [v_2]_{\mathcal{B}_2} & \dots & [v_n]_{\mathcal{B}_2} \end{pmatrix} = P.$$

Hence P is unique. Thus, we have the following definition. □

Definition 16.1. (Transition Matrix): Let V be an n dimensional vector space and $\mathcal{B}_1, \mathcal{B}_2$ two ordered basis for V . Then, there exists a unique $n \times n$ matrix P whose j th column is the coordinate vector of the j th element of \mathcal{B}_1 relative to \mathcal{B}_2 , such that, for every $v \in V$ $[v]_{\mathcal{B}_2} = P[v]_{\mathcal{B}_1}$, P is called the transition matrix from \mathcal{B}_1 -basis to \mathcal{B}_2 - basis. It is also written as $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$.

The relation is also shown in the following diagram



By definition, the columns of the transition matrix P are the coordinate vectors of a basis and are therefore linearly independent. Thus P is non-singular.

Example 16.4. Let $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ and $\mathcal{B}_2 = \{w_1, w_2, w_3\}$ be two basis for \mathbb{R}^3 , where $v_1 = (0, 3, -1)^t$, $v_2 = (-2, 0, 1)^t$, $v_3 = (-1, 0, 1)^t$, $w_1 = (3, 1, 2)^t$, $w_2 = (5, 0, 1)^t$, $w_3 = (4, 1, 1)^t$. Let $v = (-3, 3, 1)^t$.

- (i) Compute the transition matrix $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$.
- (ii) Verify that $[v]_{\mathcal{B}_2} = P[v]_{\mathcal{B}_1}$.

Solution: (i) To compute $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$, find $\alpha_1, \alpha_2, \alpha_3$ such that

$$\alpha_1 w_1 + \alpha_2 w_2 + \alpha_3 w_3 = v_1$$

Thus we need to solve the linear system with augmented matrix

$$\left(\begin{array}{ccc|c} w_1 & w_2 & w_3 & v_1 \end{array} \right).$$

Similarly we need to find $\beta_1, \beta_2, \beta_3$ and $\gamma_1, \gamma_2, \gamma_3$ such that

$$\beta_1 w_1 + \beta_2 w_2 + \beta_3 w_3 = v_2$$

$$\gamma_1 w_1 + \gamma_2 w_2 + \gamma_3 w_3 = v_3.$$

This gives two linear system whose augmented matrices are

$$\left(\begin{array}{ccc|c} w_1 & w_2 & w_3 & v_2 \end{array} \right) \text{ and } \left(\begin{array}{ccc|c} w_1 & w_2 & w_3 & v_3 \end{array} \right).$$

Since the coefficient matrix of all the 3 system is the same, namely

$\left(\begin{array}{ccc} w_1 & w_2 & w_3 \end{array} \right)$, we can transform the three augmented matrices to reduced echelon form simultaneously by transforming the matrix

$\left(\begin{array}{ccc|ccc} w_1 & w_2 & w_3 & v_1 & v_2 & v_3 \end{array} \right)$ to reduced echelon matrix. Thus we transform the matrix

$$A = \left(\begin{array}{ccc|ccc} 3 & 5 & 4 & : & 0 & -2 & -1 \\ 1 & 0 & 1 & : & 3 & 0 & 0 \\ 2 & 1 & 1 & : & -1 & 1 & 1 \end{array} \right) \quad \dots(1)$$

into reduced echelon form. The transformed matrix is

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & : & -8/6 & 7/6 & 1 \\ 0 & 1 & 0 & : & -16/6 & -1/6 & 0 \\ 0 & 0 & 1 & : & 26/6 & -7/6 & -1 \end{array} \right) \quad \dots(2)$$

Hence the transformed matrix from \mathcal{B}_1 -basis to \mathcal{B}_2 -basis is

$$\begin{aligned} P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} &= \begin{pmatrix} -8/6 & 7/6 & 1 \\ -16/6 & -1/6 & 0 \\ 26/6 & -7/6 & -1 \end{pmatrix} \\ &= 1/6 \begin{pmatrix} -8 & 7 & 6 \\ -16 & -1 & 0 \\ 26 & -7 & -6 \end{pmatrix} \end{aligned}$$

(ii) To find $[v]_{\mathcal{B}_1}$, we write

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$$

So we solve the linear system with the augmented matrix

$$\left(\begin{array}{ccc|c} v_1 & v_2 & v_3 & : & v \end{array} \right) \quad \dots(3)$$

$$\text{i.e.} \left(\begin{array}{ccc|ccc} 0 & -2 & -1 & : & -3 \\ 3 & 0 & 0 & : & 3 \\ -1 & 1 & 1 & : & 1 \end{array} \right)$$

Reducing it to the reduced echelon form, we get

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & : & 1 \\ 0 & 1 & 0 & : & 1 \\ 0 & 0 & 1 & : & 1 \end{array} \right) \quad \dots(4)$$

$$\therefore v = 1v_1 + 1v_2 + 1v_3$$

$$\text{Hence } [v]_{\mathcal{B}_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

We know that

$$\begin{aligned} [v]_{\mathcal{B}_2} &= P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} [v]_{\mathcal{B}_1} \\ &= 1/6 \begin{pmatrix} -8 & 7 & 6 \\ -16 & -1 & 0 \\ 26 & -7 & -6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= 1/6 \begin{pmatrix} 5 \\ -17 \\ 13 \end{pmatrix} \\ &= \begin{pmatrix} 5/6 \\ -17/6 \\ 13/6 \end{pmatrix} \end{aligned}$$

Let us now find $[v]_{\mathcal{B}_2}$ directly.

$$\text{If } v = \alpha_1 w_1 + \alpha_2 w_2 + \alpha_3 w_3$$

then the associated augmented system is

$$\left(\begin{array}{ccc|c} w_1 & w_2 & w_3 & : & v \end{array} \right)$$

On reducing to the reduced row echelon form we get

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & : & 5/6 \\ 0 & 1 & 0 & : & -17/6 \\ 0 & 0 & 1 & : & 13/6 \end{array} \right)$$

$$[v]_{\mathcal{B}_2} = \begin{pmatrix} 5/6 \\ -17/6 \\ 13/6 \end{pmatrix}$$

which is the same as obtained by the formula.

Hence verified. Observe that in the above problem equation (1) is

$$\left(\begin{array}{ccc|ccc} w_1 & w_2 & w_3 & : & v_1 & v_2 & v_3 \end{array} \right)$$

This matrix is transformed to reduced new echelon form to get

$$\left(\begin{array}{ccc|ccc} I & : & P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} & & & & \end{array} \right) \text{ which is equation (2).}$$

To get $[v]_{\mathcal{B}_1}$, we write equation (3) as

$$\left(\begin{array}{ccc|c} B_1 & : & & v \end{array} \right)$$

This is transformed to

$$\left(\begin{array}{ccc|c} I & : & & [v]_{\mathcal{B}_1} \end{array} \right)$$

by applying row operations. $[v]_{\mathcal{B}_2}$ is obtained by the formula

$$[v]_{\mathcal{B}_2} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} [v]_{\mathcal{B}_1}.$$

Corollary 16.2. *In particular, if \mathcal{B}_2 is the standard basis \mathcal{B} and $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ is any basis, then for any $v \in V$,*

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}_1}$$

where the j th column of P is $[v_j]_{\mathcal{B}}$

i.e. P is the matrix whose columns are the coordinate vectors of elements of \mathcal{B}_1 relative to the standard basis. We write P as $P_{\mathcal{B}_1}$.

Corollary 16.3. *If $P_{\mathcal{B}}$ is the transition matrix from any \mathcal{B} -basis to the standard basis. Then the transition matrix from standard basis to the \mathcal{B} -basis is $P_{\mathcal{B}}^{-1}$.*

Proof: $P_{\mathcal{B}}$ is the transition matrix from basis \mathcal{B} to the standard basis.

\therefore for any $v \in V$

$$[v] = P_{\mathcal{B}} [v]_{\mathcal{B}}$$

$$\Rightarrow P_{\mathcal{B}}^{-1} [v] = [v]_{\mathcal{B}}$$

$$\Rightarrow [v]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} [v] \quad \square$$

Corollary 16.4. *If $\mathcal{B}_1, \mathcal{B}_2$ be any two bases of a vector space V and $P_{\mathcal{B}_1}, P_{\mathcal{B}_2}$ be the transition matrices from bases $\mathcal{B}_1, \mathcal{B}_2$ respectively to the standard basis of V . Then the transition matrix from basis \mathcal{B}_1 to \mathcal{B}_2 is $P_{\mathcal{B}_2}^{-1} P_{\mathcal{B}_1}$.*

Proof: Let $P_{\mathcal{B}_1}, P_{\mathcal{B}_2}$ be the transition matrices from bases $\mathcal{B}_1, \mathcal{B}_2$ to the standard basis respectively. Then for any vector $v \in V$

$$[v] = P_{\mathcal{B}_1} [v]_{\mathcal{B}_1}$$

$$\text{and } [v] = P_{\mathcal{B}_2} [v]_{\mathcal{B}_2}.$$

This gives

$$P_{\mathcal{B}_1} [v]_{\mathcal{B}_1} = P_{\mathcal{B}_2} [v]_{\mathcal{B}_2}$$

By uniqueness of the transition matrix from \mathcal{B}_1 to \mathcal{B}_2 we get

$$P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = P_{\mathcal{B}_2}^{-1} P_{\mathcal{B}_1}. \quad \square$$

Example 16.5. *Consider the same bases $\mathcal{B}_1, \mathcal{B}_2$ for \mathbb{R}^3 as in the previous illustration. Find $P_{\mathcal{B}_1}, P_{\mathcal{B}_2}$ also verify that $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = P_{\mathcal{B}_2}^{-1} P_{\mathcal{B}_1}$.*

Solution: Let \mathcal{B} be the standard basis for \mathbb{R}^3 . Then $P_{\mathcal{B}_1} = P_{\mathcal{B} \leftarrow \mathcal{B}_1}$

$$= \begin{pmatrix} 0 & -2 & -1 \\ 3 & 0 & 0 \\ -1 & 1 & 1 \end{pmatrix} \because \text{the columns of } P \text{ are the coordinate vectors of elements}$$

of \mathcal{B}_2 relative to \mathcal{B} . Similarly

$$P_{\mathcal{B}_2} = \begin{pmatrix} 3 & 5 & 4 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

By Corollary 16.4., $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = P_{\mathcal{B}_2}^{-1} P_{\mathcal{B}_1}$

To find $P_{\mathcal{B}_2}^{-1}$ we proceed as follows:

$$A = (P_{\mathcal{B}_2} \quad : \quad I)$$

By applying E -operations, reduce the above matrix to the echelon form

$$A = \left(\begin{array}{ccc|ccc} 3 & 5 & 4 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/6 & -1/6 & 5/6 \\ 0 & 1 & 0 & 1/6 & -5/6 & 1/6 \\ 0 & 0 & 1 & 1/6 & 7/6 & -5/6 \end{array} \right)$$

$$\therefore P_{\mathcal{B}_2}^{-1} = 1/6 \begin{pmatrix} -1 & -1 & 5 \\ 1 & -5 & 1 \\ 1 & 7 & -5 \end{pmatrix}$$

$$\therefore P_{\mathcal{B}_2}^{-1} P_{\mathcal{B}_1} = 1/6 \begin{pmatrix} -1 & -1 & 5 \\ 1 & -5 & 1 \\ 1 & 7 & -5 \end{pmatrix} \begin{pmatrix} 0 & -2 & -1 \\ 3 & 0 & 0 \\ -1 & 1 & 1 \end{pmatrix}$$

$$= 1/6 \begin{pmatrix} -8 & 7 & 6 \\ -16 & -1 & 0 \\ 26 & -7 & -6 \end{pmatrix} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$$

Hence verified.

16.3 Procedure to Compute Transition Matrix $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ from Basis \mathcal{B}_1 to Basis \mathcal{B}_2

Step 1 Let $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$, $\mathcal{B}_2 = \{w_1, w_2, \dots, w_n\}$ be the two bases and \mathcal{B} the standard basis. Let $B_1 = ([v_1]_{\mathcal{B}} \quad [v_2]_{\mathcal{B}} \quad \dots \quad [v_n]_{\mathcal{B}})$, $B_2 = ([w_1]_{\mathcal{B}} \quad [w_2]_{\mathcal{B}} \quad \dots \quad [w_n]_{\mathcal{B}})$

Step 2 Write $A = (B_2 \quad : \quad B_1)$ (I)

By applying elementary row operations, obtain the row reduced echelon form of A . Thus (I) becomes

$$(I \quad : \quad B)$$

where B is the transition matrix from basis \mathcal{B}_1 to \mathcal{B}_2 .

Example 16.6. Find the transition matrix from the standard basis \mathcal{B} to the given basis $\mathcal{B}_1 = \{t^2 - t + 1, t + 1, t^2 + 1\}$. Hence find $[v]_{\mathcal{B}_1}$, where $v = t + 4$. The standard basis is $\mathcal{B} = \{e_1, e_2, e_3\}$ where $e_1 = 1, e_2 = t, e_3 = t^2$.

Solution: $\mathcal{B}_1 = \{v_1, v_2, v_3\}$, where $v_1 = 1 - t + t^2, v_2 = 1 + t, v_3 = 1 + t^2$
We first find $P_{\mathcal{B}_1 \leftarrow \mathcal{B}}$.

$$\begin{aligned} \text{Step 1 } A &= (\mathcal{B}_1 \quad : \quad \mathcal{B}) = (v_1 \quad v_2 \quad v_3 \quad : \quad e_1 \quad e_2 \quad e_3) = \\ & \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \end{aligned}$$

Step 2 Reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 & 0 & : & 1 & -1 & -1 \\ 0 & 1 & 0 & : & 1 & 0 & -1 \\ 0 & 0 & 1 & : & -1 & 1 & 2 \end{pmatrix}$$

Thus $P_{\mathcal{B}_1 \leftarrow \mathcal{B}} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ -1 & 1 & 2 \end{pmatrix}$.

We know that

$$[v]_{\mathcal{B}_1} = P_{\mathcal{B}_1 \leftarrow \mathcal{B}}[v]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -3 \end{pmatrix}$$

$\therefore [v]_{\mathcal{B}_1} = \begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix}$

We shall now prove that the inverse of the transition matrix $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ is $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$, i.e. the transition matrix from the basis \mathcal{B}_2 to basis \mathcal{B}_1 .

Theorem 16.5. *Let $\mathcal{B}_1, \mathcal{B}_2$ be two ordered basis of n dimensional vector space V . Let $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ be the transition matrix from \mathcal{B}_1 – basis to \mathcal{B}_2 – basis. Then $(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$ is the transition matrix from the basis \mathcal{B}_2 to basis \mathcal{B}_1 .*

Proof: Let $v \in V$, then $[v]_{\mathcal{B}_2} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}[v]_{\mathcal{B}_1}$. since $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ is invertible.

$$\therefore [v]_{\mathcal{B}_1} = (P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}[v]_{\mathcal{B}_2}$$

So it follows that the transition matrix from basis \mathcal{B}_2 to basis \mathcal{B}_1 is $(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$. □

Summarizing, we have that if $\mathcal{B}_1, \mathcal{B}_2$ are two bases of a vector space V with standard basis \mathcal{B} and $v \in V$, then $[v]_{\mathcal{B}_2} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}[v]_{\mathcal{B}_1}$

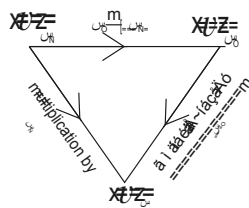
$$[v]_{\mathcal{B}} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}}[v]_{\mathcal{B}_2}$$

$$[v]_{\mathcal{B}} = P_{\mathcal{B}_1 \leftarrow \mathcal{B}}[v]_{\mathcal{B}_1},$$

$$\Rightarrow P_{\mathcal{B}_1 \leftarrow \mathcal{B}_1} P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = (P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$$

$$P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} = (P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1} \quad \dots(\text{by Corollary 16.4})$$

These relationships can be represented diagrammatically by



Example 16.7. *Find the vector v determined by the coordinate vector relative to the basis $\mathcal{B} = \{v_1, v_2, v_3\}$ where $v_1 = (0, 1, 1)^t$, $v_2 = (1, 1, 0)^t$, $v_3 = (1, -1, 0)^t$ and $[v]_{\mathcal{B}} = (2, 1, -3)^t$.*

If $v \in V$, then $[v] = P_{\mathcal{B}}[v]_{\mathcal{B}}$

where the columns of $P_{\mathcal{B}}$ are the coordinate vectors of \mathcal{B} relative to the standard basis, i.e.

$$P_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\therefore [v] = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix}.$$

Example 16.8. Let $\mathcal{B} = \{v_1, v_2, v_3\}$ be a basis of \mathbb{P}_2 , $v_1 = 3 - x - 2x^2$, $v_2 = x + 3x^2$, $v_3 = 2x^2$ if $v = 2 - x + 3x^2$, compute $[v]_{\mathcal{B}}$.

Given vector v and a basis \mathcal{B} , we have the relation $[v] = P_{\mathcal{B}}[v]_{\mathcal{B}}$ where

$$P_{\mathcal{B}} = (v_1 \ v_2 \ v_3) = \begin{pmatrix} 3 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 3 & 2 \end{pmatrix}$$

$\therefore [v]_{\mathcal{B}} = (P_{\mathcal{B}})^{-1}[v]$ to compute $(P_{\mathcal{B}})^{-1}$ we reduce the matrix $(P_{\mathcal{B}} \ : \ [v])$ into reduced echelon form

$$(P_{\mathcal{B}} \ : \ [v]) = \begin{pmatrix} 3 & 0 & 0 & : & 2 \\ -1 & 1 & 0 & : & -1 \\ -2 & 3 & 2 & : & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & : & 2/3 \\ 0 & 1 & 0 & : & -1/3 \\ 0 & 0 & 1 & : & 8/3 \end{pmatrix}$$

$$\therefore [v]_{\mathcal{B}} = (P_{\mathcal{B}})^{-1}[v] = \begin{pmatrix} 2/3 \\ -1/3 \\ 8/3 \end{pmatrix}.$$

Example 16.9. Let $P = \begin{pmatrix} 2 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ be the transition matrix from \mathcal{B}_1 - basis to \mathcal{B}_2 - basis. If $\mathcal{B}_2 = \{w_1, w_2, w_3\}$ where $w_1 = (2, 0, 1)^t$, $w_2 = (1, 2, 0)^t$, $w_3 = (1, 1, 1)^t$ find the basis \mathcal{B}_1 .

Solution: Let $\mathcal{B}_1 = \{u_1, u_2, u_3\}$ be the required ordered basis. The columns of the transition matrix are

$$[u_1]_{\mathcal{B}_2}, [u_2]_{\mathcal{B}_2}, [u_3]_{\mathcal{B}_2}$$

$$\therefore 2w_1 + w_2 + w_3 = u_1$$

$$2w_1 - w_2 + w_3 = u_2$$

$$w_1 + 2w_2 + w_3 = u_3$$

$$\text{Substituting for } w_1, w_2, w_3, \text{ we get } u_1 = \begin{pmatrix} 6 \\ 3 \\ 3 \end{pmatrix}, u_2 = \begin{pmatrix} 4 \\ -1 \\ 3 \end{pmatrix},$$

$$u_3 = \begin{pmatrix} 5 \\ 5 \\ 2 \end{pmatrix}.$$

Aliter: We know that

$$P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = (P_{\mathcal{B}_2})^{-1}P_{\mathcal{B}_1}$$

$$\text{Here } P_{\mathcal{B}_2} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \text{ and } P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

$$\therefore P_{\mathcal{B}_1} = P_{\mathcal{B}_2}P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 4 & 5 \\ 3 & -1 & 5 \\ 3 & 3 & 2 \end{pmatrix}.$$

$$\text{Hence } \mathcal{B}_1 = \left\{ \begin{pmatrix} 6 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 5 \\ 2 \end{pmatrix} \right\}.$$

Example 16.10. Let $P = \begin{pmatrix} 2 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ be the transition matrix from \mathcal{B}_1 -basis to \mathcal{B}_2 -basis. If $\mathcal{B}_1 = \{u_1, u_2, u_3\}$ where $u_1 = (6, 3, 3)^t$, $u_2 = (4, -1, 3)^t$, $u_3 = (5, 5, 2)^t$, find the basis \mathcal{B}_2 .

Solution: We know that the transition matrix from \mathcal{B}_2 -basis to \mathcal{B}_1 -basis

$$=(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1} = P^{-1} = \begin{pmatrix} 3/2 & 1/2 & -5/2 \\ -1/2 & -1/2 & 3/2 \\ -1 & 0 & 2 \end{pmatrix} = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} \text{ (say)}$$

If $\mathcal{B}_2 = \{w_1, w_2, w_3\}$, then the columns of $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$ are $[w_1]_{\mathcal{B}_1}$, $[w_2]_{\mathcal{B}_1}$, $[w_3]_{\mathcal{B}_1}$.

Thus $(3/2)u_1 - (1/2)u_2 - u_3 = w_1$

$(1/2)u_1 - (1/2)u_2 + (0)u_3 = w_2$

$(-5/2)u_1 + (3/2)u_2 + 2u_3 = w_3$

Substituting for u_1, u_2, u_3 we get

$$w_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Aliter: We know that

$$P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = (P_{\mathcal{B}_2})^{-1} P_{\mathcal{B}_1}$$

We are given $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ and $P_{\mathcal{B}_1}$. So we write

$$P_{\mathcal{B}_2} = P_{\mathcal{B}_1} (P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$$

We first find $(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$. We reduce the augmented matrix $(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} : I)$ into the reduced echelon form. The last three columns of the reduced echelon form of $(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} : I)$ form $(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$.

$$(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} : I) = \begin{pmatrix} 2 & 2 & 1 & : & 1 & 0 & 0 \\ 1 & -1 & 2 & : & 0 & 1 & 0 \\ 1 & 1 & 1 & : & 0 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & : & 3/2 & 1/2 & -5/2 \\ 0 & 1 & 0 & : & -1/2 & -1/2 & 3/2 \\ 0 & 0 & 1 & : & -1 & 0 & 2 \end{pmatrix}$$

$$\text{Thus } (P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1} = \begin{pmatrix} 3/2 & 1/2 & -5/2 \\ -1/2 & -1/2 & 3/2 \\ -1 & 0 & 2 \end{pmatrix}$$

$$\text{So } P_{\mathcal{B}_2} = \begin{pmatrix} 6 & 4 & 5 \\ 3 & -1 & 5 \\ 3 & 3 & 2 \end{pmatrix} \begin{pmatrix} 3/2 & 1/2 & -5/2 \\ -1/2 & -1/2 & 3/2 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

$$\text{Hence } \mathcal{B}_2 = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Example 16.11. Let V be an n -dimensional vector space and $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ be 3 ordered basis for V . Give a relation between the transition matrices $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}, P_{\mathcal{B}_3 \leftarrow \mathcal{B}_2}, P_{\mathcal{B}_3 \leftarrow \mathcal{B}_1}$.

Let $A = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}, B = P_{\mathcal{B}_3 \leftarrow \mathcal{B}_2}, C = P_{\mathcal{B}_3 \leftarrow \mathcal{B}_1}$ and let $v \in V$. Then

$$[v]_{\mathcal{B}_2} = A[v]_{\mathcal{B}_1} \quad \dots (1)$$

$$[v]_{\mathcal{B}_3} = B[v]_{\mathcal{B}_2} \quad \dots (2)$$

$$[v]_{\mathcal{B}_3} = C[v]_{\mathcal{B}_1} \quad \dots (3)$$

(1) and (2) imply that

$$[v]_{\mathcal{B}_3} = BA[v]_{\mathcal{B}_1} \quad \dots (4)$$

Comparing (3) and (4) we get $BA = C$

Example 16.12. Let V be an n -dimensional vector space and $\mathcal{B} = \{v_1, v_2, \dots,$

$v_n\}$ be any basis of V and $[v] = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ be the coordinates of v relative to the

standard basis. Then there is a unique matrix $P_{\mathcal{B}}$ such that $[v] = P_{\mathcal{B}}[v]_{\mathcal{B}}$

$$\text{Let } [v]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

Then $v = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n$

$$\Rightarrow [v] = [\beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n] = \beta_1 [v_1] + \beta_2 [v_2] + \dots + \beta_n [v_n]$$

$$= \begin{pmatrix} [v_1] & [v_2] & \dots & [v_n] \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

$$= P_{\mathcal{B}}[v]_{\mathcal{B}}$$

where $P_{\mathcal{B}}$ is the matrix whose j th column is the coordinate vector of v_j relative to the standard basis.

Uniqueness

Let Q be any matrix such that $[v] = Q[v]_{\mathcal{B}}$.

Taking $v = v_j$

$$[v_j] = Q[v_j]_{\mathcal{B}} = Q \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \text{ (} j\text{th row)} \\ \vdots \\ 0 \\ 0 \end{pmatrix} = j\text{th column of } Q$$

since $[v_j]$ is the j th column of $P_{\mathcal{B}}$, \therefore j th column of Q is the same as the j th column of $P_{\mathcal{B}}$ for all $j = 1, 2, \dots, n$.

Hence $Q = P_{\mathcal{B}}$.

16.4 Exercise

1. Find the vector X determined by the coordinate vector relative to the given basis $\mathcal{B} = \{b_1, b_2\}$

- (i) $b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $[X]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$
- (ii) $b_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $[X]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$
- (iii) \mathcal{B} is standard basis and $[X]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$
- (iv) $b_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $b_2 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$, $[X]_{\mathcal{B}} = \begin{pmatrix} -4 \\ 7 \end{pmatrix}$
- (v) $b_1 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$, $b_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $[x]_{\mathcal{B}} = \begin{pmatrix} -4 \\ 7 \end{pmatrix}$
2. Find the coordinates of $p = x^3 - 3x^2 + 1$ relative to the ordered basis
- (i) $\mathcal{B}_1 = \{1, x, x^2, x^3\}$
- (ii) $\mathcal{B}_2 = \{x^3, x^2, x, 1\}$
- (iii) $\mathcal{B}_3 = \{x^2, x^3, 1, x\}$
3. Let $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ Find the coordinates of $A_1 = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 2 & -3 \\ -6 & 5 \end{pmatrix}$ relative to the ordered basis
- (i) $\mathcal{B}_1 = \{E_{11}, E_{12}, E_{21}, E_{22}\}$
- (ii) $\mathcal{B}_2 = \{E_{11}, E_{21}, E_{12}, E_{22}\}$
4. Let $\mathcal{B} = \{v_1, v_2, v_3, v_4\}$ is a basis of \mathbb{P}_3 , where $v_1 = 1 + x$, $v_2 = 2 + x^2$, $v_3 = x + 2x^2 - x^3$, $v_4 = x - x^2$. If $w = 1 + 2x - 6x^2 + 2x^3$, then
- (i) find the matrix $P_{\mathcal{B}}$ such that $[v] = P_{\mathcal{B}}[v]_{\mathcal{B}}$, for all $v \in V$.
- (ii) find $[w]_{\mathcal{B}}$ directly.
- (iii) verify the relation (i) for w .
5. Let $\mathcal{B} = \{v_1, v_2\}$ is an ordered basis of \mathbb{R}^2 , where $v_1 = (1, 2)^t$, $v_2 = (-1, 2)^t$.
- (i) Find $v \in \mathbb{R}^2$ such that $[v]_{\mathcal{B}} = (1, -1)^t$.
- (ii) If $v = (1, -1)^t$, find $[v]_{\mathcal{B}}$.
- (iii) Find transition matrix from \mathcal{B} to the standard basis.
- (iv) Find the transition matrix from the standard basis to \mathcal{B} .
6. Let V be the vector space of all polynomials of degree ≤ 2 . Let $p_1 = 1$, $p_2 = 1 + x$, $p_3 = (1 + x)^2$ and $\mathcal{B} = \{p_1, p_2, p_3\}$, obtain the coordinates of $p = 2 + 3x - x^2$ relative to \mathcal{B} .
7. Let $\mathcal{B} = \{(1, 1)^t, (0, 1)^t\}$ be a basis for \mathbb{R}^2 .
- (i) Find v when $[v]_{\mathcal{B}} = (1, -2)^t$.
- (ii) If $w = (1, -2)^t$, find $[w]_{\mathcal{B}}$.

8. If $P_{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ -3 & -5 \end{pmatrix}$ is the change of coordinates matrix from \mathcal{B} to the standard basis, find the basis \mathcal{B} in \mathbb{R}^2 .
9. If $P = \begin{pmatrix} 1 & 2 \\ -3 & -5 \end{pmatrix}$ is the change of coordinate matrix from the standard basis to some basis \mathcal{B} , find \mathcal{B} in \mathbb{R}^2 .
10. If $\mathcal{B}_1, \mathcal{B}_2$ are two bases for \mathbb{R}^2 , where $\mathcal{B}_1 = \{(1, 2)^t, (2, 5)^t\}$, $\mathcal{B}_2 = \{(4, 5)^t, (3, 4)^t\}$. If $v = (1, 1)^t$, then
- find $[v]_{\mathcal{B}_1}$.
 - find $[v]_{\mathcal{B}_2}$.
 - change of coordinate matrix from \mathcal{B}_1 to \mathcal{B}_2 .
 - change of coordinate matrix from \mathcal{B}_2 to \mathcal{B}_1 .
 - verify (b) using (a) and (c).
 - verify (a) using (b) and (d).
11. Let \mathcal{B} be the standard basis for \mathbb{R}^3 . If $\mathcal{B}_1 = \{v_1, v_2, v_3\}$, where $v_1 = (1, 1, 0)^t$, $v_2 = (1, -1, 0)^t$, $v_3 = (0, 1, 1)^t$, is another basis for \mathbb{R}^3 , then
- write the transition matrix $P_{\mathcal{B} \leftarrow \mathcal{B}_1}$.
 - write the transition matrix $P_{\mathcal{B}_1 \leftarrow \mathcal{B}}$.
 - verify that product of the matrices in (a) and (b) is the identity matrix I .
12. Let $\mathcal{B}_1 = \{v_1, v_2\}$ and $\mathcal{B}_2 = \{w_1, w_2\}$ be two bases for \mathbb{R}^2 , where $v_1 = (1, 0)^t$, $v_2 = (1, -3)^t$, $w_1 = (1, -1)^t$, $w_2 = (1, 1)^t$.
- Compute $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$.
 - Verify $[v]_{\mathcal{B}_1} = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} [v]_{\mathcal{B}_2}$ for $v = (5, 1)^t$.
 - Compute $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ directly and verify that $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}^{-1}$.
13. Let $V = \mathbb{P}_2$, consider three bases for V , namely $\mathcal{B}_1 = \{1 + x + x^2, 1 + x, 1\}$, $\mathcal{B}_2 = \{1 + x^2, 1, x\}$, $\mathcal{B}_3 = \{1, 1 + x^2, x\}$, for $p = (1 + x)^2$. Find
- $[p]_{\mathcal{B}_1}$.
 - $[p]_{\mathcal{B}_2}$.
 - $[p]_{\mathcal{B}_3}$.
 - The transition matrix of \mathcal{B}_3 relative to \mathcal{B}_2 .
 - The transition matrix of \mathcal{B}_1 relative to \mathcal{B}_2 .
 - The transition matrix of \mathcal{B}_3 relative to \mathcal{B}_1 .
14. Let $\mathcal{B}_1 = \{v_1, v_2\}$, $\mathcal{B}_2 = \{w_1, w_2\}$ be two bases for \mathbb{R}^2 , where $w_1 = (1, 0)^t$, $w_2 = (1, -1)^t$. If the transition matrix from \mathcal{B}_1 to \mathcal{B}_2 is $\begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix}$, find \mathcal{B}_1 .

15. Let $\mathcal{B}_1 = \{v_1, v_2\}$, $\mathcal{B}_2 = \{w_1, w_2\}$ be two bases for \mathbb{P}_1 .
- If $v_1 = 1 + 2t$, $v_2 = t$ and the transition matrix from \mathcal{B}_1 to \mathcal{B}_2 is $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Determine \mathcal{B}_2 .
 - If $w_1 = t - 1$, $w_2 = t + 1$ and the transition matrix from \mathcal{B}_2 to \mathcal{B}_1 is $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. Determine \mathcal{B}_1 .
16. If W is the subspace of $\mathbb{C}^3(\mathbb{C})$ spanned by $\mathcal{B} = \{v_1, v_2\}$, where $v_1 = (1, 0, i)^t$, $v_2 = (1 + i, 1, -1)^t$, find $[v]_{\mathcal{B}}$ for
- $v = (1, 1, 0)^t$.
 - $v = (1, i, 1 + i)^t$.
17. Find $[v]_{\mathcal{B}}$, for $v = (10, 5, 10)^t$, $\mathcal{B} = \{(1, 0, -1)^t, (1, 2, 1)^t, (0, -3, 2)^t\}$ by finding the transition matrix from the standard basis to \mathcal{B} .
18. Let $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$
and $\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\}$
be the ordered bases for $\mathbb{M}_{2 \times 2}$, let $M = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$.
Find
- the coordinate vector of M with respect to basis to \mathcal{B}_1 .
 - the coordinate vector of M with respect to basis to \mathcal{B}_2 , using the transition matrix.
19. Let $\mathcal{B}_1 = \{v_1, v_2, v_3, v_4\}$ and $\mathcal{B}_2 = \{w_1, w_2, w_3, w_4\}$ be two bases for \mathbb{R}^4 , where $v_1 = (0, 0, 1, -1)^t$, $v_2 = (0, 0, 1, 1)^t$, $v_3 = (0, 1, 1, 0)^t$, $v_4 = (1, 0, -1, 0)^t$, $w_1 = (0, 1, 0, 0)^t$, $w_2 = (0, 0, -1, 1)^t$, $w_3 = (0, -1, 0, 2)^t$, $w_4 = (1, 1, 0, 0)^t$, let \mathcal{B} be the standard basis, then find the following transition matrices
- from \mathcal{B}_2 to \mathcal{B}_1 .
 - from \mathcal{B}_1 to \mathcal{B}_2 .
 - from \mathcal{B}_1 to \mathcal{B} .
 - from \mathcal{B} to \mathcal{B}_2 .
 - verify (b), using (c) and (d). What is the relation between the matrices in (a) and (b)?
20. Let $V = \mathbb{P}_3$ and $\mathcal{B}_1, \mathcal{B}_2$ be two ordered bases for V , where $\mathcal{B}_1 = \{1, x + x^2, 2x + x^3, x^2 + x^3\}$, $\mathcal{B}_2 = \{1 + x, x^2, x^3, 1\}$.
- Find the transition matrix from basis \mathcal{B}_2 to \mathcal{B}_1 .
 - If $v = 1 + x + x^2 + x^3$, find $[v]_{\mathcal{B}_1}$, $[v]_{\mathcal{B}_2}$.

16.5 Matrix of a Linear Transformation

Earlier we showed that every linear transformation from \mathbb{R}^n to \mathbb{R}^m can be represented by an $m \times n$ matrix, i.e. given a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ we can find a $m \times n$ matrix A such that $T(v) = Av$ for all $v \in \mathbb{R}^n$. In that case we had taken both the basis for \mathbb{R}^n and for \mathbb{R}^m to be the standard bases. We now generalize this result for a linear transformation from an arbitrary finite dimensional vector space to another finite dimensional vector space. Moreover the bases considered need not necessarily be the standard bases. Let us first consider an example.

Example 16.13. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be defined by $T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \\ y \end{pmatrix}$.

Let $\mathcal{B} = \{e_1, e_2\}$ be the standard basis for \mathbb{R}^2 and $\mathcal{B}_1 = \{w_1, w_2, w_3\}$ where $w_1 = (1, 0, 1)^t$, $w_2 = (0, 1, 1)^t$, $w_3 = (1, 1, 1)^t$ is a basis for \mathbb{R}^3

Let $v \in \mathbb{R}^2$, $v = \begin{pmatrix} x \\ y \end{pmatrix}$, then $T(v) = \begin{pmatrix} x+y \\ x-y \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

Find the coordinates of $T(v)$ relative to the basis \mathcal{B}_1 . Since the coordinate mapping is a linear transformation, therefore to find $[T(v)]_{\mathcal{B}_1}$, it is sufficient to find the coordinates of $(1, 1, 0)^t$ and $(1, -1, 0)^t$ relative to the basis \mathcal{B}_1 . To do this, we proceed as follows:

$$A = \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 0 & 0 \end{array} \right)$$

Transform A to reduced row echelon form, then

$$A \sim \left(\begin{array}{ccc|cc} 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 2 & 0 \end{array} \right)$$

$$\text{Thus } [T(v)]_{\mathcal{B}_1} = \begin{pmatrix} -1 & 1 \\ -1 & -1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = P[v]_{\mathcal{B}} \text{ where } P = \begin{pmatrix} -1 & 1 \\ -1 & -1 \\ 2 & 0 \end{pmatrix}$$

This shows that there exists a 3×2 matrix P , such that $[T(v)]_{\mathcal{B}_1} = P[v]_{\mathcal{B}}$.

This matrix P depends on the bases \mathcal{B} and \mathcal{B}_1 of \mathbb{R}^2 and \mathbb{R}^3 respectively. As will be seen later, if the bases are changed P will change.

We now state and prove the general result.

Theorem 16.6. Let V and W be two n and m dimensional vector spaces over the field F and $T : V \rightarrow W$ be a linear transformation of V into W . Let $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$ and $\mathcal{B}_2 = \{w_1, w_2, \dots, w_m\}$ be ordered bases for V and W respectively, then there exist a unique $m \times n$ matrix A such that $[T(v)]_{\mathcal{B}_2} = A[v]_{\mathcal{B}_1}$, for all $v \in V$

Proof: Let $[v_i]_{\mathcal{B}_1}$ denote the coordinate vector of v_i , $1 \leq i \leq n$. Let $v \in V$, then there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_n \in F$ such that $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$

then $[v]_{\mathcal{B}_1} = (\alpha_1, \alpha_2, \dots, \alpha_n)^t$

$\therefore T$ is a linear transformation

$\therefore T(v) = T(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \dots + \alpha_n T(v_n)$

for $1 \leq i \leq n$, $T(v_i) \in W$, so that $T(v_i)$ can be expressed as a linear combination of the basis \mathcal{B}_2 . So, there exists $r_{1i}, r_{2i}, \dots, r_{mi} \in F$ such that

$$T(v_i) = r_{1i}w_1 + r_{2i}w_2 + \dots + r_{mi}w_m$$

$$\therefore [T(v_j)]_{\mathcal{B}_2} = \begin{pmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{mj} \end{pmatrix}$$

$$\text{Let } A = (r_{ij})_{m \times n} = ([T(v_1)]_{\mathcal{B}_2} \quad [T(v_2)]_{\mathcal{B}_2} \quad \dots \quad [T(v_n)]_{\mathcal{B}_2})$$

We show that A is the desired matrix,

$$\begin{aligned} A[v]_{\mathcal{B}_1} &= A \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = ([T(v_1)]_{\mathcal{B}_2} \quad [T(v_2)]_{\mathcal{B}_2} \quad \dots \quad [T(v_n)]_{\mathcal{B}_2}) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \\ &= \alpha_1[T(v_1)]_{\mathcal{B}_2} + \alpha_2[T(v_2)]_{\mathcal{B}_2} + \dots + \alpha_n[T(v_n)]_{\mathcal{B}_2} = [\alpha_1T(v_1) + \alpha_2T(v_2) + \dots + \alpha_nT(v_n)]_{\mathcal{B}_2} \\ &\quad \{\text{as coordinate mapping is a linear transformation}\} \\ &= [T(v)]_{\mathcal{B}_2} \end{aligned}$$

$$\therefore [T(v)]_{\mathcal{B}_2} = A[v]_{\mathcal{B}_1}$$

This proves the existence of a matrix A .

Uniqueness:

Let $B = (b_{ij})_{m \times n}$ be a matrix such that $[T(v)]_{\mathcal{B}_2} = B[v]_{\mathcal{B}_1}$, for all $v \in V$ for each $i = 1, 2, \dots, n$.

$$[T(v_i)]_{\mathcal{B}_2} = B[v_i]_{\mathcal{B}_1} = B \begin{pmatrix} 0 \\ \vdots \\ 1 \text{ (ith coordinate)} \\ \vdots \\ 0 \end{pmatrix} = \text{ith column of } B, \text{ but } [T(v_i)]_{\mathcal{B}_2} =$$

ith column of A . Hence for each $i = 1, 2, \dots, n$ the i th column of A and B are identical.

$\therefore A = B$. □

The matrix A is called the matrix of the linear transformation T relative to the basis \mathcal{B}_1 and \mathcal{B}_2 . The j th column of A is the coordinate vector of the $T(v_j)$ relative to the basis \mathcal{B}_2 , where v_j is the j th element of the ordered basis \mathcal{B}_1 . Thus, we have the following definition.

Definition 16.2. Let $V(F)$ and $W(F)$ be two vector spaces with dimensions n and m respectively, and $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$, $\mathcal{B}_2 = \{w_1, w_2, \dots, w_m\}$ be ordered basis for V and W respectively. Let $T : V \rightarrow W$ be a linear transformation.

For $1 \leq j \leq n$, if

$$T(v_j) = \sum_{i=1}^m \alpha_{ij} w_i$$

Then $A = (\alpha_{ij})_{m \times n}$ is called the matrix of T relative to the basis \mathcal{B}_1 and \mathcal{B}_2 . It is denoted by $[T]_{\mathcal{B}_1\mathcal{B}_2}$. Observe that the j th column of A is the j th coordinate vector $[T(v_j)]_{\mathcal{B}_2}$. If both \mathcal{B}_1 and \mathcal{B}_2 are taken as the standard basis then $[T]_{\mathcal{B}_1\mathcal{B}_2}$ is called the standard matrix of T . If T is a linear operation on a n -dimensional vector space V and \mathcal{B} is an ordered basis for V , then $[T]_{\mathcal{B}\mathcal{B}}$ is the matrix of T relative to the basis \mathcal{B} . It is denoted by $[T]_{\mathcal{B}}$.

Example 16.14. Consider the linear transformation in Example 16.13. Let $\mathcal{B}' = \{v_1, v_2\}$ where $v_1 = (1, -1)^t$, $v_2 = (1, 2)^t$ be a basis for \mathbb{R}^2 . Consider some basis \mathcal{B}_1 for \mathbb{R}^3 . We want to find the matrix of T relative to \mathcal{B}' and \mathcal{B}_1 .

$$T(v_1) = (0, 2, 1)^t$$

$$T(v_2) = (3, -1, 2)^t$$

To find $[T(v_1)]_{\mathcal{B}_1}$ and $[T(v_2)]_{\mathcal{B}_1}$, we reduce the matrix

$$P = \begin{pmatrix} w_1 & w_2 & w_3 & T(v_1) & T(v_2) \end{pmatrix}$$

to the reduced echelon form

$$P = \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & 1 & 2 & -1 \\ 1 & 1 & 1 & -1 & 2 \end{array} \right) \sim \left(\begin{array}{ccc|cc} 1 & 0 & 0 & -3 & 3 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 3 & 0 \end{array} \right)$$

Hence

$$[T(v_1)]_{\mathcal{B}_1} = \begin{pmatrix} -3 \\ -1 \\ 0 \end{pmatrix}$$

$$[T(v_2)]_{\mathcal{B}_1} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$$

Matrix of T relative to \mathcal{B}' and $\mathcal{B}_1 = [T]_{\mathcal{B}'\mathcal{B}_1} = \begin{pmatrix} [T(v_1)]_{\mathcal{B}_1} & [T(v_2)]_{\mathcal{B}_1} \end{pmatrix} =$

$$\begin{pmatrix} -3 & 3 \\ -1 & -1 \\ 3 & 0 \end{pmatrix}$$

Comparing with Example 16.13 we see that by changing the basis, the matrix of the linear transformation has also changed.

16.6 Working Rule to Obtain $[T]_{\mathcal{B}_1\mathcal{B}_2}$

Let $V(F)$ and $W(F)$ be two vector spaces over F of dimension n and m respectively. Let $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$ and $\mathcal{B}_2 = \{w_1, w_2, \dots, w_m\}$ be ordered bases for V and W respectively. Let T be a linear transformation of V into W . To find $[T]_{\mathcal{B}_1\mathcal{B}_2}$:

Step 1: For $1 \leq i \leq n$, find $[T(v_i)]$, the coordinate vector of $T(v_i)$, relative to the standard basis.

For $1 \leq i \leq n$, find $[w_i]$.

Step 2: Let $P = \begin{pmatrix} [w_1] & [w_2] & \dots & [w_m] & [T(v_1)] & [T(v_2)] & \dots & [T(v_n)] \end{pmatrix}$

Transform P to reduced echelon form so that

$$P \sim \begin{pmatrix} I & : & B \end{pmatrix}$$

Then $B = [T]_{\mathcal{B}_1\mathcal{B}_2}$.

Example 16.15. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ be defined by $T(a + bx) = a + bx + ax^2$, let $\mathcal{B}_1 = \{1, x\}$ and $\mathcal{B}_2 = \{1 + x, 1 - x, x^2\}$ be ordered basis for \mathbb{P}_1 and \mathbb{P}_2 respectively. Find $[T]_{\mathcal{B}_1\mathcal{B}_2}$.

Let $\mathcal{B}_1 = \{v_1, v_2\}$, where $v_1 = 1, v_2 = x$ and $\mathcal{B}_2 = \{w_1, w_2, w_3\}$, where $w_1 = 1 + x, w_2 = 1 - x, w_3 = x^2$.

We have

$$\begin{aligned} T(v_1) &= 1 + x^2 \\ T(v_2) &= x \end{aligned}$$

To find $[T(v_1)]_{\mathcal{B}_2}, [T(v_2)]_{\mathcal{B}_2}$ we reduce the matrix

$P = ([w_1] \ [w_2] \ [w_3] : [T(v_1)] \ [T(v_2)])$ to row reduced echelon form

$$P = \left(\begin{array}{ccc|cc} 1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 & -1/2 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right)$$

$$\text{Hence } [T]_{\mathcal{B}_1\mathcal{B}_2} = ([T(v_1)]_{\mathcal{B}_2} \ [T(v_2)]_{\mathcal{B}_2}) = \left(\begin{array}{cc} 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1 & 0 \end{array} \right)$$

The following theorem gives a relation between the matrix of a linear transformation in two different bases.

Theorem 16.7. (Change of basis) Let V be an n -dimensional vector space and $\mathcal{B}_1, \mathcal{B}_2$ be two bases for V . If T is a linear transformation on V , then

$$[T]_{\mathcal{B}_2} = P^{-1}[T]_{\mathcal{B}_1}P \text{ where } P = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}.$$

Proof: Let $v \in V$, and $P = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$ then $[T(v)]_{\mathcal{B}_1} = P[t(v)]_{\mathcal{B}_2}$.

Since P is invertible

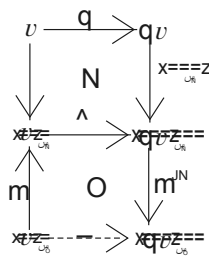
$$\therefore [T(v)]_{\mathcal{B}_2} = (P)^{-1}[T(v)]_{\mathcal{B}_1} = (P)^{-1}[T]_{\mathcal{B}_1}[v]_{\mathcal{B}_1} \quad \dots \text{by definition of } [T]_{\mathcal{B}_1}$$

$$= (P)^{-1}[T]_{\mathcal{B}_1}P[v]_{\mathcal{B}_2} \quad \dots \text{by definition of } P.$$

Also $[Tv]_{\mathcal{B}_2} = [T]_{\mathcal{B}_2}[v]_{\mathcal{B}_2}$ by uniqueness of $[T]_{\mathcal{B}_2}$ we get

$$[T]_{\mathcal{B}_2} = (P)^{-1}[T]_{\mathcal{B}_1}P \quad \square$$

A diagrammatic representation of the above theorem is given here:



Let $[T]_{\mathcal{B}_1} = A, [T]_{\mathcal{B}_2} = B$

Figure(1) commutes, so $A[v]_{\mathcal{B}_1} = [Tv]_{\mathcal{B}_2}$

Figure(2) commutes, so $B[T]_{\mathcal{B}_2} = [Tv]_{\mathcal{B}_2} = P^{-1}AP[v]_{\mathcal{B}_2}$

$$\therefore B = P^{-1}AP$$

Example 16.16. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined by $T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ and

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

be the ordered basis for \mathbb{R}^2 . Let P be the transition matrix from \mathcal{B}_2 to \mathcal{B}_1 . Find $[T]_{\mathcal{B}_1}$, $[T]_{\mathcal{B}_2}$ and verify that $[T]_{\mathcal{B}_2} = P^{-1}[T]_{\mathcal{B}_1}P$

Solution: To find $[T]_{\mathcal{B}_1}$, let $v_1 = (1 \ -1)^t$, $v_2 = (0 \ 1)^t$

$$\text{then } T(v_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, T(v_2) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

reduce the matrix

$$([v_1][v_2]:[T(v_1)]T(v_2)) = \begin{pmatrix} 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 0 \end{pmatrix} = M(\text{say})$$

to reduced echelon form. Then $M \sim \begin{pmatrix} 1 & 0 & : & 1 & -1 \\ 0 & 1 & : & 2 & -1 \end{pmatrix}$

$$\text{so that } [T]_{\mathcal{B}_1} = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$$

$$\text{To find } [T]_{\mathcal{B}_2} \text{ let } w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\text{then } T(w_1) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, T(w_2) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$\text{Reduce the matrix } ([w_1] [w_2] : [T(w_1)] [T(w_2)]) = \begin{pmatrix} 1 & 1 & : & -1 & -2 \\ 1 & 2 & : & 1 & 1 \end{pmatrix} =$$

$$N(\text{say}) \text{ to reduced echelon form so that } [T]_{\mathcal{B}_2} = \begin{pmatrix} -3 & -5 \\ 2 & 3 \end{pmatrix}.$$

To find the transition matrix P from \mathcal{B}_2 -basis to \mathcal{B}_1 -basis, the matrix

$$\begin{pmatrix} 1 & 0 & : & 1 & 1 \\ -1 & 1 & : & 1 & 2 \end{pmatrix} = Q(\text{say}) \text{ is reduced to the reduced echelon form. Then}$$

$$Q \sim \begin{pmatrix} 1 & 0 & : & 1 & 1 \\ 0 & 1 & : & 2 & 3 \end{pmatrix} \text{ so that } P = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}. \text{ Then}$$

$$P^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \text{ and } P^{-1}[T]_{\mathcal{B}_1}P = \begin{pmatrix} -3 & -5 \\ 2 & 3 \end{pmatrix} = [T]_{\mathcal{B}_2}$$

Hence verified.

Example 16.17. Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_4$ be defined by $T(a+bt+ct^2) = t^2(a+bt+ct^2)$. If $\mathcal{B}_1 = \{1, t, t^2\}$, $\mathcal{B}_2 = \{1, t, t^2, t^3, t^4\}$ are bases for \mathbb{P}_2 and \mathbb{P}_4 respectively. Find $[T]_{\mathcal{B}_1\mathcal{B}_2}$.

Solution: First we find the images of the elements of \mathcal{B}_1 .

$$T(1) = t^2, T(t) = t^3, T(t^2) = t^4, \text{ so}$$

$$[T(1)]_{\mathcal{B}_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, [T(t)]_{\mathcal{B}_2} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, [T(t^2)]_{\mathcal{B}_2} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{Hence } [T]_{\mathcal{B}_1\mathcal{B}_2} = ([T(1)]_{\mathcal{B}_2} \ [T(t)]_{\mathcal{B}_2} \ [T(t^2)]_{\mathcal{B}_2}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Example 16.18. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by $T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y \\ x-z \end{pmatrix}$

Let $\mathcal{B}_1 = \{v_1, v_2, v_3\}$, $\mathcal{B}_2 = \{w_1, w_2, w_3\}$ be two bases for \mathbb{R}^3 , where

$v_1 = (1, 1, 1)^t$, $v_2 = (1, 2, 3)^t$, $v_3 = (0, 1, 0)^t$, $w_1 = (1, 2, 2)^t$, $w_2 = (2, 1, 3)^t$, $w_3 = (1, 1, 0)^t$.

Find $[T]_{\mathcal{B}_1}$, $[T]_{\mathcal{B}_2}$ also verify that $[T]_{\mathcal{B}_2} = P^{-1}[T]_{\mathcal{B}_1}P$, where P is the transition matrix from \mathcal{B}_2 -basis to \mathcal{B}_1 -basis.

We have $T(v_1) = (2, 1, 0)^t$, $T(v_2) = (3, 2, -2)^t$, $T(v_3) = (1, 1, 0)^t$ to obtain $[T]_{\mathcal{B}_1}$ we reduce the matrix

$$\left(\begin{array}{ccc|ccc} [v_1] & [v_2] & [v_3] & [T(v_1)] & [T(v_2)] & [T(v_3)] \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & 3 & 1 \\ 1 & 2 & 1 & 1 & 2 & 1 \\ 1 & 3 & 0 & 0 & -2 & 0 \end{array} \right) =$$

L (say) to reduced echelon form. Then

$$L \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 7 & 3/2 \\ 0 & 1 & 0 & -1 & -3 & -1/2 \\ 0 & 0 & 1 & 0 & 1 & 1/2 \end{array} \right)$$

$$\text{Thus } [T]_{\mathcal{B}_1} = \left(\begin{array}{ccc} 3 & 7 & 3/2 \\ -1 & -3 & -1/2 \\ 0 & 1 & 1/2 \end{array} \right)$$

To obtain $[T]_{\mathcal{B}_2}$

$$T(w_1) = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, T(w_2) = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}, T(w_3) = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

To obtain $[T]_{\mathcal{B}_2}$ reduce the matrix

$$\left(\begin{array}{ccc|ccc} [w_1] & [w_2] & [w_3] & [T(w_1)] & [T(w_2)] & [T(w_3)] \end{array} \right)$$

$$\text{i.e. } \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 3 & 3 & 2 \\ 2 & 1 & 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & -1 & -1 & 1 \end{array} \right) = M(\text{say})$$

to reduced echelon form. Then

$$M \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -4/5 & -7/5 & -2/5 \\ 0 & 1 & 0 & 1/5 & 3/5 & 3/5 \\ 0 & 0 & 1 & 17/5 & 16/5 & 6/5 \end{array} \right)$$

$$\therefore [T]_{\mathcal{B}_2} = \left(\begin{array}{ccc} -4/5 & -7/5 & -2/5 \\ 1/5 & 3/5 & 3/5 \\ 17/5 & 16/5 & 6/5 \end{array} \right) = 1/5 \left(\begin{array}{ccc} -4 & -7 & -2 \\ 1 & 3 & 3 \\ 17 & 16 & 6 \end{array} \right)$$

The transition matrix can be calculated. We get

$$P = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} = 1/2 \left(\begin{array}{ccc} 1 & 3 & 3 \\ 1 & 1 & -1 \\ 1 & -3 & 1 \end{array} \right)$$

$$\text{Then } P^{-1} = 1/5 \left(\begin{array}{ccc} 1 & 6 & 3 \\ 1 & 1 & -2 \\ 2 & -3 & 1 \end{array} \right)$$

The following theorem expresses $[T]_{\mathcal{B}_1 \mathcal{B}_2}$ in terms of the standard matrix of T and the transition matrices.

Theorem 16.8. Let $V(F)$ and $W(F)$ be n and m dimensional vector spaces respectively, and $T : V \rightarrow W$ be a linear transformation, and \mathcal{B}_1 , \mathcal{B}_2 are standard bases for V and W respectively. Let A be the matrix of T relative to the standard bases for V and W . Then

$$[T]_{\mathcal{B}_1 \mathcal{B}_2} = P_{\mathcal{B}_2}^{-1} A P_{\mathcal{B}_1}$$

Proof: Let $v \in V$ then

$$[T(v)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_1\mathcal{B}_2}[v]_{\mathcal{B}_1} \dots (1)$$

$$[v] = P_{\mathcal{B}_1}[v]_{\mathcal{B}_1} \dots (2)$$

$$[T(v)] = P_{\mathcal{B}_2}[T(v)]_{\mathcal{B}_2} \dots (3)$$

$$[T(v)] = A[v] \dots (4)$$

$$(3) \Rightarrow [T(v)]_{\mathcal{B}_2} = P_{\mathcal{B}_2}^{-1}[T(v)] = P_{\mathcal{B}_2}^{-1}A[v] \dots (\text{using (4)})$$

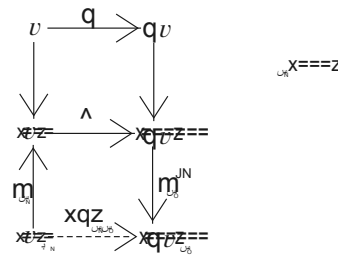
$$= P_{\mathcal{B}_2}^{-1}AP_{\mathcal{B}_1}[v]_{\mathcal{B}_1} \dots (\text{using (2)})$$

$$\therefore [T(v)]_{\mathcal{B}_2} = P_{\mathcal{B}_2}^{-1}AP_{\mathcal{B}_1}[v]_{\mathcal{B}_1} \dots (5)$$

By the uniqueness of $[T]_{\mathcal{B}_1\mathcal{B}_2}$, (1) and (5) gives

$$[T]_{\mathcal{B}_1\mathcal{B}_2} = P_{\mathcal{B}_2}^{-1}AP_{\mathcal{B}_1}.$$

A diagrammatic representation of the above theorem is given below:



□

Example 16.19. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 - 2x_2 \\ x_1 + 2x_2 \end{pmatrix}$$

Let $\mathcal{B}_1 = \{v_1, v_2\}$ and $\mathcal{B}_2 = \{w_1, w_2\}$ be two bases of \mathbb{R}^2 ,

$$\text{where } v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

$$\text{We have } P_{\mathcal{B}_1} = \begin{pmatrix} [v_1] & [v_2] \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$P_{\mathcal{B}_2} = \begin{pmatrix} [w_1] & [w_2] \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & -3 \end{pmatrix}$$

$$A = \begin{pmatrix} T(e_1) & T(e_2) \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}$$

$$P_{\mathcal{B}_2}^{-1}AP_{\mathcal{B}_1} = 1/3 \begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = 1/3 \begin{pmatrix} 8 & 0 \\ 1 & -3 \end{pmatrix}$$

... (1) Let us now calculate $[T]_{\mathcal{B}_1\mathcal{B}_2}$.

$$M = \begin{pmatrix} [w_1] & [w_2] & : & [T(v_1)] & [T(v_2)] \end{pmatrix} = \begin{pmatrix} 1 & 1 & : & 3 & -1 \\ 0 & -3 & : & -1 & 3 \end{pmatrix}$$

Then M is reduced to the row reduced echelon form

$$M \sim \begin{pmatrix} 1 & 0 & : & 8/3 & 0 \\ 0 & 1 & : & 1/3 & -1 \end{pmatrix}$$

$$\text{so that } [T]_{\mathcal{B}_1\mathcal{B}_2} = \begin{pmatrix} 8/3 & 0 \\ 1/3 & -1 \end{pmatrix} = P_{\mathcal{B}_2}^{-1}AP_{\mathcal{B}_1} \dots \text{from (1)}$$

Hence verified.

Example 16.20. Let $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ be the linear transformation whose matrix relative to the basis \mathcal{B} , \mathcal{B}' is

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ -1 & 0 & 3 & 0 \\ 0 & 1 & 1 & 4 \end{pmatrix}$$

where \mathcal{B} is the standard basis and $\mathcal{B}' = \{v_1, v_2, v_3\}$ with $v_1 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$, $v_2 =$

$$\begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$$

Find $T(e_1)$, $T(e_2)$, $T(e_3)$ and $T(e_4)$ relative to the standard basis.

Solution: We have $[T]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} 1 & 2 & 1 & -1 \\ -1 & 0 & 3 & 0 \\ 0 & 1 & 1 & 4 \end{pmatrix}$

The columns of the above matrix are $[T(e_i)]_{\mathcal{B}'}$, where e_i 's are the vectors of the standard basis of \mathbb{R}^4 . Thus

$$[T(e_1)]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, [T(e_2)]_{\mathcal{B}'} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, [T(e_3)]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, [T(e_4)]_{\mathcal{B}'} =$$

$$\begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}$$

$$\therefore T(e_1) = 1v_1 - 1v_2 + 0v_3 = v_1 - v_2 = (0, -1, 0)^t$$

$$\therefore T(e_1) \text{ relative to the standard basis is } \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\text{i.e. } [T(e_1)] = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\text{Similarly } T(e_2) = 2v_1 - v_3 = (-2, -1, -1)^t$$

$$\therefore [T(e_2)] = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}$$

$$T(e_3) = v_1 + 3v_2 - v_3 = (-4, 0, -3)^t$$

$$\therefore [T(e_3)] = \begin{pmatrix} -4 \\ 0 \\ -3 \end{pmatrix}$$

$$T(e_4) = -v_2 + 4v_3 = (1, -3, -3)^t$$

$$\therefore [T(e_4)] = \begin{pmatrix} 1 \\ -3 \\ -3 \end{pmatrix}.$$

Example 16.21. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear transformation determined by the matrix

$$A = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix}$$

Let $\mathcal{B} = \{v_1, v_2, v_3\}$ be a basis for \mathbb{R}^3 where $v_1 = (1, 2, 0)^t$, $v_2 = (0, 1, -1)^t$, $v_3 = (0, 0, 2)^t$. Find the matrix of T relative to the basis \mathcal{B} .

Solution: We know that $[T]_{\mathcal{B}} = ([T(v_1)]_{\mathcal{B}} \ [T(v_2)]_{\mathcal{B}} \ [T(v_3)]_{\mathcal{B}})$
 We reduce the matrix $M = (v_1 \ v_2 \ v_3 : [T(v_1)] \ [T(v_2)] \ [T(v_3)])$
 to the reduced echelon form. The linear transformation determined by A is
 $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that $Tv = Av$

$$\therefore T(v_1) = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 4 \end{pmatrix}$$

$$T(v_2) = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$T(v_3) = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & 0 & 0 & : & 0 & -1 & 0 \\ 2 & 1 & 0 & : & 5 & 1 & 1 \\ 0 & -1 & 1 & : & 4 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & : & 0 & -1 & 0 \\ 0 & 1 & 0 & : & 5 & 3 & 1 \\ 0 & 0 & 1 & : & 9 & 4 & 2 \end{pmatrix}$$

$$\text{Hence } [T]_{\mathcal{B}} = \begin{pmatrix} 0 & -1 & 0 \\ 5 & 3 & 1 \\ 9 & 4 & 2 \end{pmatrix}.$$

Example 16.22. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear transformation defined by
 $T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 + 3x_2 \\ x_1 - 3x_2 \end{pmatrix}$, let \mathcal{B}_1 be the standard basis for \mathbb{R}^2 and $\mathcal{B}_2 =$
 $\{v_1, v_2\}$, where $v_1 = (1, 1)^t$, $v_2 = (1, -1)^t$. Find $[T]_{\mathcal{B}_1\mathcal{B}_2}$, compute
 $T\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$ using the matrix $[T]_{\mathcal{B}_1\mathcal{B}_2}$ and verify the result by direct computation.

Solution: Let $\mathcal{B}_1 = \{e_1, e_2\}$. Then $T(e_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $T(e_2) = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$.

Let $A = ([v_1] \ [v_2] : [T(e_1)] \ [T(e_2)])$

The matrix A is reduced to the reduced echelon form

$$A = \begin{pmatrix} 1 & 1 & : & 1 & 3 \\ 1 & -1 & : & 1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & : & 1 & 0 \\ 0 & 1 & : & 0 & 3 \end{pmatrix}$$

$$\text{Then } [T]_{\mathcal{B}_1\mathcal{B}_2} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

Let $v = (1, -1)^t$. We know that $[T(v)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_1\mathcal{B}_2}[v]_{\mathcal{B}_1} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} =$

$$\begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

$$\therefore T(v) = 1v_1 - 3v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\therefore T(v) = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

By direct computation

$$T(v) = T\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 1 - 3 \\ 1 + 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

Hence verified.

Example 16.23. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_1$ be a linear transformation and $\mathcal{B}_1 =$
 $\{v_1, v_2\}$,
 $\mathcal{B}_2 = \{w_1, w_2\}$ be two bases for \mathbb{P}_1 , where $v_1 = 1 - x$, $v_2 = x$, $w_1 = 1 + 2x$, $w_2 =$

$1-x$ if $[T]_{\mathcal{B}_1} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$, obtain $[T]_{\mathcal{B}_2}$. Also obtain $[T]_{\mathcal{B}_2}$ from the standard matrix of T .

Solution: We know that $[T]_{\mathcal{B}_2} = P^{-1}[T]_{\mathcal{B}_1}P$, where $P = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$

First we find P . To do this transform the matrix $A = \begin{pmatrix} [v_1] & [v_2] \\ [w_1] & [w_2] \end{pmatrix}$ to the reduced echelon form.

$$A = \begin{pmatrix} 1 & 0 & : & 1 & 1 \\ -1 & 1 & : & 2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & : & 1 & 1 \\ 0 & 1 & : & 3 & 0 \end{pmatrix}$$

$$\text{Thus } P = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} = \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix}$$

$$\begin{aligned} \therefore [T]_{\mathcal{B}_2} &= P^{-1}[T]_{\mathcal{B}_1}P = -1/3 \begin{pmatrix} 0 & -1 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix} = \\ &1/3 \begin{pmatrix} 2 & -1 \\ 19 & 4 \end{pmatrix} \end{aligned}$$

Let us obtain the $[T]_{\mathcal{B}_2}$ from the standard matrix of T . Let \mathcal{B} denote the standard basis for \mathbb{P}_1 . To obtain the standard matrix $[T]_{\mathcal{B}}$ of T , we proceed as follows:

$$[T]_{\mathcal{B}} = Q^{-1}[T]_{\mathcal{B}_1}Q, \text{ where } Q = P_{\mathcal{B}_1 \leftarrow \mathcal{B}}$$

$$\text{But } Q = P_{\mathcal{B} \leftarrow \mathcal{B}_1}^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\therefore [T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -3 & -1 \end{pmatrix}$$

We now find $[T]_{\mathcal{B}_2}$, using $[T]_{\mathcal{B}_1}$ by $[T]_{\mathcal{B}_2} = R^{-1}[T]_{\mathcal{B}}R$, where $R = P_{\mathcal{B} \leftarrow \mathcal{B}_2} = P_{\mathcal{B}_2}$

$$\therefore R = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}, \text{ so that } R^{-1} = 1/3 \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$$

$$\text{Hence } [T]_{\mathcal{B}_2} = 1/3 \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = 1/3 \begin{pmatrix} 2 & 1 \\ 19 & 4 \end{pmatrix}$$

This is the same as the one obtained before.

Example 16.24. Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_1$ be defined by $T(a + bx + cx^2) = (a + b + c) + (a + 2b + 3c)x$, let $\mathcal{B}_1 = \{v_1, v_2, v_3\}$, $\mathcal{B}_2 = \{w_1, w_2\}$, where $v_1 = x^2$, $v_2 = -1 + x$, $v_3 = 1 + x$, $w_1 = 1 + x$, $w_2 = x$ be the basis for \mathbb{P}_2 and \mathbb{P}_1 respectively. Find the matrix of T relative to \mathcal{B}_1 and \mathcal{B}_2 . Also compute $T(1 + 2x + 3x^2)$ using the matrix $[T]_{\mathcal{B}_1 \mathcal{B}_2}$ and directly.

Solution: We know that $[T]_{\mathcal{B}_1 \mathcal{B}_2} = \begin{pmatrix} [T(v_1)]_{\mathcal{B}_2} & [T(v_2)]_{\mathcal{B}_2} & [T(v_3)]_{\mathcal{B}_2} \end{pmatrix}$

Thus we have to find $[T(v_i)]_{\mathcal{B}_2}$, $i = 1, 2, 3$.

$$T(v_1) = T(x^2) = 1 + 3x$$

$$\Rightarrow [T(v_1)] = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$T(v_2) = T(-1 + x) = -1 + 1 + (-1 + 2)x = x$$

$$\Rightarrow [T(v_2)] = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$T(v_3) = T(1 + x) = 1 + 1 + (1 + 2)x = 2 + 3x$$

$$\Rightarrow [T(v_3)] = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

To find $[T(v_1)]_{\mathcal{B}_2}$, $[T(v_2)]_{\mathcal{B}_2}$, $[T(v_3)]_{\mathcal{B}_2}$ we reduce the matrix

$A = \begin{pmatrix} [w_1] & [w_2] & : & [T(v_1)] & [T(v_2)] & [T(v_3)] \end{pmatrix}$ to the reduced echelon form

$$A = \begin{pmatrix} 1 & 0 & : & 1 & 0 & 2 \\ 1 & 1 & : & 3 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & : & 1 & 0 & 2 \\ 0 & 1 & : & 2 & 1 & 1 \end{pmatrix}$$

$$\text{Hence } [T]_{\mathcal{B}_1\mathcal{B}_2} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix}$$

$$\text{Since } [T(v)]_{\mathcal{B}_2} = [T]_{\mathcal{B}_1\mathcal{B}_2}[v]_{\mathcal{B}_1}$$

$$\therefore [T(1+2x+3x^2)]_{\mathcal{B}_2} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix} (1+2x+3x^2)_{\mathcal{B}_1}$$

To find $(1+2x+3x^2)_{\mathcal{B}_1}$, we form the matrix $B = ([v_1] \ [v_2] \ [v_3] : [v])$ and obtain its row reduced echelon form. Now

$$B = \begin{pmatrix} 0 & -1 & 1 & : & 1 \\ 0 & 1 & 1 & : & 2 \\ 1 & 0 & 0 & : & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & : & 3 \\ 0 & 1 & 0 & : & \frac{1}{2} \\ 0 & 0 & 1 & : & \frac{3}{2} \end{pmatrix}$$

$$\therefore [v]_{\mathcal{B}_1} = \begin{pmatrix} 3 \\ \frac{1}{2} \\ \frac{3}{2} \end{pmatrix}$$

$$\therefore [v]_{\mathcal{B}_2} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ \frac{1}{2} \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$$

$$\text{Hence } T(1+2x+3x^2) = 6w_1 + 8w_2 = 6(1+x) + 8x = 6 + 14x$$

$$\text{Directly } T(1+2x+3x^2) = (1+2+3) + (1+4+9)x = 6 + 14x$$

Hence the matrix $[T]_{\mathcal{B}_1\mathcal{B}_2}$ is verified.

Example 16.25. Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be a linear transformation and $\mathcal{B}_1 = \{(1, 1, 0)^t, (0, 1, 1)^t, (1, 0, 1)^t\}$, $\mathcal{B}_2 = \{(1, 2)^t, (2, 1)^t\}$ be ordered basis of \mathbb{R}^3 and \mathbb{R}^2 respectively. If matrix of T relative to $\mathcal{B}_1, \mathcal{B}_2$ is $\begin{pmatrix} 0 & 1 & \frac{1}{3} \\ 0 & -1 & \frac{1}{3} \end{pmatrix}$. Determine T .

Solution: Let P, Q be the transition matrices of $\mathcal{B}_1, \mathcal{B}_2$ to the standard basis, then

$$P = P_{\mathcal{B}_1} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$Q = P_{\mathcal{B}_2} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

If A is the standard matrix of T , then

$$Q^{-1}AP = T_{\mathcal{B}_1\mathcal{B}_2}$$

$$\Rightarrow A = Q \begin{pmatrix} 0 & 1 & \frac{1}{3} \\ 0 & -1 & \frac{1}{3} \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & \frac{1}{3} \\ 1 & -1 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

To find P^{-1} transform $(P : I)$ to $(I : R)$

$$\text{Then } P^{-1} = R \text{ thus } (P : I) = \begin{pmatrix} 1 & 0 & 1 & : & 1 & 0 & 0 \\ 1 & 1 & 0 & : & 0 & 1 & 0 \\ 0 & 1 & 1 & : & 0 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & : & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & : & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & : & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\text{Thus } P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\text{So } A = \frac{1}{2} \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow T \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_3 \end{pmatrix}$$

This defines T .

Example 16.26. Let V be a n -dimensional vector space and $\mathcal{B}_1, \mathcal{B}_2$ two bases for V . If I is the identity operator on V . Then prove that $[I]_{\mathcal{B}_1\mathcal{B}_2} = P_{\mathcal{B}_2\leftarrow\mathcal{B}_1}$.

Solution: Let $\mathcal{B}_1 = \{v_1, v_2, \dots, v_n\}$, $\mathcal{B}_2 = \{w_1, w_2, \dots, w_n\}$. Then
 $[I]_{\mathcal{B}_1\mathcal{B}_2} = \begin{pmatrix} [T(v_1)]_{\mathcal{B}_2} & [T(v_2)]_{\mathcal{B}_2} & \dots & [T(v_n)]_{\mathcal{B}_2} \\ [v_1]_{\mathcal{B}_2} & [v_2]_{\mathcal{B}_2} & \dots & [v_n]_{\mathcal{B}_2} \end{pmatrix}$
 $= \begin{pmatrix} [v_1]_{\mathcal{B}_2} & [v_2]_{\mathcal{B}_2} & \dots & [v_n]_{\mathcal{B}_2} \end{pmatrix} = P_{\mathcal{B}_2\leftarrow\mathcal{B}_1}$... (by the definition of transition matrix)
Hence $[I]_{\mathcal{B}_1\mathcal{B}_2} = P_{\mathcal{B}_2\leftarrow\mathcal{B}_1}$.

Remark 16.1. The preceding problem shows that the matrix of the identity operator need not be the identity matrix.

16.7 Exercise

- Let $\mathcal{B}_1 = \{v_1, v_2\}$ and $\mathcal{B}_2 = \{w_1, w_2\}$ be the bases for vector spaces V and W respectively. Let $T : V \rightarrow W$ be a linear transformation such that
 $T(v_1) = 2w_1 - 3w_2$
 $T(v_2) = -w_1 + 4w_2$
Find the matrix of T relative to the basis \mathcal{B}_1 and \mathcal{B}_2 .
- Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be defined by
 $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
Let $\mathcal{B} = \{e_1, e_2\}$, $\mathcal{B}' = \{(1, 2), (1, 1)\}$.
Find $[T]_{\mathcal{B}}$, $[T]_{\mathcal{B}'}$, $[T]_{\mathcal{B}\mathcal{B}'}$, $[T]_{\mathcal{B}'\mathcal{B}}$.
- Let $D : \mathbb{P}_2 \rightarrow \mathbb{P}_2$ be the differentiation operator. Let \mathcal{B} be the standard basis $\{1, x, x^2\}$ and $\mathcal{B}' = \{1, 1+x, (1+x)^2\}$
Find $[D]_{\mathcal{B}}$, $[D]_{\mathcal{B}'}$, $[D]_{\mathcal{B}\mathcal{B}'}$.
- Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ be the linear transformation whose matrix relative to the standard basis is
 $\begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$
Find $T(e_1)$, $T(e_2)$, $T(e_3)$ where $\{e_1, e_2, e_3\}$ is the standard basis for \mathbb{R}^3 .
- Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear operator whose matrix relative to the basis \mathcal{B} , \mathcal{B}' is
 $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 3 & 1 \end{pmatrix}$

where \mathcal{B} is the standard basis and $\mathcal{B}' = \{v_1, v_2, v_3\}$ with $v_1 = (1, 0, 1)^t$, $v_2 = (0, 1, 1)^t$, $v_3 = (1, 1, 1)^t$.

Find $T(e_1)$, $T(e_2)$, $T(e_3)$ relative to the standard basis of \mathbb{R}^3 .

6. Let $A = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 2 & 5 \\ 2 & 1 & 3 \end{pmatrix}$ and let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a linear operator defined

by $T(x) = Ax$. If $\mathcal{B} = \{e_1, e_2, e_3\}$ is the standard ordered basis of \mathbb{R}^3 and $\mathcal{B}' = \{e_2, e_3, e_1\}$.

(i) Find $[T]_{\mathcal{B}'}$.

(ii) Find $[T]_{\mathcal{B}'\mathcal{B}}$.

(iii) Find $[T]_{\mathcal{B}\mathcal{B}'}$.

7. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear operator whose matrix relative to $\mathcal{B}' = \{v_1, v_2\}$, $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is $\begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$

Find the matrix of T relative to the standard basis.

8. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

Find the matrix of T relative to basis \mathcal{B} , \mathcal{B}' where

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}.$$

9. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear operator defined by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}. \quad \text{Let } \mathcal{B}_1 = \{v_1, v_2\}, \quad \mathcal{B}_2 = \{w_1, w_2\}, \text{ where}$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad w_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Find the matrix of T relative to

(i) \mathcal{B}_1 .

(ii) \mathcal{B}_2 .

(iii) \mathcal{B}_1 and \mathcal{B}_2 .

(iv) \mathcal{B}_2 and \mathcal{B}_1 .

10. Let $T : \mathbb{R}^2 \rightarrow \mathbb{M}_{2 \times 2}(\mathbb{R})$ be defined by

$$T \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b & a-b \\ 2a-b & a+2b \end{pmatrix}$$

Find the matrix of T relative to the standard ordered basis.

11. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_3$ be defined by $T(a + bt) = at^2 + bt^3$

Let $\mathcal{B}_1 = \{x, x+1\}$ be basis for \mathbb{P}_1 and $\mathcal{B}_2 = \{x+1, x^2-1, x^3\}$ be basis for \mathbb{P}_3 . Find the matrix of T relative to the \mathcal{B}_1 and \mathcal{B}_2

12. Let $T : \mathbb{M}_{2 \times 2}(\mathbb{C}) \rightarrow \mathbb{M}_{2 \times 2}(\mathbb{C})$ be defined by $T(A) = A^t$.

Let $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ and \mathcal{B}_2 be the standard basis $\{E_{11}, E_{12}, E_{21}, E_{22}\}$. Find the matrix of T relative to \mathcal{B}_1 and \mathcal{B}_2

13. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined $T \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $T \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
- Find the matrix of T relative to the ordered basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$
 - Find $T \begin{pmatrix} a \\ b \end{pmatrix}$
14. Let $\mathbb{P}_1 \rightarrow \mathbb{P}_2$ be the transformation defined by $T(p(x)) = (x-1)p(x)$, for all $p(x) \in \mathbb{P}_1$
- Show that T is a linear transformation.
 - Find $[T]_{\mathcal{B}_1 \mathcal{B}_2}$, where \mathcal{B}_1 and \mathcal{B}_2 are the standard ordered bases for \mathbb{P}_1 and \mathbb{P}_2 respectively, namely $\{1, x\}$ and $\{1, x, x^2\}$ respectively.
15. Let $\mathbb{P}_1 \rightarrow \mathbb{P}_1$ be a linear operator whose matrix relative to basis $\mathcal{B} = \{1, -x, x\}$ is $\begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}$
- Find the matrix of T relative to the standard basis $\{1, x\}$.
 - Find $T(2 - 3x)$.
16. Let $T : \mathbb{M}_{2 \times 2}(\mathbb{R}) \rightarrow \mathbb{R}^3$ be defined by $T \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+d \\ b+c \\ a+b+c+d \end{pmatrix}$
- Prove that T is a linear transformation.
 - Find the matrix of T relative to the standard basis for $\mathbb{M}_{2 \times 2}(\mathbb{R})$, namely $\{E_{11}, E_{12}, E_{21}, E_{22}\}$.
17. Let $t : \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}(\mathbb{R})$ be defined by $T(x + iy) = x - iy$. Let $\mathcal{B}_1 = \{1 + i, 1 - i\}$, $\mathcal{B}_2 = \{1, 1 + i\}$. Find the matrix of T relative to
- \mathcal{B}_1 and \mathcal{B}_2 .
 - \mathcal{B}_2 and \mathcal{B}_1 .
 - \mathcal{B}_1 .
 - \mathcal{B}_2 .
18. Let $T : \mathbb{M}_{2 \times 2}(R) \rightarrow \mathbb{M}_{2 \times 2}(R)$ be defined by $T(A) = AB - BA$, where $B = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$. Let $\mathcal{B}_1 = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ and $\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$. Find the matrix of T relative to bases $\mathcal{B}_1, \mathcal{B}_2$
19. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear operator and $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$ be an ordered basis of \mathbb{R}^2 . If the matrix of T relative to \mathcal{B} is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Determine T . What is special about the vectors $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

20. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ be defined by $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \\ 0 \end{pmatrix}$. Let $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, $\mathcal{B}_1' = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ and \mathcal{B}_2 be the standard ordered basis and $\mathcal{B}_2' = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ be another basis for \mathbb{R}^4 .

(i) Find the matrix $[T]_{\mathcal{B}_1 \mathcal{B}_2}$.

(ii) Find the matrix $[T]_{\mathcal{B}_1' \mathcal{B}_2'}$.

(iii) Find $T \begin{pmatrix} -3 \\ 3 \end{pmatrix}$ using the definition of T and using the matrix obtained in part (a) and (b).

21. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_1$ be a linear transformation. Suppose that the matrix of T relative to basis $\mathcal{B}_1 = \{v_1, v_2\}$ is $\begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix}$, where $v_1 = 1 + 2x$, $v_2 = 1 - x$.

(i) Find $[T(v_1)]_{\mathcal{B}_1}$ and $[T(v_2)]_{\mathcal{B}_1}$.

(ii) Find $T(v_1)$, $T(v_2)$.

(iii) Find $T(6 + 9x)$.

22. Let $\mathcal{B} = \{v_1, v_2\}$ and $\mathcal{B}' = \{w_1, w_2\}$ be two ordered basis for \mathbb{R}^2 , where $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $w_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $w_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear operator such that $[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$.

(i) Find $[T]_{\mathcal{B}'}$ from $[T]_{\mathcal{B}}$.

(ii) Find matrix of T relative to standard bases of \mathbb{R}^2 .

(iii) Obtain $[T]_{\mathcal{B}'}$ from the standard matrix of T .

16.8 Supplementary Exercises

1. State whether the following are true or false. Justify the false ones:

(i) The coordinate vector of $3x^3 - 3x^2 + 2$ is relative to the ordered basis

$$\{1, x, x^2, x^3\} \text{ is } \begin{pmatrix} 4 \\ -3 \\ 0 \\ 2 \end{pmatrix}.$$

- (ii) The coordinate mapping from a n -dimensional V to \mathbb{R}^n is a linear transformation.
- (iii) The coordinates of $\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ relative to the basis $\{E_{22}, E_{21}, E_{12}, E_{11}\}$ is $\begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \end{pmatrix}$, where E_{ij} is the 2×2 , matrix with (i, j) th entry 1 and all others 0.
- (iv) The transition matrix can be a singular matrix.
- (v) If V is a finite dimensional vector space and \mathcal{B} is any basis then for all $v \in V$, $[v]_{\mathcal{B}}$ is well defined.
- (vi) If \mathcal{B}_1 and \mathcal{B}_2 are two ordered bases of an n -dimensional vector space V , the the matrix $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ is the matrix of identity transformation from V with basis \mathcal{B}_1 to V with basis \mathcal{B}_2 .
- (vii) If $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ and $\mathcal{B}_2 = \{w_1, w_2, w_3\}$ are two ordered bases for \mathbb{R}^3 , then the last three columns of the reduced echelon form of the matrix $[v_1 \ v_2 \ v_3 \ w_1 \ w_2 \ w_3]$ is $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$.
- (viii) If $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ and $\mathcal{B}_2 = \{w_1, w_2, w_3\}$ are two ordered bases for \mathbb{R}^3 and the matrix $[v_1 \ v_2 \ v_3 \ w_1 \ w_2 \ w_3]$ is reduced to $[A \ I]$ then the matrix A is $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$.
- (ix) The matrix $P_{\mathcal{B}}$ denotes the transition matrix from the standard basis to basis \mathcal{B} .
- (x) The matrix $P_{\mathcal{B} \leftarrow \mathcal{B}}$ is $P_{\mathcal{B}}$.
- (xi) If $\mathcal{B}_1, \mathcal{B}_2$ and \mathcal{B}_3 are three ordered bases of an n -dimensional vector space V , and $A = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ and $B = P_{\mathcal{B}_3 \leftarrow \mathcal{B}_2}$ then $P_{\mathcal{B}_3 \leftarrow \mathcal{B}_1} = AB$.
- (xii) If $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ are three ordered bases of a finite dimensional vector space V , and $A = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$, $B = P_{\mathcal{B}_3 \leftarrow \mathcal{B}_2}$, $C = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_3}$, then $ABC = I$.
- (xiii) If T is a linear transformation from a m -dimensional vector space to a n -dimensional vector space, then the matrix of T is a $m \times n$ matrix.
- (xiv) If $\mathcal{B}_1, \mathcal{B}_2$ are ordered bases of vector spaces $V(F)$ and $W(F)$ respectively, where V and W are m - and n -dimensional vector spaces over F and $T : V \rightarrow W$ is a linear transformation, and $A = [T]_{\mathcal{B}_1 \mathcal{B}_2}$, then multiplication by A defines a mapping from V to W .
- (xv) The matrix of a linear operator is always non-singular.
- (xvi) The matrix of the identity operator of a finite dimensional vector space is always the identity matrix.
- (xvii) If T is a linear operator on a finite dimensional vector space V , and $\mathcal{B}_1, \mathcal{B}_2$ are two ordered bases for V , the $[T]_{\mathcal{B}_1} = P^{-1}[T]_{\mathcal{B}_2}P$, where $P = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$.
- (xviii) If T is a linear operator on a finite dimensional vector space V and $\mathcal{B}_1, \mathcal{B}_2$ are two ordered bases for V , then $[T]_{\mathcal{B}_1 \mathcal{B}_2} [T]_{\mathcal{B}_2 \mathcal{B}_1} = I$.

- (xix) If $\mathcal{B}_1, \mathcal{B}_2$ are two ordered bases of a finite dimensional vector space V , the the matrix of the zero transformation on V relative to $\mathcal{B}_1, \mathcal{B}_2$ is always the null matrix.
- (xx) Let T be a linear operator on R^2 and $\mathcal{B} = \{v_1, v_2\}$ be an ordered basis. If $[T]_{\mathcal{B}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $[T]_{\mathcal{B}'} = [T]_{\mathcal{B}}^t$, where $\mathcal{B}' = \{v_2, v_1\}$.
2. Let $\mathcal{B} = \{v_1, v_2, v_3, v_4\}$ be a basis for \mathbb{R}^4 , where
- $$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix}, v_4 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$
- Let $v = \begin{pmatrix} 1 \\ 2 \\ -6 \\ 2 \end{pmatrix}$, find
- $P_{\mathcal{B}}$, transition matrix from \mathcal{B} to the standard basis.
 - $[v]_{\mathcal{B}}$ directly.
 - verify the relation $[u] = P_{\mathcal{B}}[u]_{\mathcal{B}}$ for the v .
3. Let $T : \mathbb{R}^3 \rightarrow \mathbb{M}_{2 \times 2}(R)$ be defined by $T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 + x_3 & 2x_1 - 3x_2 \\ 3x_1 + 4x_2 & 2x_2 - x_3 \end{pmatrix}$. Find the matrix of T relative to the standard basis.
4. If $\begin{pmatrix} 1 & 2 \\ -3 & -5 \end{pmatrix}$ is change of coordinates matrix from a basis \mathcal{B} to the standard basis, find \mathcal{B} in \mathbb{P}_1 .
5. If $\begin{pmatrix} 1 & 2 \\ -3 & -5 \end{pmatrix}$ is the change of coordinates matrix from the standard basis $\{1, x\}$ to some basis \mathcal{B} , find \mathcal{B} in \mathbb{P}_1 .
6. Let \mathcal{B} be the standard basis for \mathbb{R}^3 and $\mathcal{B}_1, \mathcal{B}_2$ two other bases given by $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ and $\mathcal{B}_2 = \{w_1, w_2, w_3\}$, where $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $w_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$, $w_2 = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$, $w_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.
- Find the transition matrix of
- \mathcal{B}_1 relative to \mathcal{B} .
 - \mathcal{B}_2 relative to \mathcal{B} .
 - \mathcal{B} relative to \mathcal{B}_1 .
 - \mathcal{B} relative to \mathcal{B}_2 .
 - \mathcal{B}_1 relative to \mathcal{B}_2 .

- (vi) \mathcal{B}_2 relative to \mathcal{B}_1 .
- (vii) Find $[v]_{\mathcal{B}_2}$, if $[v]_{\mathcal{B}_1} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$
- (viii) Find $[v]_{\mathcal{B}_1}$, if $[v]_{\mathcal{B}_2} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$
7. Let $\mathcal{B}_1 = \{v_1, v_2, v_3\}$ and $\mathcal{B}_2 = \{w_1, w_2, w_3\}$ be two bases for \mathbb{R}^3 , where $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $w_1 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$, $w_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $w_3 = \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$. Find the transition matrix from
- \mathcal{B}_2 to \mathcal{B}_1 .
 - \mathcal{B}_1 to \mathcal{B}_2 .
 - Standard basis to \mathcal{B}_1 .
 - Standard basis to \mathcal{B}_2 .
 - \mathcal{B}_1 to standard basis.
 - \mathcal{B}_2 to standard basis.
8. Let $\mathcal{B}_1 = \{x^2 + 1, x - 2x^2, x + 3x^2\}$ and $\mathcal{B}_2 = \{2 + x, 3x^2 + 1, x\}$ be bases for \mathbb{P}_2 . Let $p_1 = 6x^2 - 4x + 8$ and $p_2 = 9x^2 - x + 7$.
- Find the coordinate vectors of p_1 and p_2 with respect to basis \mathcal{B}_2 .
 - Find the transition matrix $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$ from \mathcal{B}_2 - basis to \mathcal{B}_1 - basis.
 - Find the coordinate vectors of p_1 and p_2 with respect to \mathcal{B}_1 - basis using $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$
 - Obtain (c) by direct calculation.
 - $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$, the transition matrix from \mathcal{B}_1 - basis to \mathcal{B}_2 - basis
 - Obtain (a) using $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$
9. Let $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\}$ and $\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$ be two ordered bases for $\mathbb{M}_{2 \times 2}(\mathbb{R})$. Let $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Find
- coordinate vector of M with respect to basis \mathcal{B}_2 .
 - $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$, the transition matrix from \mathcal{B}_2 - basis to \mathcal{B}_1 - basis.
 - coordinate vector of M with respect to basis \mathcal{B}_1 using transition matrix.
 - Transition matrix $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$

- (v) What is the relation between the matrices obtained in (b) and (d).
10. Let $\mathcal{B} = \{1+x+x^2, x+x^2, x^2\}$ be basis for \mathbb{P}_2 . By finding the transition matrix from \mathcal{B} to the standard basis for \mathbb{P}_2 , find the polynomial p whose coordinate vector relative to \mathcal{B} is $\begin{pmatrix} 3 \\ -2 \\ -5 \end{pmatrix}$.
11. Let $T : \mathbb{P}_1 \rightarrow \mathbb{P}_2$ be defined by $T(a+bx) = (a+b) + (a-b)x + bx^2$ and let $\mathcal{B}_1 = \{1, x\}$, $\mathcal{B}_2 = \{1, x, x^2\}$ be the standard ordered bases for \mathbb{P}_1 and \mathbb{P}_2 respectively. Let $\mathcal{B}_3 = \{1-x, 1+2x\}$ and $\mathcal{B}_4 = \{1+x^2, x+x^2, 1+x+x^2\}$ be a basis for \mathbb{P}_1 and \mathbb{P}_2 respectively.
- $[T]_{\mathcal{B}_1\mathcal{B}_2}$
 - $[T]_{\mathcal{B}_1\mathcal{B}_4}$
 - $[T]_{\mathcal{B}_3\mathcal{B}_4}$
12. Let $\mathcal{B}_1 = \{v_1, v_2, v_3\}$, $\mathcal{B}_2 = \{w_1, w_2\}$ be bases for vector spaces V and W respectively. Let $T : v \rightarrow W$ be a linear transformation such that $T(v_1) = 2w_1 - 3w_2$, $T(v_2) = 3w_1$, $T(v_3) = 4w_2$. Find $[T]_{\mathcal{B}_1\mathcal{B}_2}$.
13. Let $\mathcal{B} = \{v_1, v_2, v_3\}$ be an ordered basis of a vector space V . Define $T : V \rightarrow V$ by $T(v_1) = v_2$, $T(v_2) = v_3$, $T(v_3) = 0$.
- Find $[T]_{\mathcal{B}}$, $[T^2]_{\mathcal{B}}$, $[T^3]_{\mathcal{B}}$
 - Deduce from (i) that $T^2 \neq 0$, $T^3 = 0$.
14. Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_4$ be a transformation defined by $T(p(x)) = (x^2 + 5x - 1)P(x)$.
- Show that T is a linear transformation.
 - Find the matrix of T relative to the standard bases for \mathbb{P}_2 and \mathbb{P}_4 .
15. Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_2$ be a linear transformation whose matrix relative to the basis $\mathcal{B} = \{1, (x+1), (x+1)^2\}$ is $\begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 1 \\ 2 & 3 & 1 \end{pmatrix}$. Find the image of $(x^2 + 2)$ under T .
16. Let $T : \mathbb{P}_2 \rightarrow \mathbb{R}^3$ be defined by $T(p(x)) = \begin{pmatrix} p(1) \\ p(-1) \\ p(2) \end{pmatrix}$.
- Show that T is a linear transformation.
 - Find the matrix of T relative to the standard bases $\{1, x, x^2\}$ and $\{e_1, e_2, e_3\}$ of \mathbb{P}_2 and \mathbb{R}^3 respectively.
17. If a basis \mathcal{B} consists of the eigen vectors of a linear operator prove that $[T]_{\mathcal{B}}$ is a diagonal matrix.
18. Let A be the matrix of the linear operator T , relative to an ordered basis $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$

- (i) If $\mathcal{B}_1 = \{v_2, v_1, v_3, \dots, v_n\}$, can $[T]_{\mathcal{B}_1}$ be obtained from $[T]_{\mathcal{B}}$? If yes, how.
- (ii) If \mathcal{B}_2 is the basis obtained from \mathcal{B} by interchanging the i th and j th vectors in the basis \mathcal{B} , what is the relation between $[T]_{\mathcal{B}_2}$ and $[T]_{\mathcal{B}}$?
19. Let $\mathcal{B} = \{e_1, e_2, e_3\}$ be the standard basis for \mathbb{R}^3 and \mathcal{B}' be a basis for a vector space V and $T : \mathbb{R}^3 \rightarrow V$ be a linear transformation such that
- $$T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (x_2 - x_1)v_1 + (x_1 + x_3)v_2 + 2x_1v_3 + 3x_3v_4$$
- (i) Find $T(e_1)$, $T(e_2)$, $T(e_3)$.
- (ii) Find $[T(e_1)]_{\mathcal{B}'}$, $[T(e_2)]_{\mathcal{B}'}$, $[T(e_3)]_{\mathcal{B}'}$.
- (iii) Find the matrix of T relative to \mathcal{B} , \mathcal{B}' .
20. Let V be a n -dimensional vector space and I be the identity operator on V . If \mathcal{B} is the standard basis for V and \mathcal{B}_1 any other basis for V , find
- (i) $[I]_{\mathcal{B}_1\mathcal{B}_1}$.
- (ii) $[I]_{\mathcal{B}\mathcal{B}}$.
- (iii) $[I]_{\mathcal{B}_1\mathcal{B}}$.
- (iv) $[I]_{\mathcal{B}\mathcal{B}_1}$.
21. Let $T : \mathbb{R}^3 \rightarrow \mathbb{P}_3$ be defined by $T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (5x_1 - 2x_2) + (x_1 - 3x_2)x + (2x_1 - x_2 + x_3)x^2 + 4x_3x^3$.
- (i) Prove that T is a linear transformation.
- (ii) Find $[T]_{\mathcal{B}_1\mathcal{B}_2}$, where $\mathcal{B}_1 = \{e_1, e_2, e_3\}$ and $\mathcal{B}_2 = \{1, x, x^2, x^3\}$ are the standard bases for \mathbb{R}^3 and \mathbb{P}_3 respectively.
22. Let $D : \mathbb{P}_3 \rightarrow \mathbb{P}_3$ be the differentiation operator. Let $\mathcal{B} = \{1, x, x^2, x^3\}$ be the standard ordered basis for \mathbb{P}_3 and $\mathcal{B}' = \{1, 1+x, (1+x)^2, (1+x)^3\}$ be the another ordered basis for \mathbb{P}_3 . Find $[D]_{\mathcal{B}}$ and $[D]_{\mathcal{B}'}$.
23. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by $T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_1 + x_3 \\ -2x_1 + x_2 \\ -x_1 + 2x_2 + 4x_3 \end{pmatrix}$.
- Let \mathcal{B} be the standard basis for \mathbb{R}^3 and $\mathcal{B}' = \{v_1, v_2, v_3\}$, where $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$. Find
- (i) $[T]_{\mathcal{B}}$.
- (ii) $[T]_{\mathcal{B}'}$.
- (iii) Give a rule to define T^{-1} .
24. Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ be a basis of a vector space V . Let T be the linear transformation on V such that
- $$T(v_i) = v_{i+1}, \quad i = 1, 2, 3, \dots, (n-1)$$
- $$T(v_n) = 0.$$

Find

- (i) $[T]_{\mathcal{B}}$.
- (ii) $[T^n]_{\mathcal{B}}$.
- (iii) Is $T^n = 0$? Justify.

25. Let v_1 and v_2 be the eigenvectors of matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Find the matrix of the transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ determined by the matrix A .

16.9 Answers to Exercises

Exercise - 16.4

1. (i) $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$
 (ii) $\begin{pmatrix} -1 \\ -2 \end{pmatrix}$
 (iii) $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$
 (iv) $\begin{pmatrix} 20 \\ 23 \end{pmatrix}$
 (v) $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$
2. (i) $(1 \ 0 \ -3 \ 1)^t$
 (ii) $(1 \ -3 \ 0 \ 1)^t$
 (iii) $(-3 \ 1 \ 1 \ 0)^t$
3. (i) $(1 \ -1 \ 2 \ 0)^t$, $(2 \ -3 \ -6 \ 5)^t$
 (ii) $(1 \ 2 \ -1 \ 0)^t$, $(2 \ -6 \ -3 \ 5)^t$
4. (i) $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$
 (ii) $(3 \ -1 \ -2 \ 1)^t$
5. (i) $(2 \ 0)^t$
 (ii) $(\frac{1}{4} \ -\frac{3}{4})^t$
 (iii) $\begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}$
 (iv) $\begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & \frac{1}{4} \end{pmatrix}$
6. $(-2 \ 5 \ -1)^t$

7. (i) $\begin{pmatrix} 1 & -1 \end{pmatrix}^t$
 (ii) $\begin{pmatrix} 1 & -3 \end{pmatrix}^t$
8. $\{(1 \ 3)^t, (2 \ -5)^t\}$
9. $\{(-5 \ 3)^t, (-2 \ 1)^t\}$ *Hint: elements of \mathcal{B} are columns of P_{-1}*
10. (i) $\begin{pmatrix} 3 & -1 \end{pmatrix}^t$
 (ii) $\begin{pmatrix} 1 & -1 \end{pmatrix}^t$
 (iii) $\begin{pmatrix} -2 & -7 \\ 3 & 10 \end{pmatrix}$
 (iv) $\begin{pmatrix} 10 & 7 \\ -3 & -2 \end{pmatrix}$
11. (i) $\frac{1}{2} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$
 (ii) $\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$
12. (i) $\frac{1}{3} \begin{pmatrix} 2 & 4 \\ 1 & -1 \end{pmatrix}$
 (ii) $\frac{1}{2} \begin{pmatrix} 1 & 4 \\ 1 & -2 \end{pmatrix}$
13. $v_1 = (1 \ 1)^t, v_2 = (5 \ -2)^t$
14. (i) $\{(-1 \ 3)^t, (-1 \ 2)^t\}$
 (ii) $\{(5 \ -1)^t, (-3 \ 1)^t\}$
15. (i) $(-i \ 1)^t$
 (ii) $(2 - i \ i)^t$
16. $(3 \ 7 \ 3)^t$
17. (i) $(2 \ -2 \ 1 \ -1)^t$
 (ii) $(1 \ -\frac{4}{3} \ \frac{5}{3} \ -\frac{2}{3})^t, P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & 0 \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$
18. (i) $\begin{pmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{3}{2} & 0 \\ 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
 (ii) $C = \frac{1}{2} \begin{pmatrix} 0 & 2 & 3 & -3 \\ -2 & -2 & -2 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

$$(iii) A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & -1 \\ -1 & 1 & 0 & 0 \end{pmatrix}$$

$$(iv) B = \begin{pmatrix} -1 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & -1 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

(v) *Hint:* $BA = C$

$$19. (i) \frac{1}{3} \begin{pmatrix} 3 & 0 & 0 & 3 \\ 1 & 2 & -2 & 0 \\ 1 & -1 & 1 & 0 \\ -1 & 1 & 2 & 0 \end{pmatrix}$$

$$(ii) (1 \ \frac{1}{3} \ \frac{1}{3} \ \frac{1}{3} \ \frac{2}{3})^t, (1 \ 1 \ 1 \ 0)^t$$

Exercise - 16.7

$$1. \begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix}$$

$$7. \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}$$

$$8. \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

$$9. (i) \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$(ii) \begin{pmatrix} d & c \\ b & a \end{pmatrix}$$

$$(iii) \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

$$(iv) \begin{pmatrix} b & a \\ d & c \end{pmatrix}$$

$$10. \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}$$

$$11. \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$12. \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$13. \quad \text{(i)} \begin{pmatrix} 0 & \frac{3}{2} \\ 1 & \frac{1}{2} \end{pmatrix}$$

$$\text{(ii)} \begin{pmatrix} (3a-b)/2 \\ -b \end{pmatrix}$$

$$14. \quad \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$$

$$15. \quad \text{(i)} \begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}$$

$$\text{(ii)} -7x$$

$$16. \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$17. \quad \text{(i)} \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}$$

$$\text{(ii)} \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 1 \end{pmatrix}$$

$$\text{(iii)} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\text{(iv)} \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

$$18. \quad \begin{pmatrix} 0 & -2 & 2 & 0 \\ -1 & -2 & -1 & 0 \\ -1 & -2 & -1 & 0 \\ 1 & 8 & 0 & 0 \end{pmatrix}$$

$$19. \quad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}.$$

v_1, v_2 are eigenvectors of T and 1, -1 are corresponding eigenvalues.

$$20. \quad \text{(i)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\text{(ii)} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}$$

$$\text{(iii)} \begin{pmatrix} -3 \\ 3 \\ 0 \\ 0 \end{pmatrix}$$

$$21. \quad \text{(i)} \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \end{pmatrix}$$

- (ii) $1 + 5x, 1 - 10x$
 (iii) $6 + 15x$
22. (i) $\frac{1}{3} \begin{pmatrix} 2 & -1 \\ 19 & 4 \end{pmatrix}$
 (ii) $\begin{pmatrix} 3 & 2 \\ -3 & -1 \end{pmatrix}$

Supplementary Exercises-16.8

1. (i) F, \mathcal{B} must be ordered basis.
 (ii) F, $(2 \ 0 \ -3 \ 4)^t$
 (iii) T
 (iv) F, $(3 \ -1 \ 2 \ 1)^t$
 (v) F, always non-singular.
 (vi) T
 (vii) F, $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$
 (viii) T
 (ix) F, from \mathcal{B} - basis to standard basis.
 (x) F
 (xi) F, $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_3} = BA$
 (xii) F, $CBA = I$
 (xiii) F, $n \times m$
 (xiv) F, mapping from \mathbb{R}^n to \mathbb{T}^m
 (xv) F, matrix of the zero operator is the null matrix which is singular.
 (xvi) F
 (xvii) F, $P = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$
 (xviii) F, $[T]_{\mathcal{B}_1 \mathcal{B}_2}$ can be singular.
 (xix) T
 (xx) F, $\begin{pmatrix} b & a \\ d & c \end{pmatrix} = [T]_{\mathcal{B}'}$

$$2. \quad (i) \begin{pmatrix} 1 & 2 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

$$(ii) (3 \ -1 \ -2 \ 1)^t$$

$$3. \quad \begin{pmatrix} -1 & 0 & 1 \\ 2 & -3 & 0 \\ 3 & 4 & 0 \\ 0 & 2 & -1 \end{pmatrix}$$

$$4. \quad \{1 + 3x, 2 - 5x\}$$

$$5. \quad \{-5 + 3x, -2 + x\}$$

$$6. \quad (v) \frac{1}{5} \begin{pmatrix} 1 & 6 & 3 \\ 1 & 1 & -2 \\ 2 & -3 & 1 \end{pmatrix}$$

$$(vi) \frac{1}{2} \begin{pmatrix} 1 & 3 & 3 \\ 1 & 1 & -1 \\ 1 & -3 & 1 \end{pmatrix}$$

$$(vii) \frac{1}{5}(1 \ -4 \ 7)^t$$

$$(viii) (2 \ -1 \ 3)^t$$

$$7. \quad (i) \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

$$8. \quad (i) [p_1]_{\mathcal{B}_2} = \begin{pmatrix} 3 \\ 2 \\ -7 \end{pmatrix}, \quad [p_2]_{\mathcal{B}_2} = \begin{pmatrix} 2 \\ 3 \\ -3 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 2 & 1 & 0 \\ 1 & -\frac{2}{5} & \frac{3}{5} \\ 0 & \frac{2}{5} & \frac{3}{5} \end{pmatrix}$$

$$(iii) [p_1]_{\mathcal{B}_1} = \begin{pmatrix} 8 \\ -2 \\ -2 \end{pmatrix}, \quad [p_2]_{\mathcal{B}_1} = \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix}$$

(iv) Same as (c)

$$(v) \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{1}{2} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{3}{2} \end{pmatrix}$$

(vi) Same as (a)

$$9. \quad (i) (1 \ 1 \ 1 \ 0)^t$$

$$(ii) \begin{pmatrix} 1 & 0 & 0 & 1 \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & 0 \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$$

$$(iii) (1 \ \frac{1}{3} \ \frac{1}{3} \ \frac{2}{3})^t$$

$$(iv) \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & -2 & 0 \end{pmatrix}$$

(v) Inverse of each other.

10. $3 + x - 4x^2$

11. (i) $\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$

(ii) $\begin{pmatrix} -1 & 1 \\ -1 & -1 \\ 2 & 0 \end{pmatrix}$

(iii) $\begin{pmatrix} -3 & -3 \\ -1 & -1 \\ 3 & 0 \end{pmatrix}$

12. $\begin{pmatrix} 2 & 3 & 0 \\ -3 & 0 & 4 \end{pmatrix}$

13.

14. $\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 1 & 5 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$

15. $x^2 + x + 6$

16. (i) Yes. By interchanging the 1st and 2nd column of A .(ii) $[T]_{\mathcal{B}_2}$ is obtained by interchanging the i th and j th columns of $[T]_{\mathcal{B}}$

17. (i) $-v_1 + v_2 + 2v_3, v_1, v_2 + 3v_4$

(ii) $\begin{pmatrix} -1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

(iii) $\begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

18. (i) I (ii) I (iii) $P_{\mathcal{B}_1}$ (iv) $P_{\mathcal{B}_1}^{-1}$

19. *Hint:* use $[D]_{\mathcal{B}'} = P^{-1}[D]_{\mathcal{B}}P$

$$[D]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} = [D]_{\mathcal{B}'}$$

20. (i) $\begin{pmatrix} 3 & 0 & 1 \\ -2 & 1 & 1 \\ -1 & 2 & 4 \end{pmatrix}$

(ii) *Hint:* $[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}$

21. (ii) $O_{n \times n}$

(iii) Yes. The null matrix is the matrix of transformation of the zero transformation.

Chapter 17

Eigenvectors and Eigenvalues

Let us consider a transformation of \mathbb{R}^2 , say projection on the x-axis. Under this transformation, every vector along x-axis remains invariant. Similarly under the reflection in the y-axis, every vector along the y-axis remains invariant. Under dilation every non-zero vector is stretched by a factor. Thus, a transformation may move some vectors parallel to themselves, that is, $v \rightarrow \alpha v$ for some scalar α . Such vectors are called eigen vectors and are important for a transformation, and in this chapter we will learn to find them.

17.1 Eigenvectors and Eigenspace

Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be a linear transformation. If $v = (2, 1)^t$, then $Tv = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2v$. Thus $0 \neq v \in \mathbb{R}^2$ is such that $Tv = 2v$, i.e. Tv is a multiple of v . We now define such vectors.

Definition 17.1. (Eigenvector and Eigenvalue):

Let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear operator.

- (i) A non-zero vector v in \mathbb{R}^n is called an eigenvector of T if

$$Tv = \lambda v$$

for some $\lambda \in \mathbb{R}$. The scalar λ is called the eigenvalue of T , associated with v .

- (ii) A scalar λ is an eigenvalue of T if there exists some non-zero $v \in V$ such that $Tv = \lambda v$. The vector v is an eigenvector associated with the eigenvalue λ .

Eigenvalues are also called latent roots, characteristic roots or characteristic values. The eigenvectors are also called latent vectors or characteristic vectors. In the above example, since

$$Tv = 2v$$

2 is an eigenvalue of T associated with the eigenvector v .

Let A be a $n \times n$ matrix. Then

$$T : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \text{defined by}$$

$$TX = AX$$

is a linear operator on \mathbb{R}^n .

Thus, if $v \in \mathbb{R}^n$, then

$$Tv = \lambda v \iff Av = \lambda v$$

Definition 17.2. (Eigenvector and Eigenvalues of a matrix):

Let A be a $n \times n$ matrix.

- (i) A non-zero vector $v \in \mathbb{R}^n$ is called an eigenvector of A if there exists a scalar λ such that

$$Av = \lambda v$$

λ is called an eigenvalue of A . v is called an eigenvector corresponding to the eigenvalue λ .

- (ii) A scalar λ is called an eigenvalue of A if there exists a non-zero vector $v \in V$ such that $Av = \lambda v$ i.e. if $(A - \lambda I)v = 0$ has a non zero solution. Equivalently $|A - \lambda I| = 0$. The vector v is an eigenvector associated with the eigenvalue λ .

Since the eigenvector and eigenvalues of a linear operator are the same as those of the corresponding matrix, and vice versa, therefore it is sufficient to study them for a matrix only. We postpone the method of finding the eigenvalues. We verify whether a given scalar is a eigenvalue and a given vector is an eigenvector.

Example 17.1. Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix}$. Verify that

- (i) $(-1, 1, 2)^t$ is an eigenvector of A . Find the corresponding eigenvalue.

- (ii) 2 is an eigenvalue of A .

$v \in \mathbb{R}^3$ is an eigenvector if

$$Av = \lambda v \text{ for some scalar } \lambda.$$

$$Av = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} = v$$

$$\therefore Av = 1v$$

$\therefore v$ is an eigenvector of A and the corresponding eigenvalue is 1.

- (ii) 2 is an eigenvalue of A if there exists a non-zero vector v such that $Av = 2v$

$$\implies (A - 2I)v = 0 \text{ has a non-zero solution.} \quad (17.1)$$

$$A - 2I = \begin{pmatrix} -1 & 2 & -1 \\ 1 & -2 & 1 \\ 4 & -4 & 3 \end{pmatrix} \quad \text{Reducing it to echelon}$$

form

$$A - 2I \sim \begin{pmatrix} -1 & 2 & -1 \\ 0 & 4 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus there are only 2 pivot columns. Since every column of $A - 2I$ is not a pivot column, therefore (17.1) has a non trivial solution. Hence 2 is an eigenvalue of A .

All solutions are

$$\begin{aligned} v &= \begin{pmatrix} -3x_2 - 4x_3 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} x_3 \end{aligned}$$

Thus, the set of all eigenvectors corresponding to the eigenvalue 2 is

$$\left(\left(\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} x_3 \mid x_2, x_3 \in \mathbb{R}, x_2, x_3 \text{ not both zero} \right) \right)$$

The above example shows that there may be more than one eigenvector corresponding to a given eigenvalue. In fact, we have a general result.

Theorem 17.1. *Let A be a n -rowed square matrix. For each eigenvalue λ of A , the set of all eigenvectors corresponding to λ together with the zero vector is a subspace of \mathbb{R}^n .*

Proof: Let $W_\lambda = \{ 0 \neq v \in \mathbb{R}^n \mid Av = \lambda v \} \cup \{ 0 \}$
 $= \{ v \in \mathbb{R}^n \mid Av = \lambda v \}$
 $\because 0 \in W_\lambda \therefore W_\lambda \neq \phi$
 Let $v_1, v_2 \in W_\lambda, \alpha \in \mathbb{R}$
 $\therefore Av_1 = \lambda v_1$
 $Av_2 = \lambda v_2$
 Then

$$\begin{aligned} A(\alpha v_1 + v_2) &= \alpha Av_1 + Av_2 \\ &= \alpha(\lambda v_1) + \lambda v_2 \\ &= \lambda(\alpha v_1 + v_2) \end{aligned}$$

$$\therefore A(\alpha v_1 + v_2) = \lambda(\alpha v_1 + v_2)$$

so that $(\alpha v_1 + v_2) \in W_\lambda$

Hence W_λ is a subspace of \mathbb{R}^n . □

This theorem gives us the following definition.

Definition 17.3. (Eigenspace): *If λ is an eigenvalue of a matrix A , then the set of all eigenvectors of A corresponding to λ together with the zero vector is called the eigenspace of λ .*

We denote the eigenspace associated with λ by W_λ . Thus $v \in W_\lambda$

$$\iff Av = \lambda v$$

$$\iff (A - \lambda I)v = 0$$

$$\iff v \text{ is a solution of } (A - \lambda I)X = 0$$

Hence the eigenspace of λ is the solution space of $(A - \lambda I)X = 0$ and $\dim W_\lambda =$ number of linearly independent solutions of $(A - \lambda I)X = 0$

$$= \text{nullity } (A - \lambda I)$$

$$\text{Thus } \dim W_\lambda = \text{nullity } (A - \lambda I).$$

We have proved the following result.

Theorem 17.2. Let λ be an eigenvalue of a matrix A . The dimension of the eigenspace associated with λ is the nullity of $A - \lambda I$.

Example 17.2. Consider the linear operator T on \mathbb{R}^3 , which is the projection on the X_1X_2 plane.

The matrix of T is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = A$ (say).

Check that 1 is an eigenvalue of A . The corresponding eigenspace is

$$W_1 = \text{span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

ie. W_1 is the X_1X_2 plane.

We see that under the projection map every vector in the X_1X_2 plane is mapped to itself, and so must be a eigenvector. This is precisely what has been calculated above.

Definition 17.4. (Invariant subspace): Let T be a linear operator on \mathbb{R}^n . A subspace W of \mathbb{R}^n is said to be invariant under T if $Tv \in W$ for all $v \in W$.

Theorem 17.3. If λ is an eigenvalue of a linear operator T , then the eigenspace W_λ is invariant under T .

Proof: Let $v \in W_\lambda$

Then $Tv = \lambda v$

$$\implies Tv \in W_\lambda \quad \because \lambda v \in W_\lambda$$

Hence W_λ is invariant under T . □

Theorem 17.4. The eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof: Let $\lambda_1, \lambda_2, \dots, \lambda_m$ be distinct eigenvalues of a $n \times n$ matrix A and v_1, v_2, \dots, v_m be the corresponding eigenvectors.

Since v_i is an eigenvector corresponding to the eigenvalue λ_i , therefore $Av_i = \lambda_i v_i$, and $v_i \neq 0$, $i = 1, 2, \dots, m$. If possible, let $\{v_1, v_2, \dots, v_m\}$

be linearly dependent. $\because v_1 \neq 0, \therefore$ some vectors v_i is linear combination of the preceding vectors.

Let k be the least index such that v_{k+1} is the linear combination of the preceding v_i 's. Consequently $\{v_1, v_2, \dots, v_k\}$ is linearly independent. Hence there are scalars $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k$ such that

$$v_{k+1} = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k \tag{17.2}$$

Multiplying both sides by A , we get

$$\begin{aligned}
 Av_{k+1} &= \alpha_1 Av_1 + \alpha_2 Av_2 + \dots + \alpha_k Av_k \\
 \implies \lambda_{k+1} v_{k+1} &= \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_k \lambda_k v_k & (17.3) \\
 & (17.2) \times \lambda_{k+1} - (17.3) \\
 \implies 0 &= \alpha_1 (\lambda_{k+1} - \lambda_1) v_1 + \dots + \alpha_k (\lambda_{k+1} - \lambda_k) v_k \\
 \implies \alpha_i (\lambda_{k+1} - \lambda_i) &= 0, \quad i = 1, 2, \dots, k \text{ as } (v_1, v_2, \dots, v_k) \text{ is linearly independent.} \\
 \implies \alpha_i &= 0, \quad i = 1, 2, \dots, k \text{ as } \lambda_i \text{'s are distinct} \\
 \implies v_{k+1} &= 0 \text{ using (17.2)}
 \end{aligned}$$

which contradicts the fact that $v_{k+1} \neq 0$

Hence our assumption is wrong, so that v_1, v_2, \dots, v_m are linearly independent. \square

17.2 Solved Problems

Problem 17.1. If $A = \begin{pmatrix} -8 & -9 & -12 \\ 2 & 1 & 4 \\ 2 & 3 & 2 \end{pmatrix}$, find the eigenspace of A associated with eigenvalue -2 . Also find a basis for the eigen space.

Solution: If v is an eigenvector of A associated with the eigenvalue -2 then

$$\begin{aligned}
 Av &= -2v \\
 \implies (A + 2I)v &= 0 & (17.4)
 \end{aligned}$$

Reducing to echelon form

$$\begin{aligned}
 A + 2I &\sim \begin{pmatrix} 2 & 3 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ so system (17.4) gives} \\
 v &= \begin{pmatrix} -\frac{3}{2}k_2 - 2k_1 \\ k_2 \\ k_1 \end{pmatrix}, \text{ where } k_1, k_2 \in \mathbb{R} \\
 &= \begin{pmatrix} -\frac{3}{2} \\ 1 \\ 0 \end{pmatrix} k_2 + \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} k_1, \quad k_1, k_2 \in \mathbb{R}
 \end{aligned}$$

The eigenspace associated with -2 is $\left\{ \begin{pmatrix} -\frac{3}{2} \\ 1 \\ 0 \end{pmatrix} k_2 + \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} k_1, |k_1, k_2 \in \mathbb{R} \right\}$
 $(-\frac{3}{2}, 1, 0)^t, (-2, 0, 1)^t$ form a basis for the eigenspace.

Problem 17.2. Prove that a matrix A is a scalar matrix if and only if every non-zero vector is an eigenvector of A .

Solution: Let $A = \begin{pmatrix} d & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & d \end{pmatrix}$ be a scalar matrix.

Let $X = (x_1, x_2, \dots, x_n)^t$ be any non-zero vector. Then

$$AX = (dx_1, dx_2, \dots, dx_n)^t = d(x_1, x_2, \dots, x_n)^t$$

$\therefore AX = dX$ so that X is an eigenvector of A .

Conversely, let every non-zero vector be an eigenvector. If e_1, e_2, \dots, e_n are the columns of the $n \times n$ unit matrix, then each e_i being non-zero, is an eigenvector. Let λ_i be the eigenvalue corresponding to the eigenvector e_i . Thus

$$Ae_i = \lambda_i e_i, \quad i=1,2,\dots,n.$$

so that

$$A(e_1 + \dots + e_n) = \lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n \quad (17.5)$$

Since, $e_1 + e_2 + \dots + e_n$ is a non-zero vector, therefore it is an eigenvector. Thus there exists μ such that

$$A(e_1 + e_2 + \dots + e_n) = \mu(e_1 + e_2 + \dots + e_n)$$

Using (17.5), we get

$$\lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n = \mu(e_1 + e_2 + \dots + e_n)$$

$$\Rightarrow (\lambda_1 - \mu)e_1 + \dots + (\lambda_n - \mu)e_n = 0$$

$$\Rightarrow \lambda_1 - \mu = \dots = \lambda_n - \mu = 0, \text{ as the } e_i\text{'s are linearly independent.}$$

$$\therefore \lambda_1 = \dots = \lambda_n = \mu$$

Hence, $Ae_i = \mu e_i, 1 \leq i \leq n$

$$\text{so that } A = \begin{pmatrix} \mu & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \mu \end{pmatrix}$$

17.3 Exercise

- For the following matrices A determine whether v is an eigenvector of A . If yes, find the corresponding eigenvalue.

$$(i) \quad A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$(ii) \quad A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix}, \quad v = \begin{pmatrix} -2 \\ -3 \\ 2 \end{pmatrix}$$

$$(iii) A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix}, \quad v = \begin{pmatrix} 8 \\ -5 \\ 2 \end{pmatrix}$$

2. For the following matrices A determine whether the given scalar λ is an eigenvalue of A .

$$(i) A = \begin{pmatrix} 3 & -2 & 2 \\ 0 & 3 & 0 \\ 0 & -2 & 5 \end{pmatrix}, \quad \lambda = 3$$

$$(ii) A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda = -1$$

$$(iii) A = \begin{pmatrix} 4 & 3 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 5 & 2 \end{pmatrix}, \quad \lambda = 2$$

3. Find a basis for the eigen space corresponding to the listed eigenvalue.

$$(i) A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}, \quad \lambda = 1$$

$$(ii) A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}, \quad \lambda = 2$$

$$(iii) A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \lambda = 2$$

4. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear operator defined below. Find eigenvalues of T and the associated eigen spaces without calculations.

- (i) Projection on x_1 axis.
- (ii) Reflection in x_2 axis.
- (iii) Reflection in origin.
- (iv) Rotation about origin.
- (v) Reflection in the line $x_1 = x_2$.

5. Justify the following statements:

- (i) A square A matrix is singular if and only if 0 is an eigenvalue of A .
- (ii) A steady state vector of a stochastic matrix is an eigenvector.

6. If A is any square matrix such that $A^2 = 0$. Show that '0' is the only eigenvalue of A .
7. If A is a square matrix, such that each column of A adds up to the same number say λ . show that λ is a characteristic value of A .
8. If A is a diagonal matrix with distinct entries then prove that each diagonal entry is (i) an eigenvalue, (ii) the eigenspace corresponding to each eigenvalue is of dimension 1.
9. If A is a square matrix such that the sum of the elements of each row of A is k , prove that k is a characteristic root of A .

17.4 Characteristic Equation

The eigenvectors of a matrix A are the non-zero solutions of $(A - \lambda I)X = 0$, where λ is an eigenvalue, \therefore the question arises, "How do we find the eigenvalues of a matrix?"

A scalar λ is an eigenvalue of a n -rowed square matrix A

\iff there exists some non-zero vector v , such that $Av = \lambda v$

$\iff (A - \lambda I)v = 0$

$\iff (A - \lambda I)X = 0$ has a non-zero solution.

\iff column rank $(A - \lambda I) < n$

\iff row rank $(A - \lambda I) < n$, since column rank and row rank are equal.

$\iff A - \lambda I$ is row equivalent to a matrix with a row of zeros.

$\iff |A - \lambda I| = 0$, since $|A| = \pm |B|$, if A and B are row equivalent matrices.

$\therefore |A - \lambda I|$ is a polynomial of degree n , in λ (the coefficient of λ^n is $(-1)^n$) $\therefore \lambda$ is a zero of a polynomial of degree n , namely $|A - \lambda I|$

Hence, the eigenvalues of a matrix A are the roots of the equation $|A - \lambda I| = 0$

Thus any $n \times n$ matrix has n eigenvalues in \mathbb{C} .

Definition 17.5. Let A be a n -rowed square matrix over \mathbb{R} . Then,

- (i) $|A - \lambda I|$ is called the characteristic polynomial of A .
- (ii) The equation $|A - \lambda I| = 0$ is called the characteristic equation of A .
- (iii) The roots of the characteristic equation are called the characteristic roots or eigenvalues or latent roots of A .

Since, the matrix is considered over \mathbb{R} , therefore the eigenvalues must also belong to \mathbb{R} , so we restrict ourselves to real eigenvalues only, though complex eigenvalues are also possible. The study of complex eigenvalues is beyond the scope of this book.

Problem 17.3. Let us find the characteristic polynomial, characteristic equation and characteristic roots of the matrix A , where

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix}$$

Solution: Characteristic polynomial of A is

$$|A - \lambda I| = \begin{vmatrix} 1 - \lambda & 2 & -1 \\ 1 & -\lambda & 1 \\ 4 & -4 & 5 - \lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - 11\lambda + 6$$

Characteristic equation of A is

$$|A - \lambda I| = 0 \text{ i.e. } \lambda^3 + 6\lambda^2 - 11\lambda + 6 = 0 \quad (17.6)$$

Characteristic roots of A are the roots of (17.6). They are $\lambda = 1, 2, 3$.

Example 17.3. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

Characteristic equation of A is

$$\begin{aligned} & |A - \lambda I| = 0 \\ \Rightarrow & \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = 0 \\ \Rightarrow & \lambda^2 + 1 = 0 \\ & \lambda = \pm i \end{aligned}$$

Since there are no real roots, hence there are no eigenvalues of A .

The characteristic roots of a matrix in special form can be easily found.

Theorem 17.5. The characteristic roots of a triangular matrix are its diagonal elements.

Proof: Let A be an upper triangular matrix. Then

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ 0 & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ 0 & 0 & a_{33} & \cdot & \cdot & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & a_{nn} \end{pmatrix}$$

Characteristic equation of A is

$$\Rightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ 0 & a_{22} - \lambda & \cdot & \cdot & \cdot & a_{2n} \\ 0 & 0 & a_{33} - \lambda & \cdot & \cdot & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & a_{nn} - \lambda \end{vmatrix} = 0 \quad |A - \lambda I| = 0$$

$$\implies (a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda) = 0$$

$$\implies \lambda = a_{11}, a_{22}, \dots, a_{nn}$$

Hence characteristic roots of an upper triangular matrix are its diagonal elements.

Similarly, if A is a lower triangular matrix, then characteristic roots are the diagonal elements. Hence the characteristic roots of a triangular matrix are its diagonal elements.

Corollary 17.6. *The characteristic roots of a diagonal matrix are its diagonal elements.*

Corollary 17.7. *The characteristic roots of a scalar matrix are its diagonal elements.*

The following theorem gives relationship of the eigenvalues and eigenvectors of A and some of its related matrix.

Theorem 17.8. *Let A be an n -rowed square matrix. Then*

- (i) A and A^t have the same characteristic roots.
- (ii) If k is any scalar, then characteristic roots of kA are k times the characteristic roots of A . Also the corresponding eigenvectors are the same.
- (iii) For any positive integer p the characteristic roots of A^p are the p^{th} power of the characteristic roots of A . Moreover the corresponding eigenvectors are the same.

Proof: Left to the readers. □

Theorem 17.9. *Similar matrices have the same characteristic roots.*

Proof: Let A and B be similar matrices. Then there exists an invertible matrix P such that

$$B = P^{-1}AP$$

Then

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda I| \\ &= |P^{-1}AP - \lambda I| \\ &= |P^{-1}AP - P^{-1}P\lambda I| \\ &= |P^{-1}AP - P^{-1}\lambda IP| \\ &= |P^{-1}(A - \lambda I)P| \\ &= |P^{-1}||A - \lambda I||P| \quad \because \det(AB) = (\det A)(\det B) \\ &= |P^{-1}||P||A - \lambda I| \quad \because \det(P^{-1}) = (\det P)^{-1} \\ &= |A - \lambda I| \end{aligned}$$

$$\therefore |B - \lambda I| = |A - \lambda I|$$

$$\text{so that } |B - \lambda I| = 0$$

$$\iff |A - \lambda I| = 0$$

Thus B and A have the same characteristic equation and therefore the same characteristic roots. □

Theorem 17.10. *If A is any square matrices then*

- (i) *Sum of the characteristic roots = trace A .*
- (ii) *Product of characteristic roots = $|A|$.*

Proof: Let $A = (a_{ij})_{n \times m}$. Then

$$\begin{aligned}
 |(A - \lambda I)| &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} - \lambda \end{vmatrix} \\
 &= (-1)^n \lambda^n + a_1 \lambda^{n-1} + \dots + a_n
 \end{aligned} \tag{17.7}$$

Characteristic equation of A is

$$\begin{aligned}
 |(A - \lambda I)| &= 0 \\
 \Rightarrow (-1)^n \lambda^n + a_1 \lambda^{n-1} + \dots + a_n &= 0
 \end{aligned} \tag{17.8}$$

If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the characteristic roots of A , then

$$\lambda_1 + \lambda_2 + \dots + \lambda_n = (-1)^{n-1} a_1 \tag{17.9}$$

$$\lambda_1 \lambda_2 \dots \lambda_n = (-1)^n a_n = a_n \tag{17.10}$$

a_1 is the coefficient of λ^{n-1} in (17.7).

In the expansion of $|(A - \lambda I)|$, λ^{n-1} occurs only in the term $(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda)$ and its coefficient is $(-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})$.

$$\therefore a_1 = (-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn}) = (-1)^{n-1} \text{ trace } A \tag{17.11}$$

(17.10) and (17.11) $\Rightarrow \lambda_1 \lambda_2 \dots \lambda_n = |A|$ Also putting $\lambda = 0$ in (17.7), we get

$$|A| = a_n \tag{17.12}$$

(17.9) and (17.11) $\Rightarrow \lambda_1 + \lambda_2 + \dots + \lambda_n = \text{trace } A$ □

Problem 17.4. *Find the eigenvalues and the corresponding eigenvector of the*

linear operator T on \mathbb{R}^3 whose matrix is $\begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix}$

Solution: Let $A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix}$

The characteristic equation of A is $|A - \lambda I| = 0$

$$\begin{aligned}
 \Rightarrow & \begin{vmatrix} 2 - \lambda & 2 & 3 \\ 1 & 2 - \lambda & 1 \\ 2 & -2 & 1 - \lambda \end{vmatrix} = 0 \\
 \Rightarrow & \lambda^3 - 5\lambda^2 + 2\lambda + 8 = 0 \\
 \Rightarrow & (\lambda - 2)(\lambda + 1)(\lambda - 4) = 0 \\
 \Rightarrow & \lambda = 2, -1, 4
 \end{aligned}$$

Thus there are 3 distinct eigenvalues. We find the eigenvectors corresponding to them. If v is an eigenvector corresponding to the eigenvalue λ then

$$\begin{aligned} (A - \lambda I)v &= 0 \\ \text{Thus } v \text{ is a solution of} & \\ (A - \lambda I)X &= 0 \end{aligned} \quad (17.13)$$

$$\text{i.e.} \quad \begin{pmatrix} 2 - \lambda & 2 & 3 \\ 1 & 2 - \lambda & 1 \\ 2 & -2 & 1 - \lambda \end{pmatrix} X = 0 \quad (17.14)$$

Characteristic vector corresponding to $\lambda = -1$.

Putting $\lambda = -1$ in (17.14) we get

$$\begin{pmatrix} 3 & 2 & 3 \\ 1 & 3 & 1 \\ 2 & -2 & 2 \end{pmatrix} X = 0 \quad (17.15)$$

Reducing the coefficient matrix to echelon form, we get

$$\begin{pmatrix} 3 & 2 & 3 \\ 1 & 3 & 1 \\ 2 & -2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

\therefore there are only 2 pivot columns, \therefore (17.15) has non-trivial solutions, given by

$$x_1 = -k.$$

$$x_2 = 0 \quad \text{where } k \text{ is any real number.}$$

$$x_3 = k.$$

$$\therefore X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -k \\ 0 \\ k \end{pmatrix} = k \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$\therefore \text{Eigenvectors corresponding to } \lambda = -1 \text{ are } \left(k \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid k \in \mathbb{R}^* \right)$$

An eigenvector corresponding to $\lambda = 1$ is $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, obtained by taking $k = 1$.

Characteristic vector corresponding to $\lambda = 2$.

Putting $\lambda = 2$ in (17.14)

$$\Rightarrow \quad \begin{pmatrix} 0 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & -2 & -1 \end{pmatrix} X = 0 \quad (17.16)$$

The echelon form of the coefficient matrix is

$$\begin{pmatrix} \mathbf{1} & 0 & 1 \\ 0 & \mathbf{2} & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

By the same argument as before (17.16) has non-trivial solutions given by

$$\begin{aligned}x_1 &= -k \\x_2 &= -\frac{3}{2}k \\x_3 &= k\end{aligned}$$

where k is any real number. General solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} -1 \\ \frac{3}{2} \\ 1 \end{pmatrix} = k' \begin{pmatrix} -2 \\ 3 \\ 2 \end{pmatrix}, \quad k' \in \mathbb{R}$$

\therefore Eigenvectors corresponding to $\lambda = 2$ are $\left(k \begin{pmatrix} -2 \\ 3 \\ 2 \end{pmatrix} \mid k \in \mathbb{R}^* \right)$

An eigenvector corresponding to $\lambda = 2$ is $\begin{pmatrix} -2 \\ 3 \\ 2 \end{pmatrix}$, obtained by taking $k = 1$.

Characteristic vector corresponding to $\lambda = 4$.

Putting $\lambda = 4$ in (17.14)

$$\implies \begin{pmatrix} -2 & 2 & 3 \\ 1 & -2 & 1 \\ 2 & -2 & -3 \end{pmatrix} X = 0 \quad (17.17)$$

The echelon form of the coefficient matrix is

$$\begin{pmatrix} \mathbf{1} & -2 & 1 \\ 0 & \mathbf{-2} & 5 \\ 0 & 0 & 0 \end{pmatrix}$$

By the same argument as before (17.17) has non-trivial solutions given by

$$\begin{aligned}x_1 &= 8k \\x_2 &= 5k \\x_3 &= 2k\end{aligned}$$

General solution is

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} 8 \\ 5 \\ 2 \end{pmatrix}$$

\therefore Eigenvectors corresponding to $\lambda = 4$ are

$$\left(k \begin{pmatrix} 8 \\ 5 \\ 2 \end{pmatrix} \mid k \in \mathbb{R}^* \right)$$

An eigenvector corresponding to $\lambda = 4$ is $\begin{pmatrix} 8 \\ 5 \\ 2 \end{pmatrix}$, obtained by taking $k = 1$.

Thus the eigenvector corresponding to $\lambda = -1, 2, 4$ are $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 5 \\ 2 \end{pmatrix}$ respectively.

We know that the eigenvectors corresponding to distinct eigenvalue are linearly independent. In the above problem, if

$$\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = 4 \text{ and}$$

$$v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -2 \\ 3 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 8 \\ 5 \\ 2 \end{pmatrix}$$

$$\text{Let } P = (v_1 \ v_2 \ v_3), D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

$$\text{Then } AP = \begin{pmatrix} 1 & -4 & 32 \\ 0 & 6 & 20 \\ -1 & 4 & 8 \end{pmatrix}$$

$$PD = \begin{pmatrix} 1 & -4 & 32 \\ 0 & 6 & 20 \\ -1 & 4 & 8 \end{pmatrix}$$

so that $AP = PD$.

\therefore the columns of P are linearly independent, $\therefore P$ is invertible. Hence $AP = PD$

$$\implies P^{-1}AP = D$$

This will help us in solving many problems, as D is a diagonal matrix.

Problem 17.5. Let $V = \mathbb{P}_3$, the vector space of all polynomials of degree less than or equal to 3.

Let $D: \mathbb{P}_3 \rightarrow \mathbb{P}_3$ be defined by $D(f(x)) = \frac{df(x)}{dx}$

Find the matrix A of D relative to the standard basis.

Find the eigenvalues and the associated eigenspace of A .

Solution: A basis of \mathbb{P}_3 is $\{1, x, x^2, x^3\} = \{v_1, v_2, v_3, v_4\}$ (say)

$$Dv_1 = 0 = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^t$$

$$Dv_2 = 0 = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}^t$$

$$Dv_3 = 0 = \begin{pmatrix} 0 & 2 & 0 & 0 \end{pmatrix}^t$$

$$Dv_4 = 0 = \begin{pmatrix} 0 & 0 & 3 & 0 \end{pmatrix}^t$$

$$\text{Hence } A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ is the matrix of } D \text{ relative to the standard basis.}$$

Eigenvalues of A

The eigenvalues of A are the roots of the equation $|A - \lambda I| = 0$

$$\implies \begin{vmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 2 & 0 \\ 0 & 0 & -\lambda & 3 \\ 0 & 0 & 0 & -\lambda \end{vmatrix} = 0$$

$$\implies \lambda^4 = 0$$

$$\implies \lambda = 0, 0, 0, 0$$

The eigenspace corresponding to the eigenvalue $\lambda = 0$ is the solution set of

$$(A - 0 I)X = 0 \\ \implies AX = 0$$

A is in echelon form, it is

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & 0 & \mathbf{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

If $X = (x_1 \ x_2 \ x_3 \ x_4)^t$ then the above system gives
 $x_2 = x_3 = x_4 = 0$
 x_1 is arbitrary.

Hence eigenspace corresponding to $\lambda = 0$ is $\left(x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \mid x_1 \in \mathbb{R} \right)$

and a basis of the eigenspace is $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$

Problem 17.6. If A is any non-singular matrix and if λ is an eigenvalue of A , then

- (i) $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} .
- (ii) $\frac{|A|}{\lambda}$ is an eigenvalue of $\text{adj}A$.

Solution: Since A is non-singular

$\therefore |A| \neq 0$, and so every eigenvalue of A is non-zero. Let λ be an eigenvalue of A .

$\therefore \lambda \neq 0$.

Let $0 \neq X$ be eigen vector corresponding to the eigenvalue λ . Then

$$AX = \lambda X \tag{17.18}$$

- (i) Premultiplying (17.18) by A^{-1} ,
 $A^{-1}AX = \lambda A^{-1}X \implies A^{-1}X = \frac{1}{\lambda}X$
 $\implies \frac{1}{\lambda}$ is a characteristic root of A^{-1}

- (ii) Premultiplying (17.18) by $\text{adj}A$, we get
 $(\text{adj}A)AX = \lambda \text{adj}(A)X$
 $\implies |A| IX = \lambda(\text{adj}A)X$
 $\implies (\text{adj}A)X = \frac{|A|}{\lambda} X$
 $\implies \frac{|A|}{\lambda}$ is a characteristic root of $\text{adj}A$.

Problem 17.7. In a town the usage of land in the year 2005 was 30% residential, 20% commercial and 50% industrial. If x_m, y_m, z_m denotes the percentage of residential, commercial and industrial usage respectively after m years, and $X_{m+1} = PX_m$

where $X_m = (x_m \ y_m \ z_m)^t$, and

$$P = \begin{pmatrix} 0.8 & 0.1 & 0.0 \\ 0.1 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.9 \end{pmatrix}$$

Find the land used in town after 40 years.

Solution: Here the initial vector

$$X_0 = \begin{pmatrix} 0.3 \\ 0.2 \\ 0.5 \end{pmatrix}$$

We first find the eigenvalues and the eigenvectors of the matrix P.

Characteristic equation of P is

$$\begin{aligned} |P - \lambda I| &= 0 \\ \implies \lambda^3 - 2.4\lambda^2 + 1.88\lambda - 0.480 &= 0 \\ \implies (\lambda - 0.6)(\lambda - 0.8)(\lambda - 1.0) &= 0 \\ \implies \lambda &= 0.6, 0.8, 1. \end{aligned}$$

Let $\lambda_1 = 0.6$, $\lambda_2 = 0.8$, $\lambda_3 = 1$.

If v is a characteristic vector corresponding to λ , then we have to solve

$$(A - \lambda I)v = 0$$

$$\text{For } \lambda = \lambda_1 = 0.6, \text{ we get } v_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\text{For } \lambda = \lambda_2 = 0.8, \text{ we get } v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$\text{For } \lambda = \lambda_3 = 1, \text{ we get } v_3 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$$

\therefore the eigen vectors v_1, v_2, v_3 correspond to distinct eigenvalues, \therefore they must be linearly independent.

We express the initial vector X_0 in terms of v_1, v_2, v_3 . We can see that

$$X_0 = c_1 v_1 + c_2 v_2 + c_3 v_3 \\ \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 2 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.2 \\ 0.5 \end{pmatrix}$$

Solving, we get,

$$c_1 = 0.025, c_2 = 0.15, c_3 = 0.125$$

$$\begin{aligned} X_1 &= PX_0 = P(c_1 v_1 + c_2 v_2 + c_3 v_3) \\ &= c_1 P v_1 + c_2 P v_2 + c_3 P v_3 \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + c_3 \lambda_3 v_3 \\ X_2 &= P X_1 = P(c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + c_3 \lambda_3 v_3) \\ &= c_1 \lambda_1 P v_1 + c_2 \lambda_2 P v_2 + c_3 \lambda_3 P v_3 \\ &= c_1 \lambda_1^2 v_1 + c_2 \lambda_2^2 v_2 + c_3 \lambda_3^2 v_3 \end{aligned}$$

In general $X_k = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + c_3 \lambda_3^k v_3$

Land usage after 40 years is X_{40} and is given by

$$\begin{aligned} X_{40} &= c_1 \lambda_1^{40} v_1 + c_2 \lambda_2^{40} v_2 + c_3 \lambda_3^{40} v_3 \\ &= c_1 (0.6)^{40} v_1 + c_2 (0.8)^{40} v_2 + c_3 (1)^{40} v_3 \\ &= c_3 v_3 \quad \because (0.60)^{40} \approx 0, (0.8)^{40} \approx 0 \end{aligned}$$

$$\begin{aligned}
 &= 0.125 \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \\
 &= \begin{pmatrix} 0.125 \\ 0.250 \\ 0.625 \end{pmatrix}
 \end{aligned}$$

\therefore The land use in the town after 40 years is 12.5% residential, 25% commercial and 62.5% industrial.

17.5 Exercise

1. For the following matrices find the eigenvalues without calculation.

$$\begin{aligned}
 \text{(i)} & \begin{pmatrix} -1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{(iii)} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
 \text{(iv)} & \begin{pmatrix} 3 & 0 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 6 & 0 & -1 & 0 \end{pmatrix} \quad \text{(v)} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 4 & 3 & -4 & -6 \end{pmatrix}
 \end{aligned}$$

2. Find one eigenvalue of the following matrices without calculation.

$$\begin{aligned}
 \text{(i)} & \begin{pmatrix} 1 & -2 & 4 \\ 1 & -2 & 4 \\ 2 & -4 & 8 \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{pmatrix} \quad \text{(iii)} \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \\
 \text{(iv)} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 6 & 0 \\ 3 & 1 & 9 & 4 \\ -1 & 1 & 2 & 0 \end{pmatrix}
 \end{aligned}$$

3. Find the eigenvalues of the matrix A

$$\begin{aligned}
 \text{(i)} & A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \\
 \text{(ii)} & A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \\
 \text{(iii)} & A = \begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix} \\
 \text{(iv)} & A = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix} \\
 \text{(v)} & A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}
 \end{aligned}$$

4. Find the eigenvalues of the following matrices

$$\text{(i)} \begin{pmatrix} 3 & 5 \\ 1 & 4 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 0 & -1 & 2 \\ -1 & 0 & 2 \\ 2 & 2 & 3 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix}$$

5. If the characteristic equation of a non-singular 3×3 matrix A is $\lambda^3 - p\lambda^2 + q\lambda - r = 0$ prove that the characteristic equation of $\text{adj}A$ is $\lambda^3 - q\lambda^2 + rp\lambda - r^2 = 0$
6. Prove that A and A^t have the same eigenvalues. Do they necessarily have the same eigenvectors?
7. Prove that the eigenvalues of kA are k times the eigenvalues of A . Also prove that the corresponding eigenvectors are the same.
8. Prove that the eigenvalues of A^p are the p^{th} powers of the eigenvalues of A , where p is a positive integer and that the corresponding eigen vectors are the same.
9. Prove that a matrix is singular if and only if 0 is an eigenvalue.
10. If A and B are n -rowed square matrices and if A is invertible, show that $A^{-1}B$ and BA^{-1} have the same eigenvalues.
11. If A and B are the n -rowed square matrices and λ and μ are eigen values of A and B corresponding to an eigenvector X , prove that
- $\lambda + \mu$ is an eigenvalue of $A+B$.
 - $\lambda\mu$ is an eigenvalue of AB .
- Is $\lambda\mu$ an eigenvalue of BA also?
12. For $A = \begin{pmatrix} 2 & 1 & 0 \\ 9 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ find the eigenvalues and the corresponding eigenvectors. Find the eigen values and eigen vectors of A^t and hence verify Q 6.
13. For $A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix}$ find the eigenvalues and the eigenvectors. Also find the eigenvalues and eigenvectors of $-2A$ and $3A$. Hence verify Q.7.
14. For $A = \begin{pmatrix} 3 & -1 \\ -2 & 2 \end{pmatrix}$ find the eigenvalues and eigenvectors of A^2 and A^3 . Hence verify Q.8.
15. Prove that every square matrix of odd order has atleast one real eigenvalue.

16. If $A = \begin{pmatrix} 12 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$, prove that (i) each diagonal entry is an eigen-

value (ii) the eigenspace associated with the eigenvalue 12 is 2-dimensional, whereas the eigen spaces associated with 1 and 3 are each 1-dimensional.

17. Prove that the eigenspace of a $n \times n$ scalar matrix is \mathbb{R}^n .

18. For the following matrix find the characteristic polynomial, eigenvalues and the corresponding eigenspaces.

$$\begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$$

19. Let $V = P_1$, the vector space of polynomials of degree 1 or less. Let

$D: P_1 \rightarrow P_1$ defined by

$$D(f(x)) = \frac{d(f(x))}{dx}$$

Find the matrix A of D relative to the standard basis and the associated eigen-space.

Repeat the above problem for P_2 .

17.6 Diagonalization

In this section we study a very useful application of eigenvalues and eigenvectors.

In physics, chemistry, engineering, business etc. we come across situations where we need to calculate high power of a given matrix A . Such calculations are very time consuming and prone to errors. For instance, if $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ and we want to calculate A^6 , then

$$\begin{aligned} A^6 &= A A A A A A \\ &= \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 5 \\ -10 & 14 \end{pmatrix} \begin{pmatrix} -1 & 5 \\ -10 & 14 \end{pmatrix} \begin{pmatrix} -1 & 5 \\ -10 & 14 \end{pmatrix} \\ &= \begin{pmatrix} -49 & 65 \\ -130 & 146 \end{pmatrix} \begin{pmatrix} -1 & 5 \\ -10 & 14 \end{pmatrix} \\ &= \begin{pmatrix} -601 & 665 \\ -1330 & 1394 \end{pmatrix} \end{aligned}$$

We know that, there is a diagonal matrix $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ and an invertible matrix $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ such that $A = PDP^{-1}$

Then

$$\begin{aligned}
 A^6 &= (PDP^{-1})(PDP^{-1})(PDP^{-1})(PDP^{-1})(PDP^{-1})(PDP^{-1}) \\
 &= PD^6P^{-1} \\
 &= P \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^6 P^{-1} \\
 &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2^6 & 0 \\ 0 & 3^6 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 64 & 729 \\ 64 & 1458 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} -601 & 665 \\ -1330 & 1394 \end{pmatrix}
 \end{aligned}$$

The calculations have been simplified because $A = PDP^{-1}$
 $\implies A^6 = PD^6P^{-1}$

It is easy to find D^6 , as D is a diagonal matrix.

Thus we observe that if any matrix A can be expressed in the form PDP^{-1} , for some diagonal matrix D , then it is very easy to calculate any power of A . In this case we say that A is a diagonalizable matrix. Thus we have the following definition.

Definition 17.6. (Diagonalizable Matrix): A square matrix A is said to be diagonalizable if A is similar to a diagonal matrix, i.e. there exist an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

The following theorem gives a characterization of diagonalizable matrices. It also gives a way to construct the matrix P .

Theorem 17.11. An n -rowed square matrix A is diagonalizable if and only if A has n linearly-independent eigenvectors.

In this case $D = P^{-1}AP$, where D is a diagonal matrix whose diagonal elements are the eigenvalues of the A , and P is a matrix whose columns are respectively the n linearly independent eigenvectors of A .

Proof: Let A be an n -rowed diagonalizable matrix. Then there exists a diagonal matrix D and an invertible matrix P such that $A = PDP^{-1}$

$$\therefore AP = PD \tag{17.19}$$

$$\text{Let } D = \begin{pmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ \cdot & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}, P = (v_1 \ v_2 \ \cdot \ \cdot \ v_n)$$

$$\text{Then } AP = (Av_1 \ Av_2 \ \cdot \ \cdot \ Av_n) \text{ and } PD = P \begin{pmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ \cdot & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

$$= (\lambda_1 v_1 \ \lambda_2 v_2 \ \cdot \ \cdot \ \lambda_n v_n)$$

$$\text{From (17.19) we get } (Av_1 \ Av_2 \ \cdot \ \cdot \ Av_n) = (\lambda_1 v_1 \ \lambda_2 v_2 \ \cdot \ \cdot \ \lambda_n v_n)$$

$$\implies Av_i = \lambda_i v_i, i = 1, 2, \dots, n. \tag{17.20}$$

Since P is non-singular, \therefore the columns of P are non-zero and are linearly independent.

Thus from (17.20) and above we conclude that v_1, v_2, \dots, v_n are linearly independent. Thus A has n linearly-independent eigenvectors.

Conversely, suppose A has n linearly-independent eigenvectors v_1, v_2, \dots, v_n . Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues. Then $Av_1 = \lambda_1 v_1, Av_2 = \lambda_2 v_2, \dots, Av_n = \lambda_n v_n$, so that

$$(Av_1 \quad Av_2 \quad \dots \quad Av_n) = (\lambda_1 v_1 \quad \lambda_2 v_2 \quad \dots \quad \lambda_n v_n) \quad (17.21)$$

$$\text{Let } P = (v_1 \quad v_2 \quad \dots \quad v_n), \quad D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ \cdot & \lambda_2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

$\therefore v_i$'s are linearly independent, $\therefore P$ is non-singular.

(17.21) $\implies AP = PD$, so that $A = PDP^{-1}$ as P is invertible.

Hence A is diagonalizable. \square

Problem 17.8. Is $\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ is diagonalizable.

Solution: Let $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$

Characteristic equation of A is

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & -1-\lambda \end{vmatrix} \\ \implies (\lambda+1)(\lambda-1) = 0 \\ \implies \lambda = -1, 1$$

Eigen values are $-1, 1$

Let us now find the eigenvectors. Let $v = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^t$ is a eigenvector corresponding to the eigenvalue λ . Then $(A - \lambda I)v = 0$

$$\implies \begin{pmatrix} 1-\lambda & 1 \\ 0 & -1-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad (17.22)$$

Eigenvector corresponding to $\lambda = 1$

For $\lambda = 1$, equation (17.22) becomes

$$\begin{pmatrix} 0 & 1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad (17.23)$$

Echelon form of the coefficient matrix is

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

But (17.23) $\Rightarrow x_2 = 0$

\therefore Eigenspace corresponding to $\lambda = 1$ is

$$\left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix} : x_1 \in \mathbb{R} \right\} = \left\{ x \begin{pmatrix} 1 \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$$

Hence $\{v_1\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ is a basis of this eigenspace.

Similarly eigenspace corresponding to $\lambda = -1$ is

$$\left\{ \begin{pmatrix} x_1 \\ -2x_1 \end{pmatrix} : x_1 \in \mathbb{R} \right\} = \left\{ x_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} : x_1 \in \mathbb{R} \right\}$$

Thus $\{v_2\} = \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\}$ is a basis of this eigenspace.

\therefore the eigenvectors corresponding to distinct eigenvalues are linearly independent,

$\therefore A$ has 2 linearly independent vectors v_1 and v_2

Hence A is diagonalizable, with

$P = (v_1 \ v_2) = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$, and $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, such that $P^{-1}AP = D$.

Steps involved in diagonalizing a matrix

Step 1 Find the characteristic roots of A which are the roots of

$$|A - \lambda I| = 0$$

Let them be $\lambda_1, \lambda_2, \dots, \lambda_n$

Step 2 If all the roots are real then proceed to Step 3.

If not, then A is not diagonalizable.

Step 3 For each eigenvalue λ_i , find a basis for the eigenspace. This basis is the set of linearly independent eigenvectors associated with λ_i .

If the number of all linearly independent eigenvectors corresponding to the eigenvalue λ_i , is n then go to step 4, else A is not diagonalizable.

Step 4 Let v_1, v_2, \dots, v_n be the n linearly independent eigenvectors, corresponding to the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (in this order) obtained in Step 3.

$$\text{Let } P = (v_1 \ v_2 \ \dots \ v_n) \ , D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ \cdot & \lambda_2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

Then $A = PDP^{-1}$ or $P^{-1}AP = D$ is the required diagonalization of A .

Problem 17.9. (Case of repeated eigenvalues) Is the matrix

$$A = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix} \text{ diagonalizable?}$$

Solution: Let $A = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix}$

Step 1 Characteristic equation of A is

$$|A - \lambda I| = 0$$

$$\implies \lambda^3 + 3\lambda^2 - 4 = 0$$

$$\implies (\lambda - 1)(\lambda + 2)^2 = 0$$

$$\implies \lambda = 1, -2, -2.$$

One eigen value is repeated.

Step 2 Let us find the eigenvectors corresponding to the eigenvalues.

The eigenspace corresponding to $\lambda_1 = 1$ is

$$\left\{ k \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \mid k \in \mathbb{R} \right\}$$

A basis of the eigenspace is $\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\} = \{v_1\}$ (say)

The eigenspace corresponding to $\lambda_2 = -2$ is

$$\left\{ k \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \mid k \in \mathbb{R} \right\}$$

A basis of the eigen space is $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\} = \{v_2\}$ (say)

Thus there are only two linearly independent vectors v_1, v_2 and there are 3 eigen values. So A is not diagonalizable.

Problem 17.10. (Case of repeated eigenvalues) Diagonalize the matrix

$$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}, \text{ if possible.}$$

Step 1 The eigenvalues of A are the roots of the equation

$$|A - \lambda I| = 0$$

$$\implies \lambda = 1, -2, -2$$

Eigenvalues are $\lambda_1 = 1, \lambda_2 = -2$

Step 2 The eigenspace corresponding to $\lambda_1 = 1$ is

$$\left\{ k \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \mid k \in \mathbb{R} \right\}$$

A basis of the eigenspace is $\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\} = \{v_1\}$ (say)

Eigen space corresponding to $\lambda_2 = -2$ is

$$\left\{ k_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid k_1, k_2 \in \mathbb{R} \right\}$$

A basis is $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\} = \{v_2, v_3\}$ (say)

Thus there are 3 linearly independent vectors.

Step 3

$$\text{Let } P = [v_1 \quad v_2 \quad v_3] = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Then $A = PDP^{-1}$ or $D = P^{-1}AP$. so that A is diagonalizable.

Remark 17.1. Let A be a $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ (some may be repeated)

If the eigenspace of $\lambda_1, \lambda_2, \dots, \lambda_k$ are of dimensions m_1, m_2, \dots, m_k respectively, and

(i) if $m_1 + m_2 + \dots + m_k = n$, then A is diagonalizable.

(ii) If $m_1 + m_2 + \dots + m_k < n$ then A is not diagonalizable.

Example 17.4. Compute A^8 , where $A = \begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix}$

First we diagonalize A , if possible.

It can be easily seen that eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = 2$

Eigenvector corresponding to $\lambda_1 = 1$ is $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Eigenvector corresponding to $\lambda_2 = 2$ is $v_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

If $P = (v_1 \ v_2) = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}$, $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

then $D = P^{-1}AP$ so that $A = PDP^{-1}$.

$$\begin{aligned} \text{Hence } A^8 &= PD^8P^{-1} = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1^8 & 0 \\ 0 & 2^8 \end{pmatrix} \begin{pmatrix} -2 & +3 \\ +1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 766 & -765 \\ 510 & -509 \end{pmatrix} \end{aligned}$$

17.7 Exercise

1. Compute A^4 , where $A = PDP^{-1}$, where

(i) $P = \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

(ii) $P = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}$ $D = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$

(iii) $P = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ $D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$

2. If A, v_1, v_2 are given. Suppose you are told that v_1, v_2 are eigenvectors of A . Use this information to diagonalize A , where

(i) $A = \begin{pmatrix} -3 & 12 \\ -2 & 7 \end{pmatrix}$, $v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$, $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

3. Diagonalize the following matrices, if possible.

(i) $\begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}$

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

(iii) $\begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$

(iv) $\begin{pmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{pmatrix}$

(v) $\begin{pmatrix} 0 & -4 & -6 \\ -1 & 0 & -3 \\ 1 & 2 & 5 \end{pmatrix}$

$$(vi) \begin{pmatrix} -7 & -16 & 4 \\ 6 & 13 & -2 \\ 12 & 16 & 1 \end{pmatrix}$$

$$(vii) \begin{pmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{pmatrix}$$

17.8 Supplementary Exercises

1. State whether the following statements are true or false. Justify the false ones.
 - (i) If a transformation T moves a non-zero vector v parallel to itself, then v is an eigenvector of T .
 - (ii) The transformation of rotation of \mathbb{R}^2 about the origin has no eigenvector.
 - (iii) The identity matrix has a unique eigenvalue and eigenvector.
 - (iv) If λ is an eigenvalue of a $n \times n$ matrix A , then the subset S of \mathbb{R}^n consisting of all eigenvectors of A associated with the eigenvalue λ is a subspace of \mathbb{R}^n .
 - (v) The dimension of an eigenspace corresponding to an eigenvalue can be zero.
 - (vi) If A is an $n \times n$ matrix, then the sum of the dimensions of the eigenspaces associated with all the eigenvalues of A may exceed n .
 - (vii) The eigenspace associated with an eigenvalue λ of a matrix is same as the null space of $(A - \lambda I)$.
 - (viii) If W_λ is the eigenspace of A associated with eigenvalue λ , then $Av \in W_\lambda$, for all $v \in W_\lambda$.
 - (ix) If A is an $n \times n$ diagonal matrix with distinct entries d_i , $i = 1, 2, \dots, k$; d_i repeated k_i times, then d_i is an eigenvalue of A and the dimension of the eigenspace associated with d_i is less than k_i .
 - (x) If A is a scalar matrix then A has only one eigenvalue and the eigenspace is \mathbb{R}^n .
 - (xi) A square matrix whose eigenvalues are not all distinct is not similar to a diagonal matrix.
 - (xii) An $n \times n$ matrix over \mathbb{R} has exactly n real eigenvalues.
 - (xiii) The matrix whose characteristic equation is $x^3 - 3x^2 + 5x - 6 = 0$ can be a singular matrix.
 - (xiv) Every root of the characteristic equation of A a matrix over \mathbb{R} is a characteristic value of A .
 - (xv) A matrix which is similar to a diagonal matrix has exactly n eigenvalues.
 - (xvi) A matrix with eigenvalues $1, 2, 3$ cannot have its trace equal to zero.

- (xvii) The matrix $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$, $a \neq c$ is similar to a diagonal matrix.
- (xviii) If an $n \times n$ matrix has n distinct eigenvalues then the eigenspaces associated with each eigenvalue is at most 1-dimensional.
- (xix) $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ is always similar to a diagonal matrix.

2. Find the eigen values and eigen space of the matrix $\begin{pmatrix} 2 & 1 & 0 \\ 9 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Also find a basis for the eigen spaces.

3. Find the characteristic polynomial of

(i)
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \end{pmatrix}$$

(ii)
$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

4. Prove that the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $b \neq 0$ has no real eigenvalue.
5. Prove that the characteristic roots of a real symmetric matrix are real.
6. Prove that the only real characteristic roots of a real skew symmetric matrix is zero.
7. Prove that the characteristic vectors corresponding to distinct real roots of a real symmetric matrix are orthogonal.
8. Prove that the only real characteristic roots of an orthogonal matrix are ± 1 .
9. Prove that the only eigenvalue of nilpotent matrix is zero.
10. If A is a diagonal matrix with an entry d_1 repeated k times then prove that d_1 is an eigenvalue with multiplicity k and that the dimension of the eigenspace associated with the eigenvalue d_1 is k .
11. For square matrices A and B , show that the following are equivalent
- Zero is an eigenvalue of AB .
 - Either A or B is singular but not both.
 - Zero is an eigen value of BA .
12. If A and B are square matrices then AB and BA have the same eigen values.
13. If A and B are 2×2 matrices such that $\text{trace } A = \text{trace } B$, $|A| = |B|$ then prove that A and B have the same eigenvalues.
14. Let A and B are $n \times n$ matrices. If $I - AB$ is invertible the $I - BA$ is invertible and $(I - BA)^{-1} = I + B(I - AB)^{-1}A$

17.9 Answers to Exercises

Exercise - 17.3

6. *Hint:* $|A| = 0 \Rightarrow$ '0' is an eigenvalue of A , If λ is any other eigenvalue of A , then $Av = \lambda v \Rightarrow A^2v = \lambda Av = \lambda^2v \Rightarrow \lambda^2v = 0 \Rightarrow \lambda = 0$.
7. *Hint:* The columns of $|A - \lambda I|$ adds up to '0' so A is row equivalent to a matrix with one row of 0's. Therefore, $|A - \lambda I|$ is singular.
9. *Hint:* $\lambda = (11 \dots 1)^t$

Exercise - 17.5

1. (i) -1,1,5
 (ii) 2, 3, 0
 (iii) -1, -1, -1
 (iv) 3, -1, 3, 0
 (v) 2, 1, 0, -6.

2. (i) 0,
 (ii) 0
 (iii) 0
 (iv) 0

3. (i) 2,3
 (ii) 1,-1
 (iii) None
 (iv) 1,-2,-2
 (v) None

4. (i) $\frac{7 \pm \sqrt{21}}{2}$
 (ii) 0, $\pm i\sqrt{14}$
 (iii) 1, $1 + 2\sqrt{3}$
 (iv) 2, $\pm t$

5. *Hint:* λ is eigenvalue of $A \implies |A - \lambda I| = 0 \implies |A - \lambda \frac{A(\text{adj}A)}{|A|}| = 0$

$$A \frac{\lambda}{|A|} \left(\frac{|A|}{\lambda} I - (\text{adj}A) \right) = 0$$

$$\implies |A| \frac{\lambda}{|A|} \left((\text{adj}A) - \frac{|A|}{\lambda} I \right) = 0$$

$$|\text{adj}A - \frac{|A|}{\lambda} I| = 0 \quad \because \lambda \neq 0$$

$\frac{|A|}{\lambda}$ is an eigenvalue of A .

if $f(\lambda) = 0$ is the characteristic equation of A , then the characteristic equation of $(\text{adj}(A))$ is $f\left(\frac{|A|}{\lambda}\right) = 0$

6. $|A - \lambda I| = |(A - \lambda I)^t| = |A^t - \lambda I^t| = |A^t - \lambda I|$

7. $|kA - k\lambda I| = |k(A - \lambda I)| = k^n |A - \lambda I|$
 $\therefore |kA - k\lambda I| = 0 \iff |A - \lambda I| = 0$
 If X is an eigen vector corresponding to eigen value λ , then $AX = \lambda X$
 $\therefore k(AX) = k\lambda X \implies (kA)X = k\lambda X$
8. Use induction.
10. $(A^{-1}B - \lambda I) = A^{-1}(BA^{-1} - \lambda I)A$
 $\therefore |A^{-1}B - \lambda I| = |BA^{-1} - \lambda I|$
12. $\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = 5$
 $X_1 = (1 \ -3 \ 0)^t, X_2 = (1 \ 0 \ -9)^t, X_3 = (1 \ 3 \ 0)^t$
 for A^t $\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = 5$
 $X_1 = (9 \ -3 \ 1)^t, X_2 = (0 \ 0 \ 1)^t, X_3 = (9 \ 3 \ 1)^t$
13. $\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = 4$
 $X_1 = (1 \ 0 \ -1)^t, X_2 = (-2 \ -3 \ 2)^t, X_3 = (8 \ 5 \ 2)^t$
19. $-\lambda^3 + 18\lambda^2 - 45\lambda; 0, 3, 15$ Span $(1 \ 2 \ 2)^t, \text{Span}(-2 \ -1 \ 2)^t,$
 Span $(2 \ -2 \ 1)^t$.

Exercise - 17.7

1. (i) $\begin{pmatrix} 226 & -525 \\ 90 & -209 \end{pmatrix}$ (ii) $\frac{1}{16} \begin{pmatrix} 151 & 90 \\ 225 & -134 \end{pmatrix}$
- (iii) $\begin{pmatrix} 171 & -85 \\ -170 & 86 \end{pmatrix}$
2. $\lambda_1 = 1, \lambda_2 = 3$ $P = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ $P^{-1}AP = D$
3. (i) Not possible.
- (ii) $P = \begin{pmatrix} 1 & -3 \\ 1 & 4 \end{pmatrix}, D = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$
- (iii) $P = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$
- (iv) $P = \begin{pmatrix} -2 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix}, D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix}$
- (v) $P = \begin{pmatrix} -2 & -3 & -2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}, D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
- (vi) $P = \begin{pmatrix} -4 & 1 & -2 \\ 3 & 0 & 1 \\ 0 & 3 & 2 \end{pmatrix}, D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -3 \end{pmatrix}$

$$(vii) P = \begin{pmatrix} 1 & 8 & -58 \\ 0 & 5 & -49 \\ 0 & 0 & 14 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Supplementary Exercises

1. (i) True
 (ii) True
 (iii) False, unique eigen value, but every vector is an eigen vector.
 (iv) False, must contain zero vector also.
 (v) False
 (vi) False
 (vii) True
 (viii) True
 (ix) False
 (x) True
 (xi) False
 (xii) False
 (xiii) False
 (xiv) False, complex roots will not be characteristic value.
 (xv) True
 (xvi) True
 (xvii) True
 (xviii) False, exactly 1.
 (xix) False, if $b = 0$ it is true.
3. (i) $a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + \lambda^4 = 0$
 (ii) $\lambda^3 - \lambda^2 + \lambda - 1 = 0$
14. *Hint:* Let $(I - AB)^{-1} = C$
 $\therefore I = (I - AB)C$
 $\implies ABC + I = C$
 $\implies BABCA + BA = BCA$
 $\implies (I - BA)(I + BCA) = I$
 $\implies (I - BA)^{-1} = I + BCA$

Chapter 18

Markov Process

Definitions and Examples

Suppose that each year 6% of the population of Delhi migrate to Mumbai and 4% of the population of Mumbai migrate to Delhi. Let the population of Delhi be x_0 and that of Mumbai be y_0 . We are interested in knowing the population after 1, 2, 3 years.

Initial population vector $X_0 = [x_0 \ y_0]^t$ if x_1, y_1 are the populations of Delhi and Mumbai after 1 year.

$$\begin{aligned} \therefore \quad x_1 &= 0.94x_0 + 0.04y_0 \\ y_1 &= 0.06x_0 + 0.96y_0 \end{aligned}$$

In the matrix form, we get

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.94 & 0.04 \\ 0.06 & 0.96 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

If $X_1 = [x_1 \ y_1]^t$ is the population vector after 1 year

$$X_1 = AX_0$$

where

$$A = \begin{pmatrix} 0.94 & 0.04 \\ 0.06 & 0.96 \end{pmatrix}$$

If $X_2 = [x_2, y_2]^t$ is the population vector after 2 years, then we will get

$$X_2 = AX_1$$

or

$$X_2 = A^2X_0$$

Thus, the population vector $X_n = [x_n, y_n]^t$, after n years, is given by

$$X_n = AX_{n-1}$$

or

$$X_n = A^n X_0$$

The matrix A has the following characteristics:

- (i) Since the entries of each column of A represents the probability of a person residing in one of the two cities in the next year, therefore each entry is non-negative, and less than or equal to 1.
- (ii) Since a person has to reside in one of the two cities (because we are assuming that total population remains the same), therefore the sum of the entries in each column is 1.

Now we give the following definitions:

Definition 18.1. (Probability vector):

A vector $V = (x_1, x_2, \dots, x_n)^t$ is called a probability vector if

- (i) $x_i \geq 0 \quad i = 1, 2, \dots, n$
- (ii) $x_1 + x_2 + \dots + x_n = 1$

Definition 18.2. (Transition matrix):

Suppose that a system has n possible states, s_1, s_2, \dots, s_n . Let p_{ij} be the probability that if the given system is in state s_j at a certain period of observation, then it will be in state s_i at the next period of observation, for $i, j = 1, 2, \dots, n$. The matrix $P = (p_{ij})_{n \times n}$ is called the transition matrix of the system.

Note that we assume that p_{ij} remains the same for all time periods.

The transition matrix is also called stochastic matrix or probability matrix or Markov matrix. From the definition it follows that the columns of a transition matrix are probability vectors.

Example 18.1. $P = \begin{pmatrix} 0.2 & 0.1 & 0 \\ 0.5 & 0.8 & 0.3 \\ 0.3 & 0.1 & 0.7 \end{pmatrix}$ is a transition matrix as

- (1) All entries are non-negative and less than or equal to 1.
- (2) Sum of each column is 1.

Example 18.2. $P = \begin{pmatrix} 0 & 0.4 & 0.3 \\ 1 & 0.6 & 0.2 \\ 0 & 0.1 & 0.5 \end{pmatrix}$ is not a transition matrix as the sum of the entries of the second column is not 1.

Example 18.3. $P = \begin{pmatrix} -0.5 & 0.7 \\ 1.5 & 0.3 \end{pmatrix}$ is not a transition matrix even though the sum of the entries of each column is 1. This is because one of the entries is negative.

Definition 18.3. (Markov process):

A sequence of probability vectors X_0, X_1, \dots , and a transition matrix P such that

$$X_{i+1} = PX_i, i = 0, 1, 2, \dots$$

is called a Markov chain.

The vectors X_i in a Markov chain are called the state vectors at the i th stage.

Thus X_0 is the initial state vector. Since

$$\begin{aligned} X_{i+1} &= PX_0, i = 0, 1, 2, \dots \\ X_1 &= PX_0 \\ X_2 &= PX_1 = P^2 X_0 \\ X_3 &= PX_2 = P^3 X_0 \end{aligned}$$

In general $X_k = P^k X_0$, so that the vector at any state is expressible in terms of the initial vector.

Example 18.4. Give 3 terms of the Markov chain whose transition matrix is $\begin{pmatrix} 0.3 & 0.6 \\ 0.7 & 0.4 \end{pmatrix}$ and initial vector is $\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$.
Here

$$\begin{aligned} X_0 &= \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}, \\ P &= \begin{pmatrix} 0.3 & 0.6 \\ 0.7 & 0.4 \end{pmatrix} \end{aligned}$$

The first 3 terms are X_0, X_1, X_2 where $X_1 = PX_{i-1}, i = 0, 1, 2$
Thus

$$\begin{aligned} X_0 &= \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} \\ X_1 &= PX_0 = \begin{pmatrix} 0.42 \\ 0.58 \end{pmatrix} \\ X_2 &= PX_1 = \begin{pmatrix} 0.574 \\ 0.426 \end{pmatrix} \end{aligned}$$

Example 18.5. Two companies X and Y manufacture mobiles. Initially X has $\frac{3}{5}$ of the market while Y has $\frac{2}{5}$ of the market. Each year, company X keeps $\frac{1}{4}$ of its customers while $\frac{3}{4}$ switch to Y where as company Y keeps $\frac{2}{3}$ of its customers while $\frac{1}{3}$ switch to X . Find

- (i) Initial state vector
- (ii) The transition matrix
- (iii) The distribution of market after 1 year, 2 years and 3 years.

Solution: (i) If X_0 is the initial market share, then

$$X_0 = \begin{pmatrix} \frac{3}{5} \\ \frac{2}{5} \end{pmatrix} \quad (18.1)$$

(ii) The Transition matrix is

$$P = \begin{pmatrix} \text{From } X & Y & \text{To} \\ & X & Y \\ \frac{1}{4} & \frac{1}{3} & X \\ \frac{3}{4} & \frac{2}{3} & Y \end{pmatrix} \quad (18.2)$$

(iii) Market distribution after 1 year is given by

$$X_1 = PX_0$$

Using (18.1) and (18.2)

$$X_1 = \begin{pmatrix} \frac{17}{60} \\ \frac{43}{60} \end{pmatrix}$$

Market distribution after 2 years

$$X_2 = PX_1$$

$$X_2 = \begin{pmatrix} \frac{223}{720} \\ \frac{497}{720} \end{pmatrix}$$

Market distribution after 3 years

$$X_3 = PX_2 = \begin{pmatrix} \frac{2657}{8640} \\ \frac{5983}{8640} \end{pmatrix}$$

We now define the steady state vector:

Definition 18.4. (Steady state vector):

Let X_0, X_1, X_2, \dots be a Markov chain with transition matrix P . If the sequence X_0, X_1, X_2, \dots is convergent then $\lim_{n \rightarrow +\infty} X_n$ is called the steady state vector of the Markov chain.

Thus, if X is the steady state vector of a Markov chain with transition matrix P , then

$$X_{k+1} = PX_k$$

Taking limits as $k \rightarrow +\infty$, we get

$$X = PX$$

Equivalently,

$$(P - I)X = \tilde{0}$$

Thus the steady state vector is a solution of the homogeneous system

$$(P - I)X = \tilde{0}$$

Example 18.6. *Find the steady state vector in illustration.*

If $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is the steady state vector, then

$$x_1 + x_2 = 1 \tag{18.3}$$

as X is a probability vector. Also X is the solution of

$$\begin{aligned}
 & PX = X \Rightarrow (P - I)X = 0 \\
 \Rightarrow & \begin{pmatrix} \frac{1}{4} - 1 & \frac{1}{3} \\ \frac{3}{4} & \frac{2}{3} - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \\
 \Rightarrow & \begin{pmatrix} -\frac{3}{4} & \frac{1}{3} \\ \frac{3}{4} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \\
 \Rightarrow & -\frac{3}{4}x_1 + \frac{1}{3}x_2 = 0 \tag{18.4}
 \end{aligned}$$

$$\frac{3}{4}x_1 - \frac{1}{3}x_2 = 0 \tag{18.5}$$

Solving (18.3), (18.4) and (18.5) we get $x_1 = \frac{4}{13}, x_2 = \frac{9}{13}$.
 Steady state vector = $\begin{pmatrix} \frac{4}{13} \\ \frac{9}{13} \end{pmatrix}$

Problem 18.1. There are two brands of tea, A and B , used by persons of a certain town. Each year 30% users of Brand A start using Brand B , whereas 20% users of Brand B , start using Brand A . Initially, there are 8000 users of brand A and 2000 users of Brand B . Assuming that the total no. of users remains constant, how many users of each brand will there be after

- (i) 1 year?
- (ii) 2 years?
- (iii) What is the steady state?

Solution: The initial vector is

$$X_0 = \begin{pmatrix} \frac{8000}{10000} \\ \frac{2000}{10000} \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$$

The transition matrix is

$$P = \begin{pmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{pmatrix}$$

Vector X_k after k years is given by

$$\begin{aligned}
 & X_k = PX_{k-1}, k = 1, 2, \dots \\
 \therefore X_1 &= PX_0 = \begin{pmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{pmatrix} \begin{pmatrix} 0.80 \\ 0.20 \end{pmatrix} = \begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix} \\
 X_2 &= PX_1 = \begin{pmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{pmatrix} \begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix} = \begin{pmatrix} 0.50 \\ 0.50 \end{pmatrix}
 \end{aligned}$$

Steady state vector X is the solution of

$$PX = X$$

If $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, then

$$x_1 + x_2 = 1 \tag{18.6}$$

as X is a probability vector.

$$\begin{aligned} \text{Also } PX &= X \\ \Rightarrow (P - I)X &= 0 \\ \Rightarrow \begin{pmatrix} -1 + 0.70 & 0.20 \\ 0.30 & -1 + 0.80 \end{pmatrix} X &= 0 \end{aligned}$$

$$\Rightarrow -3x_1 + 2x_2 = 0 \tag{18.7}$$

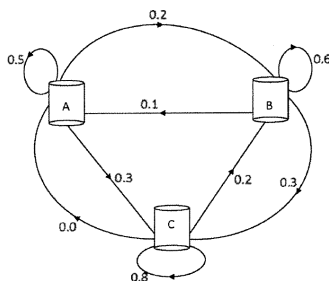
$$\Rightarrow 3x_1 - 2x_2 = 0 \tag{18.8}$$

Solving (18.6), (18.7) and (18.8) we get

$$\begin{aligned} x_1 &= \frac{2}{5} = 0.40 \\ x_2 &= \frac{3}{5} = 0.60 \end{aligned}$$

- (i) After 1 year, $X_1 = \begin{pmatrix} 0.60 \\ 0.40 \end{pmatrix}$
 Number of users of Brand $A = .6 \times 10000 = 6000$
 Number of users of Brand $B = .4 \times 10000 = 4000$
- (ii) After 2 years, $X_2 = \begin{pmatrix} 0.50 \\ 0.50 \end{pmatrix}$
 Number of users of Brand $A = .5 \times 10000 = 5000$
 Number of users of Brand $B = .5 \times 10000 = 5000$
- (iii) Steady state vector $X = \begin{pmatrix} 0.40 \\ 0.60 \end{pmatrix}$
 Number of users of Brand $A = .4 \times 10000 = 4000$
 Number of users of Brand $B = .6 \times 10000 = 6000$

Problem 18.2. In a college canteen three brands A, B, C of a soft drink are available. Every year the change of liking of the students from a particular brand to another is shown by the following diagram.



Write the transition matrix for the Markov chain.

Solution: From the figure we see that the transition matrix is:

$$P = \begin{pmatrix} \text{From } A & B & C & \text{To} \\ & 0.5 & 0.1 & 0.0 & A \\ & 0.2 & 0.6 & 0.2 & B \\ & 0.3 & 0.3 & 0.8 & C \end{pmatrix}$$

Problem 18.3. In the above problem, if initially the distribution for brands A, B, C is $\begin{pmatrix} 0.30 \\ 0.30 \\ 0.40 \end{pmatrix}$, find the distribution after 1 year. What is steady state vector?

Solution: Here $X_0 = \begin{pmatrix} 0.30 \\ 0.30 \\ 0.40 \end{pmatrix}$
Distribution vector after 1 year is

$$X_1 = PX_0 = \begin{pmatrix} 0.18 \\ 0.32 \\ 0.50 \end{pmatrix}$$

If X is the steady state vector, then X is the solution of

$$PX = X$$

$$\Rightarrow (P - I)X = 0 \quad (18.9)$$

Also $X = [x_1 \ x_2 \ x_3]^t$ is a probability vector, so that

$$x_1 + x_2 + x_3 = 1 \quad (18.10)$$

$$(1) \Rightarrow \begin{pmatrix} -0.5 & 0.1 & 0.0 \\ 0.2 & -0.4 & 0.2 \\ 0.3 & 0.3 & -0.2 \end{pmatrix} X = 0 \quad (18.11)$$

We are required to solve the equations (18.10) and (18.11), i.e. the system

$$\begin{pmatrix} 1 & 1 & 1 \\ -0.5 & 0.1 & 0.0 \\ 0.2 & -0.4 & 0.2 \\ 0.3 & 0.3 & -0.2 \end{pmatrix} X = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (18.12)$$

The reduced echelon form of the augmented matrix of system (18.12) is

$$\left(\begin{array}{ccc|c} \mathbf{1} & 0 & 0 & \frac{1}{15} \\ 0 & \mathbf{1} & 0 & \frac{1}{3} \\ 0 & 0 & \mathbf{1} & \frac{3}{5} \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Thus the solution of system (18.12) is

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{15} \\ \frac{1}{3} \\ \frac{3}{5} \end{pmatrix} \approx \begin{pmatrix} 0.07 \\ 0.33 \\ 0.60 \end{pmatrix}$$

X is the steady state vector.

Problem 18.4. Show that every 2×2 transition matrix has at least one steady state vector. There are two linearly-independent steady state vectors if the transition matrix is I_2 .

Solution: By the definition of transition matrix, any 2×2 transition matrix is of the form $\begin{pmatrix} \alpha & 1-\beta \\ 1-\alpha & \beta \end{pmatrix}$ with $0 \leq \alpha, \beta \leq 1$.

Let

$$P = \begin{pmatrix} \alpha & 1-\beta \\ 1-\alpha & \beta \end{pmatrix}$$

The steady state vector X is a solution of

$$PX = X$$

$$\Rightarrow (P - I)X = 0 \quad (18.13)$$

$$P - I = \begin{pmatrix} \alpha - 1 & 1 - \beta \\ 1 - \alpha & \beta - 1 \end{pmatrix} \sim \begin{pmatrix} \alpha - 1 & 1 - \beta \\ 0 & 0 \end{pmatrix}$$

Since the echelon form has a row of zeroes, therefore there is one free variable. Hence Eq. 18.13 has at least one solution.

Thus there is at least one steady state vector.

If $\alpha = 1, \beta = 1$ then $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

and $P - I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

so that there are two free variables. This gives two linearly-independent solution of Eq. 18.13. This gives two linearly-independent steady state vectors. (In fact every vectors in \mathbb{R}^2 is a steady state vector.)

Problem 18.5. Let P be a transition matrix. Then the following properties hold:

- (i) The rows of $(P - I)$ are linearly dependent.
- (ii) Row rank of $(P - I)$ is less than n .
- (iii) The null space of $(P - I)$ is non-zero. Thus $PX = X$ always has a non trivial solution.

Solution: Let $P = (p_{ij})_{n \times n}$ be a transition matrix. Then $0 \leq p_{ij} \leq 1 \quad \forall i, j = 1, \dots, n$

For each $j = 1, 2, \dots, n$

$$p_{1j} + p_{2j} + \dots + p_{nj} = 1 \quad (18.14)$$

$$P - I = \begin{pmatrix} p_{11} - 1 & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} - 1 & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} - 1 \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix}$$

(i) Applying $R_n \rightarrow R_n + R_1 + \cdots + R_{n-1}$, the last row is
 $[p_{11}+p_{21}+\cdots+p_{n1}-1 \quad p_{12}+p_{22}+\cdots+p_{n2}-1 \quad \cdots \quad p_{1n}+p_{2n}+\cdots+p_{nn}-1]$
 i.e. $[0 \quad 0 \quad \cdots \quad 0]$
 Thus $R_1 + \cdots + R_n = 0$
 Hence the rows of $(P - I)$ are linearly dependent.

(ii) By (i) the rows of $(P - I)$ are linearly dependent. Thus, there can be at most $n - 1$ linearly independent rows. Therefore dimension of row space $\leq n - 1$
 Since row rank = dimension of row space
 \therefore row rank of $(P - I) \leq n - 1 < n$

(iii) Null space of $(P - I)$ is the set of all solutions of

$$(P - I)X = 0 \quad (18.15)$$

Since column rank of $(P - I) = \text{row rank of } (P - I) < n$ by (ii)

\therefore Column rank of $(P - I) < n$.

Hence (18.15) has a non-trivial solution. Hence $(P - I)$ has non-zero null space.

$\therefore \exists 0 \neq X_0$ such that $(P - I)X_0 = 0$

Hence

$$PX_0 = X_0$$

$\therefore PX = X$ has a non-zero solution.

18.1 Exercise

- The land use in a city in 2009 is 30% residential, 20% commercial and 50% industrial. If the transition matrix is given by

$$\begin{pmatrix} 0.8 & 0.1 & 0.0 \\ 0.1 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.9 \end{pmatrix}$$

Find the distribution after 1 year, 2 years.

- Which of the following matrices are not transition matrices. Give reasons for your answer.

(i) $\begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$

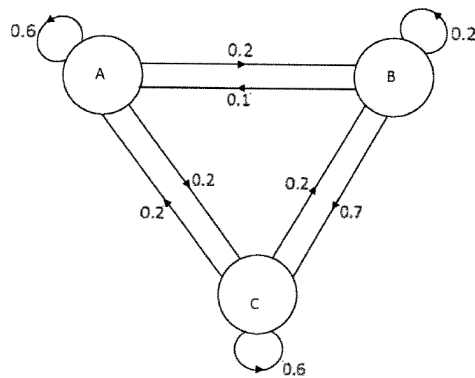
(ii) $\begin{pmatrix} 0.2 & 0.3 & 0.8 \\ 0.9 & 0.7 & 0.1 \\ 0.1 & 0.0 & 0.1 \end{pmatrix}$

$$(iii) \begin{pmatrix} 0.8 & 0.2 & 1.0 \\ 0.2 & 0.8 & 0.0 \\ 0.0 & 0.0 & 0.0 \end{pmatrix}$$

$$(iv) \begin{pmatrix} 0.2 & 0.6 & 0.3 \\ 0.2 & 0.1 & 0.1 \\ 0.4 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0.5 \end{pmatrix}$$

$$(v) \begin{pmatrix} 0.2 & 0.4 & 0.3 \\ 0.0 & -0.1 & 0.3 \\ 0.8 & 0.7 & 0.4 \end{pmatrix}$$

3. A car rental company XYZ has branch offices in three cities A, B and C. A car rented from one office can be left at any other office. The company started the business with 30 cars at each of the offices. The monthly distribution is shown by the following diagram:



Find the distribution after 1 month, 2 months. Also find the steady state.

4. Find the steady state vector for the following transition matrices:

$$(i) \begin{pmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 0.9 & 0.4 \\ 0.1 & 0.6 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 0.5 & 0.8 \\ 0.5 & 0.2 \end{pmatrix}$$

5. Find the steady state vector for the following matrices:

$$(i) \begin{pmatrix} 0.8 & 0.1 & 0.2 \\ 0.1 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.7 \end{pmatrix}$$

$$(ii) \begin{pmatrix} 0.6 & 0.1 & 0.3 \\ 0.3 & 0.7 & 0.2 \\ 0.1 & 0.2 & 0.5 \end{pmatrix}$$

6. A new means of transport, namely 'the Metro', has gone into operation in Delhi. Studies made by the transport authority predict the percentage of commuters who switch over to Metro or continue using their old means of transport. The transition matrix is $\begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}$. Suppose that the population of the commuters remains constant and initially 30% of the commuters use Metro. What percentage of the commuters will be using the Metro after 1 year, 2 years?
7. A behavioural psychologist places a rat every day in a cage with two doors, A and B. The rat can go through door A where it receives an electric shock, or through door B, where it receives food. A record is made of the door through which the rat passes. At the start of the experiment, on a Monday, the rat is equally likely to go through door A as through door B. After going through door A and receiving a shock, the probability of going through the same door on the next day is 0.3. After going through door B and receiving food, the probability of going through the same door on the next day is 0.6.
- Write the transition matrix.
 - What is the probability of going through door B Wednesday?
 - What is the probability of going through door A on Thursday?
8. The students of Vidya Mandir School are given one of the 3 drinks every day — milk, juice or coffee. If they are given milk today, the chances of getting juice or coffee tomorrow are 30% and 10% respectively. If today they get juice, then the chances of getting milk or juice tomorrow are 40% and 30% respectively. Finally if they get coffee today, then there are 40% chances of getting milk and 50% chances of getting juice on the next day.
- Write the transition matrix.
 - If today there are 50% chances of getting milk and 50% chances of getting juice, what are the chances of getting coffee tomorrow, and juice the day after tomorrow.
 - Suppose that on a Tuesday there are 40% chances of getting milk and 60% chances of getting coffee. What are the chances for
 - Juice on Thursday
 - Milk on Friday.
 - In the long run, what are chances of milk being served?

18.2 Answers to Exercises

Exercises - 18.1

- After 1 year 26% residential, 22% commercial, 52% industrial
After 2 years 23% residential, 23.2% commercial, 53.8% industrial

- (2) (i) Column Sum is not 1
(ii) Sum of elements of 1st column $\neq 1$
(iv) Not a square matrix
(v) Has a negative entry .
- (3) (27, 18, 45) in offices at A, B, C respectively after 1 month and 2 months
(.3, .2, .5) the steady state vector.
- (4) (i) $\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}$
(ii) $\begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix}$
(iii) $\begin{pmatrix} \frac{8}{13} \\ \frac{5}{13} \end{pmatrix}$
- (5) (i) $\begin{pmatrix} \frac{7}{16} \\ \frac{4}{16} \\ \frac{5}{16} \end{pmatrix}$
(ii) $\begin{pmatrix} \frac{11}{37} \\ \frac{17}{37} \\ \frac{9}{37} \end{pmatrix}$
- (6) $\begin{pmatrix} 0.35 \\ 0.65 \end{pmatrix}, \begin{pmatrix} 0.375 \\ 0.625 \end{pmatrix}$
- (7) (i) $\begin{pmatrix} 0.3 & 0.4 \\ 0.7 & 0.6 \end{pmatrix}$
(ii) 0.635
(iii) 0.3635
- (8) (i) $\begin{pmatrix} 0.6 & 0.4 & 0.4 \\ 0.3 & 0.3 & 0.5 \\ 0.1 & 0.3 & 0.1 \end{pmatrix}$
(ii) 20%, 34%
(iii) (a) 32%
(b) 49.92%

This page is intentionally left blank.

Index

- p -rowed minor, 464
- Diagonal Matrix, 442
- Scalar matrix, 442

- Abelian group, 185
- Adjoint of a Matrix, 464
- an echelon form, 363
- Angle between two vectors, 476
- antisymmetric, 26
- associates, 141
- Associative operation, 51
- Associativity, 186

- basic variables, 395
- Basis, 615
- bijective, 83
- binary operation, 49
- Binary relation, 20
- block diagonal matrix, 467
- block upper triangular matrix, 467

- Cardinality of a finite set, 111
- Cartesian product of sets, 16
- centralizer, 240
- Centralizer of a subset, 240
- Centralizer of an element, 239
- Centre of a group, 241
- Centre of a ring, 313
- characteristic equation, 483, 726
- characteristic polynomial, 483, 726
- characteristic root, 483
- characteristic roots, 726
- Closure, 186
- co-domain of transformation, 511
- Cofactor, 464
- Column Rank, 635
- Commutative Operation, 53
- Commutativity, 186
- Comparable matrices, 443
- Complement of a set, 12
- composite, 100, 141

- congruent, 166
- Conjugate of a matrix, 450
- coordinates, 617
- coordinates of a vector, 617
- Countable set, 112
- Countably infinite set, 112
- cyclic group, 271
- Cyclic subgroup, 253
- cyclic subgroup, 271

- diagonal set, 23
- Diagonalizable Matrix, 490, 738
- Difference of two sets, 10
- Dimension, 628
- Disjoint sets, 9
- Divisibility, 140
- Division ring, 318
- domain of transformation, 511
- Dot product, 476

- Eigenspace, 721
- eigenspace, 484
- eigenvalue, 483
- eigenvalues, 726
- eigenvector, 483
- Eigenvector and Eigenvalue, 719
- Eigenvector and Eigenvalues of a matrix, 720
- Empty set, 4
- Equality of Matrices, 443
- Equality of sets, 4
- Equipollent sets, 111
- Equivalence class, 28
- equivalence relation, 27
- Existence of identity, 186
- Existence of inverse, 186
- Extension of a function, 85

- Field, 318
- Finite Dimensional Vector Space, 627
- finite order, 245

- Finite set, 5, 111
- free variables, 395
- function, 77
- generator, 271
- Graph of a Relation, 21
- Greatest common divisor, 150
- group, 184
- groupoid, 184
- Hermitian matrix, 457
- horizontal shear, 540
- Identity matrix, 443
- identity relation, 23
- Identity transformation, 514
- identity transformation, 653
- image, 77
- image of transformation, 511
- infinite order, 185, 245
- injective, 83
- Integral domain, 318
- Intersection of two sets, 8
- Invariant subspace, 722
- Inverse of a Function, 97
- Inverse of a Matrix, 463
- Inverse of a relation, 22
- Invertible Elements, 52
- Kernel, 666
- latent root, 483
- latent roots, 726
- leading entry, 359
- least common multiple (lcm), 159
- left identity, 220
- left inverse, 220, 463
- linear combination, 379
- linear congruence, 168
- linear transformation, 653
- Linearly dependent vectors, 601
- linearly independent, 602
- linear operator, 525
- Linear transformation, 524
- lower triangular matrix, 443
- mapping, 77
- Markov chain, 749
- Markov matrix, 749
- Markov process, 749
- Matrix addition, 444
- matrix of the linear transformation, 692
- Matrix transformation, 513
- maximal linearly independent, 627
- Minor, 464
- monoid, 184
- Norm or length of a vector, 476
- Normal Vector, 476
- normalizer, 244
- Normalizer of a subset, 243
- Null matrix, 443
- Nullity, 667
- Nullity A, 636
- one-one, 83
- One-to-one correspondence, 111
- onto, 83
- order, 185
- Order of an element, 245
- Orthogonal matrix, 479
- Orthogonal Vectors, 477
- Orthogonal vectors, 476
- Orthonormal Vectors, 477
- partial order, 27
- partially ordered set (Poset), 27
- partition of a set, 29
- pivot column, 363
- pivot position, 363
- postfactor, 447
- Power set, 5
- pre-image of transformation, 511
- prefactor, 447
- preimage, 77
- Prime Number, 141
- probability matrix, 749
- Probability vector, 749
- Product of vectors, 446
- Quotient set, 28
- range of transformation, 511
- Rank, 666
- Rank of a matrix, 635
- reflexive, 22
- Relatively prime, 151
- remainder, 164
- Restriction of a function, 85
- Reversal law for transpose, 449
- right identity, 220
- right inverse, 220, 463

- row echelon form, 359
- Row Matrix, 442
- Row rank, 635

- Scalar Multiplication, 445
- semigroup, 184
- Set, 2
- Shear along a line, 539
- shear factor, 540
- Similar Matrices, 489
- Singleton, 4
- Singular matrix, 465
- skew Hermitian matrix, 457
- skew symmetric matrix, 456
- Square Matrix, 442
- standard matrix, 693
- state vectors, 749
- Steady State Vector, 751
- stochastic matrix, 749
- subgroup, 231
- Subset, 5
- surjective, 83
- symmetric, 24
- Symmetric difference of two sets, 14
- symmetric matrix, 455

- the reduced echelon form, 363
- Trace of a matrix, 451
- Tranjugate of a matrix, 450
- Transformation, 510
- Transition Matrix, 680, 749
- transitive, 25
- Transpose of a matrix, 449
- Triangular matrix, 443
- trivial transformation, 653

- uncountable, 112
- Union of two sets, 7
- Unit matrix, 443
- Unitary Matrix, 478
- Universal Set, 3
- upper triangular matrix, 443

- weights, 379

- Zero Divisor, 318
- Zero matrix, 443
- zero transformation, 653
- Zero transformation, 514