स्वाध्याय स्वमन्थन स्वावलम्बन

UTTAR PRADESH RAJARSHI TANDON OPEN UNIVERSITY

(Established vide U.P. Govt. Act No. 10, of 1999)

UGMM-04
Elementary Algebra

FIRST BLOCK
Solutions of Polynomial Equations

INDIRA GANDHI NATIONAL OPEN UNIVERSITY

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

Shantipuram (Sector-F), Phaphamau, Allahabad-211013

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

UGMM-04
Elementary Algebra

Block

1

SOLUTIONS OF POLYNOMIAL EQUATIONS

ELEMENTARY ALGEBRA

'Excellent', I cried. 'Elementary', said he.

Sir Arthur Conan Doyle

This is our first course in algebra. The word 'algebra' comes from 'al-jabr', the Arabic term for the following process: if one or both sides of an equation contain a negative term, then this is taken to the other side so as to make all the terms positive.

Of course, as you know from your previous studies, algebra encompasses much more. Mathematicians through the centuries have built up a huge body of knowledge. We aim to give you a glimpse into some of it that is basic for studying other mathematics courses. This will also be useful to you while applying mathematics for solving problems in other areas.

We start the course by introducing you to the basic concept of sets, which was born in the 20th century. Then we will take you back to the 16th and 17th centuries, when European mathematicians were making great discoveries in algebra. It was in this period that a group of Italian mathematicians discovered methods for solving polynomial equations of degree 3 and 4. This work was also responsible for the discovery of complex numbers. In Unit 2 we shall talk about complex numbers; and in Unit 3 about the work of the Italians, that we just mentioned.

In the second half of the course we will acquaint you with some algebra that was known to ancient Indian, Egyptian and Babylonian mathematicians, namely, the theory of linear equations. In the Bakhshali Manuscript (c. 4th century A.D.), found near Takshshila, there are several verses dealing with linear equations. We will discuss linear equations taken singly, as well as three methods for obtaining common solutions for several linear equations. One of these methods is due to the great mathematician Gauss. Another method, due to Cramer, involves the concept of determinants. We shall develop this concept and discuss Cramer's method. We will go into greater depth about systems of equations in our course on linear algebra.

We end the course with a unit on some well-known and often used inequalities. Some were known to ancient mathematicians, and some were developed in the 19th century. We shall discuss them and see how they can be usefully applied.

Now, a few words about the way we have presented the course. It is divided into two blocks. In each block we have first introduced you to the block and given a list of symbols that are used in the block. Then we have presented the units of the block. In each unit we have interspersed exercises with the text. They are meant to help you check whether you've understood the material that we have discussed in that section or sub-section. We have also given our solutions to the exercises in a section at the end of the unit.

At the end of each block we have given a set of miscellaneous exercises covering the contents of the block. Doing them will give you some practice and a better understanding of what we have done in the course, though it is not necessary for you to do them.

A special feature of Block 1 is an appendix on some commonly used mathematical symbols and methods of proof. The contents of this appendix will be of help to you for studying this course, or any other mathematics course.

Now a word about our notation. Each unit is divided into sections, which may be further divided into sub-sections. These sections / sub-sections are numbered sequentially, as are the exercises and important equations in a unit. Since the material in the different units is heavily interlinked, we will be doing a lot of cross-referencing. For this we will be using the **notation Sec. x.y to mean Section y of Unit x.**

Another compulsory component of this course is its assignment, which covers the whole course. Your academic counsellor will evaluate it and return it to you with detailed comments. Thus, the assignment is meant to be a teaching as well as an assessment aid.

The course material that we have sent you is self-sufficient. If you have a problem in understanding any portion, please ask your academic counsellor for help. Also, if you feel like studying any topic in greater depth, you may consult

Higher Algebra by Hall and Knight, Book Palace, 1991.

This book will be available at your study centre.

We hope you will enjoy this course!

BLOCK 1 SOLUTIONS OF POLYNOMIAL EQUATIONS

For centuries a basic concept in mathematics has been that of numbers. But, the mathematicians of the 20th century have developed another very basic concept, that of a set. You will be using sets and their operations in any mathematical study. Thus, it is appropriate that we begin this course with a unit on sets.

You probably know that various mathematicians created negative numbers, fractions and real numbers (like $\sqrt{2}$) in response to a need for them. In Unit 2 we shall talk of a further need that led to the creation of a new kind of number, namely, a complex number. We shall define a complex number as a binomial as well as an ordered pair. We shall also discuss the algebra of complex numbers.

In your previous mathematical studies you may have met polynomial equations of degree 1 and 2. In Unit 3 we shall briefly recall their definitions and solution sets. Then we shall discuss methods of obtaining solutions of polynomial equations of degree 3 and 4. These methods are due to several 16th century mathematicians, and are still applied for obtaining exact solutions.

After these units we have given a set of miscellaneous exercises that cover the contents of this block. Doing these exercises is optional; but the more exercises you do, the better.

Finally, at the end of the block, we have given an appendix that deals with some mathematical logic. As you know, mathematics deals with reasoning and the laws of thought. Any mathematical activity involves the use of logical processes for proving assertions or for solving problems. It also involves the use of various symbols for brevity. This is why we felt we should briefly discuss some frequently used symbols. In the appendix we have done so. We have also discussed some methods of proof there.

In the next block you will need a lot of mathematics that you will study in this block. So go through all the units carefully.

Notations and Symbols

$\{x \mid x \text{ satisfies } P\}$	the set of all x such that x satisfies the property P.
N	the set of natural numbers
$Z\ (Z^*)$	the set of integers (non-zero integers)
$Q\ (Q^*)$	the set of rational numbers (non-zero rational numbers)
$R\ (R^*)$	the set of real numbers (non-zero real numbers)
$C\ (C^*)$	the set of complex numbers (non-zero complex numbers)
ϕ	the empty set
$\in$	belongs to
$\notin$	does not belong to
$\subseteq\ (\subsetneq)$	is contained in (is properly contained in)
$\not\subseteq$	is not contained in
$A \cup B$	union of the sets A and B
$A \cap B$	intersection of the sets A and B
$A \setminus B$	the set of elements of A that are not in B
A^c	complement of A
$A \times B$	Cartesian product of A and B
$\exists$	there exists
$\forall$	for all
$\Rightarrow$	implies
$\Leftrightarrow$	implies and is implied by
iff	if and only if
$a \mid b$	a divides b
$a \nmid b$	a does not divide b
$\therefore$	therefore
i.e.	that is
$< (\leq)$	less than (less than or equal to)
$> (\geq)$	greater than (greater than or equal to)
Re z	real part of the complex number z
Im z	imaginary part of the complex number z
Arg z	the argument of the complex number z
$\bar{z}$	the complex conjugate of the complex number z
deg f	the degree of the polynomial f

Greek Alphabets

α	Alpha	κ	Kappa	σ, Σ	Sigma (capital sigma)
β	Beta	λ	Lambda	τ	Tau
γ	Gamma	μ	Mu	υ	Upsilon
δ	Delta	ν	Nu	ϕ	Phi
ε	Epsilon	ξ	Xi	χ	Chi
ζ	Zeta	o	Omicron	ψ	Psi
η	Eta	π, Π	Pi (capital pi)	ω	Omega
θ	Theta	ρ	Rho		
ι	Iota				

UNIT 1 SETS

Structure

1.1 INTRODUCTION

Consider the collection of words that are defined in a given dictionary. A word either belongs to this collection or not, depending on whether it is listed in the dictionary or not. This collection is an example of a set, as you will see in Section 1.2. When you start studying any part of mathematics, you will come into contact with one or more sets. This is why we want to spend some time in discussing some basic concepts and properties concerning sets.

In this unit we will introduce you to various examples of sets. Then we will discuss some operations on sets. We will also introduce you to Venn diagrams, a pictorial way of describing sets.

As mentioned earlier, a knowledge of the material covered in this unit is necessary for studying any mathematics course. So please study this unit carefully.

And now we will list the objectives of this unit. After going through the unit, please read this list again and make sure that you have achieved the objectives.

Objectives

After studying this unit you should be able to

- identify a set;
- represent sets by the listing method, property method and Venn diagrams;
- perform the operations of complementation, union and intersections on sets;
- prove and apply the distributive laws;
- prove and apply De Morgan's laws;
- obtain the Cartesian product of two or more sets.

1.2 SETS

You may have often come across categories, classes or collections of objects. In mathematics, a well-defined collection of objects is called a set. The adjective 'well-defined' means that given an object it should be possible to decide whether it belongs to the collection or not. For example, the collection of all women pilots is a set. This is because any person is either a woman pilot or not, and accordingly she/he does or does not belong to the collection. On the other hand, the collection of all intelligent human beings is not a set. Why? Because, a particular human being may seem intelligent to one person and not to another. In other words, there is no clear criterion for deciding on who is intelligent and who is not. So, the collection is not well-defined.

Now we give some more examples of sets which you may have come across. We will also use them a great deal in this course.

i) The set of natural numbers, denoted by N.

ii) The set of integers, denoted by Z.

iii) The set of rational numbers, denoted by Q.

iv) The set of real numbers, denoted by R.

In the next unit we will be studying another set, namely, the set of complex numbers, denoted by C.

You may like to try this exercise now.

E1) Which of the collections mentioned below are sets?

a) The collection of all good people in India.

b) The collection of all those people who have been to Mars.

c) The collection of prime numbers.

d) The collection of even numbers.

e) The collection of all rectangles that are not squares.

A prime number is a natural number other than one, whose only factors are one and itself.

An object that belongs to a set is called an element or member of that set. For example, 2 is an element of the set of natural numbers.

We normally use capital letters A, B, C, etc., to denote sets. The small letters a, b, c, x, y, etc., are usually used to denote elements of sets.

We symbolically write the statement 'a is an element of the set A' as $a \in A$.

If a is not an element of A or, equivalently, a does not belong to A, we write it as $a \notin A$.

So, for example, if A is the set of prime numbers, then $5 \in A$ and $9 \notin A$.

The symbol $\in$ stands for 'belongs to'. It was suggested by the Italian mathematician Peano (1858-1932)

Try the following exercise now.

E2) Which of the following statements are true?

a) $0.2 \in N$

b) $2 \notin N$

c) $\sqrt{2} \in R$

d) $\sqrt{2} \in Q$

e) $\sqrt{-1} \notin R$

f) Any circle is a member of the set in E1 (e).

Now, you know that a number is either rational or irrational, but not both. So, what will the set of all numbers that are rational as well as irrational be? It will not have any element.

The set which has no elements is called the **empty set** (or the **void set**, or the **null set**). It is denoted by the Greek letter ϕ (phi).

A set, which has at least one element is called a **non-empty set**. We usually describe a non-empty set in two ways—the **listing method** and the **property method**.

In the first method we list all the members of the set within curly brackets. For instance, the set of all natural numbers that are factors of 10 is $\{1, 2, 5, 10\}$.

But what if the set has too many elements to be able to write them all down? In this case we list some of the elements of the set, enough to exhibit some pattern which its elements follow. For example, the set **N** of natural numbers can be described as

$N = \{1, 2, 3, \ldots\}$,

and the set of all even numbers lying between 10 and 100 is

$\{12, 14, 16, \ldots, 98\}$.

This method of representing sets is called the **listing method** (or tabular method, or roster method)

In the second method of describing a set we describe its elements by means of a property possessed by all of them. As an example, consider the set S of all natural numbers which are multiples of 5. This set S can be written in the form

$S = \{x \mid x \in N \text{ and } x \text{ is a multiple of } 5\}$. .... (1)

The vertical bar after x denotes 'such that'. (Some authors use : instead of | for such that.)

So (1) states that S is the set of all x such that x is a natural number and x is a multiple of 5.

We can also write this in a slightly shorter form as

$S = \{x \in N \mid x \text{ is a multiple of } 5\}$, or as

$S = \{5n \mid n \in N\}$.

This method of describing the set is called the **property method** or the set - builder method.

In some cases we can use either method to describe the set under consideration.
For instance, the set E, of all natural numbers less than 10, can be described as

$E = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ (by the listing method), or
$E = \{x \mid x \text{ is a natural number less than } 10\}$ (by the property method).

You can see that both these sets are the same, since both of them have precisely the same elements. This example leads us to the following definition.

Definition : Two sets S and T are called equal, denoted by $S = T$, iff every element of S is an element of T and every element of T is an element of S.

Now, while describing a set by the listing method, you must keep the following important remarks in mind.

Remark 1: The set $\{1, 2, 3, 3\}$ is the same as $\{1, 2, 3\}$. That is, while listing the elements in a set we do not gain anything by repeating them. By convention, we do not repeat them.

Remark 2: Consider the sets $\{1, 2, 3\}$ and $\{2, 1, 3\}$. Are they equal or not? You can see that every element of the first set belongs to the second, and vice versa. Therefore, these sets are equal. This example shows that **changing the order in which the elements are listed does not alter the set.**

We would also like to emphasize an observation about the property method.

Remark 3: There can be several properties that define the same set. For example,

$\{x \mid 3x - 1 = 5\}$

$= \{x \mid x \text{ is an even prime number}\}$.

You may like to do the following exercises now.

E3) Describe the following sets by the listing method.

a) { x | x is the smallest prime number }

b) { x | x is a divisor of 12}

c) $\{ x \in \mathbf{Z} \mid x^2 = 4 \}$

d) $\{ x \mid 3x - 5 = 19 \}$

e) the set of all letters in the English alphabet.

E4) Describe the following sets by the property method.

a) {1, 4, 9, 16, ...}

b) {2, 3, 5, 7, 11, 13, 17, ...}

c) {..., −6, −4, −2, 0, 2, 4, 6, ...}

d) ϕ

E5) Give an example of a non-empty set which can be represented only by the property method.

While solving E 3 you would have come across sets consisting of exactly one element. Such a set is called a **singleton**. The singleton with element x is usually written as {x}.

Remark 4: The element x is not the same as the set {x}. In fact, $x \in \{x\}$.

A set which has a finite number of elements is called a **finite set**. By convention, the empty set is considered to be a finite set.

A set which is not finite is an infinite set.

Some examples of infinite sets are **N**, **Q**, **R** and the set of points on a given line.

The following exercise will help you in getting used to the notion of finite and infinite sets.

E6) Which of the following sets are finite, and which are infinite?

a) **Z** ,

b) ϕ ,

c) the solution set of $2x + 5 = 7$,

d) the set of points on the circumference of a circle ,

e) the set of stars in the sky .

The solution set of an equation is the set of solutions of the equation.

Now, given any two real numbers a and b, you know that either $a < b$ or $a = b$ or $a > b$. Is there a similar relationship between sets? Let us see.

1.3 SUBSETS

In this section we shall see what we mean by the terms 'is contained in' and 'contains'.

Consider two sets A and B, where

A = the set of all students of IGNOU, and

B = the set of all female students of IGNOU.

Every female student of IGNOU is a student of IGNOU. So, each element of B is also an element of A. In such a situation we say that B is contained in A.

Of course, IGNOU also has some male students!

So, there is an element x in A such that x does not belong to B. Mathematically, we write this as

$\exists\ x \in A$ such that $x \notin B$.

In this situation we say that B is properly contained in A.

In general, we have the following definitions.

Definitions : A set A is a subset of a set B if every element of A belongs to B, and we denote this fact by $A \subseteq B$.

In this situation, we also say that A is contained in B, or that B contains A, denoted by $B \supseteq A$.

If $A \subseteq B$ and $\exists\ y \in B$ such that $y \notin A$, then we say that A is a **proper subset of B** (or A is properly contained in B). We denote this by $A \subsetneq B$.

If X and Y are two sets such that X has an element x which does not belong to Y, then we say that X **is not contained in** Y. We denote this fact by $X \not\subseteq Y$.

Let us look at a few examples of what we have just defined.

Consider the set $A = \{1, 2, 3\}$. Is $A \subseteq A$? Since every element of A is in A, we find that $A \subseteq A$.

In fact, this is true for any set. In other words, any set is a subset of itself. But note that no set is a proper subset of itself.

Now consider the sets $A = \{1, -1\}$ and $B = \{0, 1, 2\}$.

We find that $\exists\ (-1) \in A$ such that $(-1) \notin B$. $\therefore A \not\subseteq B$.

Similarly, $B \not\subseteq A$.

Note that given any two sets A and B, one and only one of the following possibilities is true.

i) $A \subseteq B$, or

ii) $A \not\subseteq B$.

Using this fact we can prove by contradiction (see the appendix to this block) that

the empty set ϕ is a subset of every set.

$\phi \subseteq A$ for any set A.

To prove this, consider any set A. Suppose $\phi \not\subseteq A$. Then there must be some element in ϕ which is not in A. But this is not possible since ϕ has no elements. Thus, we reach a contradiction. Therefore, what we assumed must be wrong. That is, $\phi \not\subseteq A$ is false. Thus, $\phi \subseteq A$, for any set A.

Try the following exercises now. While doing them remember that to show that $A \subseteq B$, for any sets A and B, you must show that if $a \in A$ then $a \in B$,

i.e., $a \in A \Rightarrow a \in B$.

'$\Rightarrow$' denotes 'implies' (see the block appendix also).

Also, to show that $A \not\subseteq B$, you must show that $\exists\ x \in A$ such that $x \notin B$.

The set of all subsets of a set A is called the power set of A.

E7) Write down all the subsets of $\{1, 2, 3\}$. How many of these contain

a) no element, b) one element, c) two elements, c) three elements?

E8) Show that if $A \subsetneq B$ and $B \subsetneq C$, then $A \subsetneq C$. This shows that '$\subsetneq$' is a transitive relation.

E9) Show that '$\not\subseteq$' is not a transitive relation. For this, you need to find an example of three sets A, B and C such that $A \not\subseteq B$, $B \not\subseteq C$, but $A \subseteq C$.

Now, let us go back for a moment to the point before Remark 1, where we defined equality of sets. Let us see what equality means in terms of subsets. Consider the sets
A = the set of even natural numbers less than 10, and
$B = \{2, 4, 6, 8\}$.

They are equal since every member of A is a member of B, and vice versa. That is, $A \subseteq B$ and $B \subseteq A$.

Thus, $A = B \Leftrightarrow (A \subseteq B$ and $B \subseteq A)$.

Now, for some interesting exercises !

E10) Consider the sets $A = \{x \mid x + 1 = 3\}$, $B = \{1,2\}$ and C = the set of all even numbers that are prime. What is the relationship between

i) A and B ?

ii) A and C ?

E11) If $A = B$ and $B \supseteq C$, what is the relationship between A and C ?

So far you have seen two methods of representing sets. There is yet another way of depicting sets and the relationships between them. This is what we discuss in the next section.

1.4 VENN DIAGRAMS

It is often easier to understand a situation if we can represent it graphically. To ease our understanding of many situations involving sets and their relationships, we represent them by simple diagrams, called **Venn diagrams**. An English logician, John Venn (1834 - 1923), invented them. To be able to draw a Venn diagram, you would need to know what a universal set is.

In any situation involving two or more sets, we first look for a convenient large set which contains all the sets under discussion. We call this large set a **universal set** and denote it by U. Clearly, U is not unique. For example, if we are talking about the set D of women directors, and the set S of women scientists, then we can take our universal set U to be the set of all earning women. This is because U contains D as well as S. But, we can also take U to be the set of all women. This would also serve the same purpose.

A universal set is not unique.

Again, if we wish to work with the sets of integers and rational numbers, we could take the set of real numbers as our universal set. We could also take Q as our universal set, since it contains both Z and Q.

We usually take our universal set to be just large enough to contain all the sets under consideration.

Now, let us see how to draw a Venn diagram. Suppose we are discussing various sets A,B,C, ... We choose our universal set U. So, $A \subseteq U$, $B \subseteq U$, $C \subseteq U$, and so on. We show this situation in a Venn diagram as follows :

The interior of a rectangle represents U. The subsets A, B, C, etc., are represented by the interiors of closed regions lying completely within the rectangle. These regions may be in the form of a circle, ellipse or any other shape. To clarify what we have just said, consider the following example.

Example 1 : Draw a Venn diagram to represent the sets

$U = \{1,2,...,10\}$, $A = \{1,2,3\}$, $B = \{3,4,5\}$, $C = \{6,7\}$.

Solution : See Figure 1.

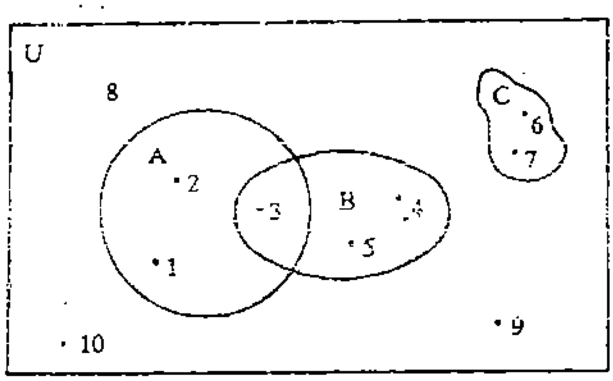


Figure 1 : A Venn diagram

We have denoted A by a circle, B by an ellipse which intersects A, and C as another closed region. The points 8, 9 and 10 don't lie in any of A, B or C. All these regions and points lie in the universal set U, which is represented by the outer rectangle.

Note that 3 belongs to both A and B. Therefore, it lies in the circle as well as the ellipse. Also note that A and C do not have any elements in common. Therefore, the regions representing them do not cut each other. For the same reason the regions representing B and C do not cut each other.

Of course, we could have drawn B and C in the shape of circles also.

Now, consider the following situation : A and B are two sets such that $A \subsetneq B$, that is, A is a proper subset of B. What will a Venn diagram corresponding to this situation look like? Well, we can just take B to be our universal set. Then the Venn diagram in Figure 2 is one such diagram. If we take another set U that properly contains B, as our universal set, then we get the Venn diagram in Figure 3.

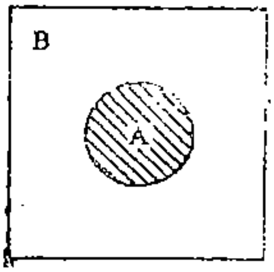


Figure 2

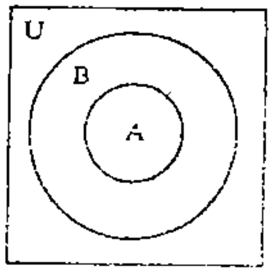


Figure 3

Try this exercise now.

E12) How would you represent the following situation by a Venn diagram?
The set of all rectangles, the set of all squares and the set of all parallelograms.

Now that you are familiar with Venn diagrams, let us discuss the various operations on sets. During the discussion we will be using Venn diagrams off and on.

1.5 OPERATIONS ON SETS

You must be familiar with the basic operations on real numbers—addition, subtraction, multiplication and division. In using these we combine two real numbers at a time, in different ways, to obtain another real number. In a similar way we can obtain new sets by applying certain operations to two given sets at a time. In this section we shall discuss the operations of complementation, intersection and union.

1.5.1 Complementation

Consider the sets N and {0, 1}. There are elements of N that do not belong to {0, 1}, like 2, 3, etc. The set of these elements is the complement of {0, 1} in N, according to the following definition.

$A \setminus B \subseteq A$

Definition: Let A and B be two sets. The complement of A in B, denoted by $B \setminus A$, is the set $\{x \in B \mid x \notin A\}$.

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

If B is the universal set U, then $B \setminus A$ is $U \setminus A$. This set is called the **complement of the set A**, and is denoted by A' or A^c.

The unshaded area in Figure 2 denotes the set $B \setminus A$ (or A', since B = U in this case). This diagram shows us that $x \in A^c$ **if and only if** $x \notin A$.

Note that in the situation of Figure 2, $A \setminus B = \phi$, since every point of A is a point of B.

Try this exercise now.

E13) a) Represent the following sets in a Venn diagram : The set P of all prime numbers, the set Z and the set $Q \setminus Z$.

b) Is the set $Z \setminus P$ finite or infinite?

E14) Let A be any set. What will $A \setminus A$, $\phi \setminus A$, $A \setminus \phi$ and $(A^c)^c$ be ?

Let us now consider another operation on sets, namely, the intersection of two or more sets.

1.5.2 Intersection

Let A and B be two subsets of a universal set U. The intersection of A and B will be the set of points that belong to both A and B. This is denoted by $A \cap B$.

Thus, $A \cap B = \{x \in U \mid x \in A \text{ and } x \in B\}$.

To clarify this idea consider the following example.

Example 2: Let S be the set of prime numbers and T be the set $\{x \in Z \mid x \text{ divides } 10\}$. What is $S \cap T$?

Solution : We take Z to be our universal set. Then $S \cap T$ = set of those integers that are prime numbers as well as divisors of 10 = $\{2, 5\}$.

You should be able to do the following exercise now.

E15) Obtain $Z \cap Q$, $Q \cap Z$, $Z \cap Z$ and $N \cap \phi$.

While solving E 15 you may have noticed certain facts about the operation of intersection. We explicitly list them in the following theorem.

Theorem 1: For any two sets A and B,

a) $A \cap B \subseteq A$

b) $A \cap B \subseteq B$

c) $A \subseteq B \Rightarrow A \cap B = A$

d) $A \cap A = A$

e) $A \cap \phi = \phi$

f) $A \cap B = B \cap A$

(f) says that the operation of intersection of sets is commutative

g) $A \setminus B = A \cap B^c$.

h) if C is a set such that $C \subseteq A$ and $C \subseteq B$, then $C \subseteq A \cap B$.

Proof : We will prove (a) and (b), and leave you to check the rest (see E 16).

So, let $x \in A \cap B$. Then $x \in A$ and $x \in B$. Thus, $A \cap B \subseteq A$ and $A \cap B \subseteq B$. So we have proved (a) and (b).

Now, using (a) and (b), try to do the following exercise.

E16) Prove (c) – (h) of Theorem 1.

E17) Is Theorem 1 (h) true if we replace '$\subseteq$' by '$\subsetneq$' everywhere? Why ?

Now, consider the set Q^- of negative rational numbers and the set Q^+ of positive rational numbers. Then $Q^- \cap Q^+ = \phi$. This shows that Q^- and Q^+ are disjoint sets, a term we now define.

Definition : Let A and B be two sets such that $A \cap B = \phi$. Then we say that they are mutually disjoint (or disjoint).

Now let us represent the intersection of sets by means of Venn diagrams. The shaded region in Figure 4 represents the set $A \cap C$, which is non-empty, as you can see. Also note that the regions representing A and B do not overlap, that is, $A \cap B = \phi$. From this diagram we can also see that neither is $A \subseteq C$, nor is $C \subseteq A$. Further, both $C \setminus A$ and $A \setminus C$ are non-empty sets. See how much information a Venn diagram can convey!

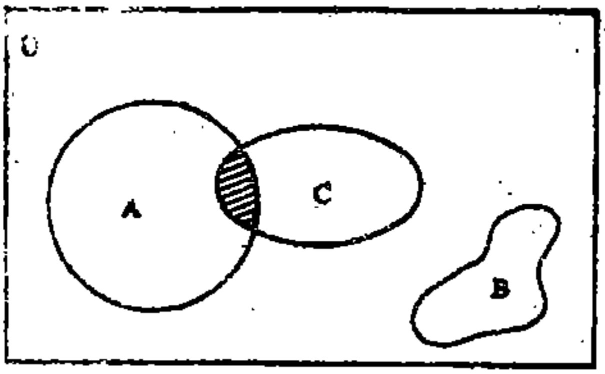


Figure 4

Now, go back to Figure 2 for a moment. What situation does it represent? It shows two sets A and B, with $A \subsetneq B$, that is, A is a proper subset of B. Then the shaded area shows $A \cap B = A$.

Try the following exercise now.

E18) Let $U = \mathbf{Z}$, $A = \{1, 3, 5\}$, B = the set of odd integers. Draw the Venn diagram to represent this situation, and shade the portion $A \cap B$.

So far we have considered the intersection of two sets. Now let us define the intersection of 3, 4 or more sets.

Definition: The **intersection** of n sets $A_1, A_2, \ldots, A_n$ is defined to be the set $\{x \mid x \in A_i \text{ for every } i = 1, \ldots, n\}$.

This is denoted by $A_1 \cap A_2 \cap \ldots \cap A_n$ or $\bigcap_{i=1}^{n} A_i$.

Let us look at an example involving the intersection of 3 sets.

Example 3 : Let A, B and C be the sets of multiples of 3, 6 and 10 in N, respectively. Obtain $A \cap B \cap C$.

Solution : $A \cap B \cap C$ will consist of all those natural numbers that belong to A, B and C. Thus,

$$A \cap B \cap C = \{x \in N \mid 3, 6 \text{ and } 10 \text{ divide } x\}$$
$$= \{x \in N \mid 30 \text{ divides } x\}$$
$$= \{30n \mid n \in N\}.$$

Note that 30 is the lowest common multiple (l.c.m) of 3, 6 and 10.

Try the following exercise now.

E19) Let $A = \{1, 2, 3, 4\}$, $B = \{2, 4, 5, 6\}$ and $C = \{1, 4, 7, 8\}$.
Determine $A \cap B \cap C$. Also verify that

a) $A \cap B \cap C = (A \cap B) \cap C$,

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

E20) If $A = \{6n \mid n \in N\}$ and $B = \{15n \mid n \in N\}$, then define $A \cap B \cap A$.

What you have shown in E 19 is not only true for those sets. It is true for any three sets A, B and C. This property of $\cap$ is called **associativity**. Using this property we can obtain the intersection of any n sets by intersecting any two adjacent sets at a time.

For example, if A, B, C, D are 4 sets, then

$$A \cap B \cap C \cap D = [(A \cap B) \cap C] \cap D$$
$$= A \cap [(B \cap C) \cap D]$$
$$= (A \cap B) \cap (C \cap D).$$

Let us now look at another operation on sets.

1.5.3 Union

'≤' denotes 'less than or equal to' and '≥' denotes 'greater than or equal to'.

Suppose we have two sets $A = \{x \in R \mid x \leq 10\}$ and $B = \{x \in R \mid x \geq 10\}$. Then any element of R belongs to either A or B, because any real number will be either less than or equal to 10 or greater than or equal to 10; and 10 will belong to both A and B. In this case we will say that R is the union of A and B.

In general, we have the following definition.

Definition : Let A and B be two sets. The set of all those elements which belong either to A or to B or to both A and B is called the **union** of A and B. It is symbolically written as $A \cup B$.

Thus

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

Before going further we make a remark.

Remark 5 : Since $A \cup B$ contains all the elements of A as well as B, it follows that

$A \subseteq A \cup B$, $B \subseteq A \cup B$.

In fact, $A \cap B \subseteq A \subseteq A \cup B$ and $A \cap B \subseteq B \subseteq A \cup B$.

Using E 8, this shows that $A \cap B \subseteq A \cup B$.

Now let us look at an example.

Example 4: Find $N \cup Z$.

Solution : $N = \{1, 2, 3, \ldots\}$ and $Z = \{\ldots -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

We want to find

$N \cup Z = \{x \mid x \in N \text{ or } x \in Z\}$.

But $N \subseteq Z$. Thus, $x \in N \Rightarrow x \in Z$. This immediately tells us that $N \cup Z = \{x \mid x \in Z\} = Z$.

'⇔' denotes 'equivalence' or 'implies and is implied by' (see appendix to block 1).

Example 4 is a particular case of the general fact that $A \subseteq B \Leftrightarrow A \cup B = B$.

You can use this fact while solving the following exercises.

E21) For any three sets A, B and C, show that

a) $A \cup A = A$,

b) $A \cup B = B \cup A$, that is, the operation of union is commutative.

c) $A \cup \phi = A$.

d) if $A \subseteq C$ and $B \subseteq C$, then $A \cup B \subseteq C$.

E22) Let U be the real line R, $A = \{x \in R \mid 0 \leq x \leq 1\}$ and $B = \{x \in R \mid 1 \leq x \leq 3\}$. Determine $A \cup B$.

E23) What can you say about A and B if $A \cup B = \phi$?

It is easy to visualise unions of sets by Venn diagrams. Consider Figure 5. In this diagram we see four sets A, B, C and D, and the universal set U. The shaded area represents $A \cup B$.

$C \cup D$ is the area enclosed by both C and D, which is just D, since $C \subseteq D$. Can you find the area in Figure 5 that represents $A \cup B \cup D$? You may be able to, once we have defined the union of 3 or more sets.

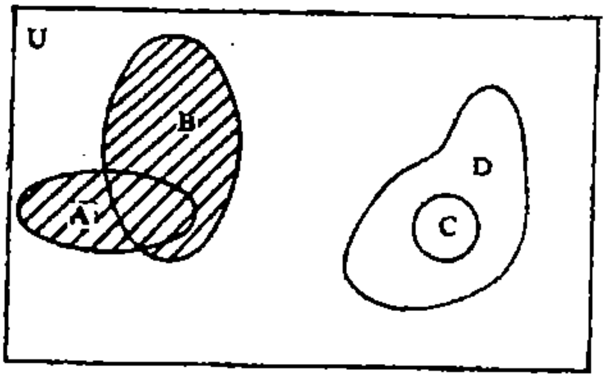


Figure 5

Definition : The union of n sets $A_1, A_2, ..., A_n$ is the set
$\{ x \mid x \in A_i$ for some i such that $1 \le i \le n \}$.

This is denoted by $A_1 \cup A_2 \cup ... \cup A_n$ or $\bigcup_{i=1}^{n} A_i$.

So now you can see that $A \cup B \cup D$ is represented in Figure 5 by the shaded area along with the area enclosed by D.

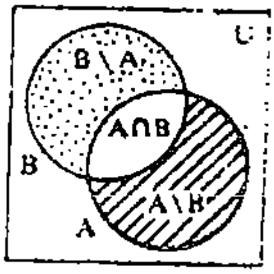
Now, let us consider $A \cup B \cup C$, where $A = \{1, 2, 3\}$, $B = \{2, 3, 4, 5\}$, $C = \{1\}$. Then
$A \cup B \cup C = \{1, 2, 3, 4, 5\}$.

You will find that this is the same as $(A \cup B) \cup C$ as well as $A \cup (B \cup C)$.

You may also like to verify that

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

These statements are particular cases of the general facts that we ask you to prove in the following exercise.



E24) a) For the sets A, B and C in Example 3, show that

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

b) For any two sets A and B, show that $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$.
(We depict this situation in the Venn diagram in Fig.6.)

What you have shown in E 24 (a) is true for any three sets, that is, the operation of union of sets is associatives. Consequently, we can obtain the union of any number of set by taking the union of any two adjacent ones at a time. For example, if A, B, C, D are four sets, then

$$A \cup B \cup C \cup D = [(A \cup B) \cup C] \cup D$$
$$= A \cup [(B \cup C) \cup D]$$
$$= (A \cup B) \cup (C \cup D).$$

By now you must be familiar with the operations that we have discussed in this section. Now let us try and prove some laws that relate them.

1.6 LAWS RELATING OPERATIONS

In this section we shall discuss two sets of laws that relate unions with intersections, Cartesian products with unions and Cartesian products with intersections. We start with the distributive laws.

1.6.1 Distributive Laws

You must be familiar with the law of distributivity that connects the operations of multiplication and addition of real numbers. It is

$a \times (b + c) = a \times b + a \times c \ \forall a, b, c \in \mathbb{R}$.

Whereas we have only one law for numbers, we have two laws of distributivity for sets, which we will now state.

Theorem 2 : Let A, B and C be three sets. Then

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

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

Proof : We will prove (a) and ask you to prove (b) (see E 29)

a) We will show that

$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$ and .... (2)

$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)$. .... (3)

Now, $x \in A \cap (B \cup C)$

$\Rightarrow x \in A$ and $x \in B \cup C$

$\Rightarrow x \in A$ and $(x \in B$ or $x \in C)$

$\Rightarrow (x \in A$ and $x \in B)$ or $(x \in A$ and $x \in C)$

$\Rightarrow x \in A \cap B$ or $x \in A \cap C$

$\Rightarrow x \in (A \cap B) \cup (A \cap C)$

So, we have proved the first inclusion, (2). To prove (3), let

$x \in (A \cap B) \cup (A \cap C)$.

$\Rightarrow x \in A \cap B$ or $x \in A \cap C$

$\Rightarrow (x \in A$ and $x \in B)$ or $(x \in A$ and $x \in C)$

$\Rightarrow x \in A$ and $(x \in B$ or $x \in C)$

$\Rightarrow x \in A$ and $x \in B \cup C$

$\Rightarrow x \in A \cap (B \cup C)$

So we have proved (3).

Note that our argument for proving (3) is just the reverse of our argument for proving (2). In fact, we could have combined the proofs of (2) and (3) as follows :

$x \in A \cap (B \cup C)$

$\Leftrightarrow x \in A$ and $x \in B \cup C$

$\Leftrightarrow x \in A$ and $(x \in B$ or $x \in C)$

$\Leftrightarrow (x \in A$ and $x \in B)$ or $(x \in A$ and $x \in C)$

$\Leftrightarrow x \in A \cap B$ or $x \in A \cap C$

$\Leftrightarrow x \in (A \cap B) \cup (A \cap C)$

This proves (a).

Now try to solve the following exercise, using a two - way implication.

E25) Prove (b) of Theorem 2.

Let us verify Theorem 2 in the following situation.

Example 5 : Verify the distributive laws for the sets N, Q and R in place of A, B and C.

Solution : We first show that $N \cap (Q \cup R) = (N \cap Q) \cup (N \cap R)$.

Now $Q \cup R = R$, since $Q \subseteq R$.

Therefore, $N \cap (Q \cup R) = N \cap R = N$, since $N \subseteq R$.

Also $N \cap Q = N$ and $N \cap R = N$.

Therefore, $(N \cap Q) \cup (N \cap R) = N \cup N = N$.

Thus, $N \cap (Q \cup R) = (N \cap Q) \cup (N \cap R)$.

Now, to verify that $N \cup (Q \cap R) = (N \cup Q) \cap (N \cup R)$, note that both sides are equal to Q. Hence the law holds.

Remark 6 : Theorem 2 (a) says that $\cap$ distributes over $\cup$, and Theorem 2 (b) says that $\cup$ distributes over $\cap$.

Let us now consider another set of laws.

1.6.2 De Morgan's Laws

We will now discuss two laws that relate the operation of finding the complement of a set to that of the intersection or union of sets. These are known as De Morgan's laws, after the British mathematician Augustus De Morgan (1806 - 1871).

Theorem 3 : For any two sets A and B in a universal set U,

a) $(A \cap B)^c = A^c \cup B^c$

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

Proof : As in the case of Theorem 2, we will prove (a), and ask you to prove (b).

a) Let $x \in (A \cap B)^c = U \setminus (A \cap B)$

$\Leftrightarrow x \notin A \cap B$

$\Leftrightarrow x \notin A$ or $x \notin B$ (because, if $x \in A$ and $x \in B$, then $x \in A \cap B$)

$\Leftrightarrow x \in A^c$ or $x \in B^c$

$\Leftrightarrow x \in A^c \cup B^c$

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

Now try the following exercises.

E26) Prove (b) of Theorem 3.

E27) Verify De Morgan's laws for A and B, where $A = \{1, 2\}$, $B = \{2, 3, 4\}$.

(You can take $U = \{1, 2, 3, 4\}$, i.e., $U = A \cup B$. Of course, the laws will continue to hold true with any other U.)

So far we have discussed operations on sets and their inter-relationship. We will now talk of a product of sets, of which the coordinate system is a special case.

1.7 CARTESIAN PRODUCT

An interesting set that can be formed from two given sets is their Cartesian product, named after the French philosopher and mathematician Rene Descartes (1596 – 1650). He also invented the Cartesian coordinate system. Let us see what this product is.

Figure 7: Rene Descartes

Let A and B be two sets. Consider the pair (a, b), in which the first element is from A and the second element is from B. Then (a, b) is called an **ordered pair**. In an ordered pair, the order in which the two elements are written is important. Thus, (a, b) and (b, a) are **different ordered pairs**. We call two ordered pairs (a, b) and (c, d) **equal** (or the same) if $a = c$ and $b = d$.

Using ordered pairs, we give the following definition.

Definition : The Cartesian product A × B, of the sets A and B, is the set of all possible ordered pairs (a, b) where $a \in A$, $b \in B$.

That is, $A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$.

For example, if $A = \{1, 2, 3\}$, $B = \{4, 6\}$, then

$A \times B = \{(1, 4), (1, 6), (2, 4), (2, 6), (3, 4), (3, 6)\}$.

Also note that

$B \times A = \{(4, 1), (4, 2), (4, 3), (6, 1), (6, 2), (6, 3)\}$.

You can see that $(1, 4) \in A \times B$, but $(1, 4) \notin B \times A$.

Therefore, $A \times B \neq B \times A$.

Try these exercises now.

E28) If $A = \{2, 5\}$ and $B = \{2, 3\}$, find $A \times B$, $B \times A$, $A \times A$.

E29) If $A \times B = \{(7, 2), (7, 3), (7, 4), (2, 2), (2, 3), (2, 4)\}$, determine A and B.

Now that we have defined the Cartesian product of two sets, let us extend the definition to any number of sets.

∀ denotes 'for every'.

Definition : Let $A_1, A_2, \ldots, A_n$ be n sets. Then their Cartesian product is the set

$A_1 \times A_2 \times \ldots \times A_n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in A_i \ \forall\, i = 1, 2, \ldots, n\}$.

For example, if **R** is the set of real numbers, then $\mathbf{R} \times \mathbf{R} = \{(a_1, a_2) \mid a_1 \in \mathbf{R}, a_2 \in \mathbf{R}\}$,

$\mathbf{R} \times \mathbf{R} \times \mathbf{R} = \{(a_1, a_2, a_3) \mid a_i \in \mathbf{R}\ \forall\, i = 1, 2, 3\}$, and so on.

It is customary to write $\mathbf{R}^2$ for $\mathbf{R} \times \mathbf{R}$ and $\mathbf{R}^n$ for $\mathbf{R} \times \ldots\ldots \times \mathbf{R}$ (n times).

In your earlier mathematical studies you must have often used the fact that **R** can be geometrically represented by a line. Are you also familiar with a geometrical representation of $\mathbf{R} \times \mathbf{R}$?

You know that every point in a plane has two coordinates, x and y, and every ordered pair (x, y) of real numbers defines the coordinates of a point in the plane. Thus, $\mathbf{R}^2$ is the Cartesian product of the x-axis and the y-axis and hence, $\mathbf{R}^2$ represents a plane. In the same way $\mathbf{R}^3$ represents three-dimensional space, and $\mathbf{R}^n$ represents n-dimensional space, for any $n \geq 1$.

Try the following exercise now.

E30) Which of the following belong to the Cartesian product $\mathbf{Q} \times \mathbf{Z} \times \mathbf{N}$? Why?

a) $(3, 0)$, b) $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$, c) $(1, 1, 1)$, d) $\left(\frac{1}{2}, -5, \sqrt{2}\right)$, e) $\{-2, 2, 3\}$.

E31) Give an example of a proper non-empty subset of $\mathbf{R} \times \mathbf{R}$.

E32) Prove that for any 3 sets A, B and C,

a) $A \times (B \cup C) = (A \times B) \cup (A \times C)$

b) $A \times (B \cap C) = (A \times B) \cap (A \times C)$

distributes over ∪ as well as over ∩.

We will end our discussion on sets here. Let us summarise what we have covered in this unit.

1.8 SUMMARY

In our discussion on sets we have brought out the following points.

1) A set is a well-defined collection of objects.
2) The listing method and property method for representing sets.
3) Given two sets A and B, what we mean by $A \subseteq B$, $A \supseteq B$ and $A = B$.
4) The pictorial representation of sets and their relationships by Venn diagrams, and its utility.
5) The operations of complementation, intersection and union of sets, and their properties.
6) The distributive laws : For any three sets A, B and C,

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

   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
7) De Morgan's laws : For any two sets A and B,

   $(A \cap B)^c = A^c \cup B^c$

   $(A \cup B)^c = A^c \cap B^c$.
8) The Cartesian product of sets.

Now, we suggest that you go back to the objectives given in Sec.1.1, and see if you have achieved them. One way of checking this is to solve all the exercises in the unit. If you would like to know what our solutions are, we have given them in the next section.

1.9 SOLUTIONS/ANSWERS

E1) (b), (c), (d), (e)

E2) (c), (e)

E3) a) {2}

b) {1, 2, 3, 4, 6, 12}

c) {2, – 2}

d) {8}

e) {a, b, c, ..., x, y, z}

E4) a) $\{x^2 \mid x \in Z\}$

b) { x | x is a prime number }

c) { x | x is an even integer}

d) We can have several representations (see Remark 2). For example,
$\phi = \{x \in N \mid x \text{ is both odd as well as even}\}$, or
$\phi = \{x \in N \mid x < 0\}$.

E5) R

E6) (b), (c), and (e) are finite.
(a) and (d) are infinite.

E7) ϕ, {1}, {2}, {3}, {1,2}, {1, 3}, {2, 3}, {1, 2, 3}.
If a finite set A has n elements, then its power set has 2^n elements.

E8) Let $x \in A$. Then
$A \subseteq B \Rightarrow x \in B$. And then
$B \subseteq C \Rightarrow x \in C$
$\therefore x \in A \Rightarrow x \in C$
$\therefore A \subseteq C$

E9) Consider $A = \{0\}$, $B = \{1\}$, $C = \{0\}$.
Then $A \nsubseteq B$, $B \nsubseteq C$, but $A \subseteq C$.

E10) Note that $A = \{2\} = C$, $B = \{1, 2\}$.
$\therefore$ $A = C$ and $A \subsetneq B$.

E11) $A \supsetneq C$

E12)

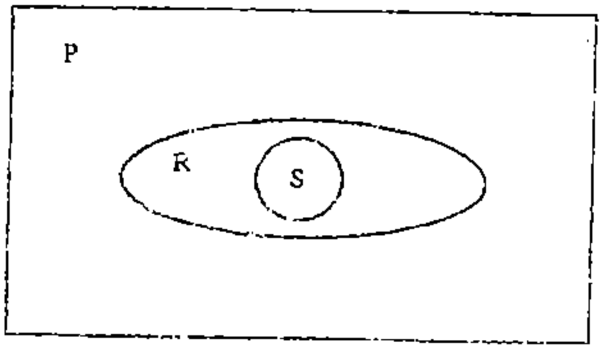


Figure 8

S, R, P are the sets of squares, rectangles and parallelograms, respectively. Here we have taken $U = P$.

E13) a)

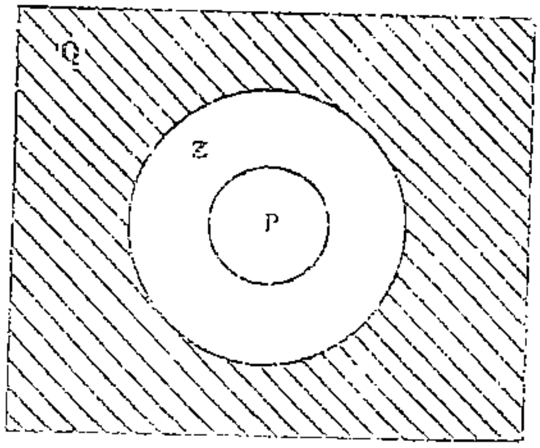


Figure 9

The shaded area is $Q \setminus Z$.

b) Infinite

E14) ϕ, ϕ, A and

$(A^c)^c = A$, since $x \in A \Leftrightarrow x \notin A^c \Leftrightarrow x \in (A^c)^c$.

E15) Z, Z, Z, ϕ.

E16) c) From (a) we know that $A \cap B \subseteq A$. We need to prove that $A \subseteq A \cap B$. For this, let $x \in A$. Then, since $A \subseteq B$, $x \in B$.

Thus, $x \in A \cap B$. $\therefore$ $A \subseteq A \cap B$.

$\therefore$ $A = A \cap B$.

d) $A \cap A \subseteq A$, applying (a).
$A \subseteq A \cap A$, as in proof of (c).

$\therefore$ $A = A \cap A$.

e) $A \cap \phi \subseteq \phi$, applying (b).
Also $\phi \subseteq A \cap \phi$, since ϕ is a subset of every set.
$\therefore A \cap \phi = \phi$.

f) $A \cap B \subseteq B$ and $A \cap B \subseteq A$. $\therefore A \cap B \subseteq B \cap A$.
Similarly, $B \cap A \subseteq A \cap B$.
$\therefore A \cap B = B \cap A$.

g) $x \in A \setminus B \Leftrightarrow x \in A$ and $x \notin B$
$\Leftrightarrow x \in A$ and $x \in B^c$
$\Leftrightarrow x \in A \cap B^c$.
$\therefore A \setminus B = A \cap B^c$.

h) Let $x \in C$. Then $C \subseteq A \Rightarrow x \in A$. Similarly, $x \in B$. Therefore, $x \in A \cap B$. Hence, $C \subseteq A \cap B$.

E17) No. For example, if $A = \{1,2,3\}$, $B = \{1,2,4\}$, $C = \{1,2\}$, then $C \subsetneq A$ and $C \subsetneq B$, but C is not properly contained in $A \cap B$; it is exactly equal to $A \cap B$.

E18)

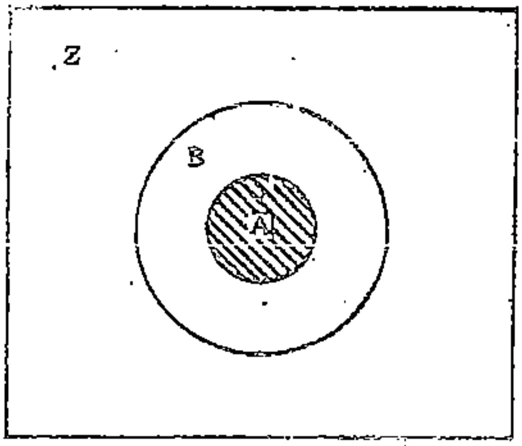


Figure 10

E19) $A \cap B \cap C = \{4\} = (A \cap B) \cap C = A \cap (B \cap C)$.

E20) $A \cap B \cap A = A \cap (B \cap A) = A \cap (A \cap B) = (A \cap A) \cap B$.
$= A \cap B = \{30n \mid n \in N\}$.

E21) a) $A \subseteq A \Rightarrow A \cup A = A$.

b) $x \in A \cup B \Leftrightarrow x \in A$ or $x \in B$
$\Leftrightarrow x \in B$ or $x \in A$.
$\Leftrightarrow x \in B \cup A$.
$\therefore A \cup B = B \cup A$.

c) $\phi \subseteq A \Rightarrow A \cup \phi = A$

d) Let $x \in A \cup B$. Then $x \in A$ or $x \in B$. In either case $x \in C$, since $A \subseteq C$ and $B \subseteq C$. Thus,
$x \in A \cup B \Rightarrow x \in C$.
$\therefore A \cup B \subseteq C$.

E22) $\{x \in R \mid 0 \le x \le 3\}$

E23) Since $A \subseteq A \cup B = \phi$, we see that $A \subseteq \phi$. Also, $\phi \subseteq A$ always.
$\therefore A = \phi$. Similarly, $B = \phi$.

E24) a) $A = \{3x \mid x \in N\}$, $B = \{6x \mid x \in N\}$, $C = \{10x \mid x \in N\}$.
Note that $B \subseteq A$.
$\therefore A \cup B = A$.

$\therefore (A \cup B) \cup C = A \cup C = \{x \in N \mid 3 \text{ divides } x \text{ or } 10 \text{ divides } x\}$
Also, $A \cup B \cup C = \{x \in N \mid 3, 6 \text{ or } 10 \text{ divide } x\}$
$= \{x \in N \mid 3 \text{ or } 10 \text{ divide } x\}$
$= (A \cup B) \cup C$.
Similarly, $A \cup B \cup C = A \cup (B \cup C)$.

b) $A \setminus B \subseteq A \subseteq A \cup B$, $B \setminus A \subseteq B \subseteq A \cup B$, $A \cap B \subseteq A \cup B$.
$\therefore (A \setminus B) \cup (A \cap B) \cup (B \setminus A) \subseteq A \cup B$.
Conversely, let $x \in A \cup B$. Then $x \in A$ or $x \in B$.

Now, there are only three possibilities for x :

i) $x \in A$ but $x \notin B$, that is, $x \in A \setminus B$, or

ii) $x \in A$ and $x \in B$, that is, $x \in A \cap B$, or

iii) $x \in B$ but $x \notin A$, that is, $x \in B \setminus A$.

Thus, $A \cup B \subseteq (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$.

So we have proved the result.

E25) $x \in A \cup (B \cap C)$
$\Leftrightarrow x \in A$ or $x \in B \cap C$
$\Leftrightarrow x \in A$ or ($x \in B$ and $x \in C$)
$\Leftrightarrow$ ($x \in A$ or $x \in B$) and ($x \in A$ or $x \in C$)
$\Leftrightarrow x \in A \cup B$ and $x \in A \cup C$
$\Leftrightarrow x \in (A \cup B) \cap (A \cup C)$.

E26) $x \in (A \cup B)^c \Leftrightarrow x \notin A \cup B$
$\Leftrightarrow x \notin A$ and $x \notin B$
$\Leftrightarrow x \in A^c$ and $x \in B^c$.
$\Leftrightarrow x \in A^c \cap B^c$.

E27) $A \cup B = U$. $\therefore (A \cup B)^c = \phi$
Also $A^c = \{3, 4\}$, $B^c = \{1\}$
$\therefore A^c \cap B^c = \phi$
$\therefore (A \cup B)^c = A^c \cap B^c$
Further $A \cap B = \{2\}$. $\quad \therefore (A \cap B)^c = \{1, 3, 4\}$
$\therefore A^c \cup B^c = (A \cap B)^c$.

E28) $A \times B = \{(2,2), (2,3), (5,2), (5,3)\}$

$B \times A = \{(2,2), (3,2), (2,5), (3,5)\}$

$A \times A = \{(2,2), (2,5), (5,2), (5,5)\}$

E29) $A = \{7, 2\}$, $B = \{2, 3, 4\}$

E30) Only (c). (a) is not, since it is only an ordered pair, and not a triple. (b) is not, since $\frac{1}{2} \notin N \cup Z$. (d) is not, since $\sqrt{2} \notin N$. (e) is not, since it is not an ordered triple; it is a set of three elements.

E31) Any subset is $A \times B$, where $A \subseteq R$, $B \subseteq R$. For a proper subset of R^2, either A or B should be a proper subset of R.

E32) a) $(x, y) \in A \times (B \cup C)$
$\Leftrightarrow x \in A$ and $y \in B \cup C$
$\Leftrightarrow x \in A$ and ($y \in B$ or $y \in C$)
$\Leftrightarrow (x, y) \in A \times B$ or $(x, y) \in A \times C$
$\Leftrightarrow (x, y) \in (A \times B) \cup (A \times C)$.

b) $(x, y) \in A \times (B \cap C)$
$\Leftrightarrow x \in A$ and $y \in B \cap C$
$\Leftrightarrow (x, y) \in A \times B$ and $(x, y) \in A \times C$
$\Leftrightarrow (x, y) \in (A \times B) \cap (A \times C)$.

UNIT 2 COMPLEX NUMBERS

Structure

2.1 INTRODUCTION

In your studies so far you must have dealt with natural numbers, integers, rational numbers and real numbers. You also know that a shortcoming in N led mathematicians of several centuries ago to define negative numbers. Hence, the set Z was born. For similar reasons Z was extended to Q and Q to R at various stages in history. Then came a point when mathematicians looked for solutions of equations like $x^2+1 = 0$. Since $x^2+1 = 0$ has no solution in R, for a long time it was accepted that this equation has no solution. The Indian mathematicians Mahavira (in 850 A.D.) and Bhaskara (in 1150 A.D.) clearly stated that the square root of a negative quantity does not exist. Then, in the 16th century the Italian mathematician Cardano tried to solve the quadratic equation $x^2-10x+40 = 0$. He found that $x_1 = 5+\sqrt{-15}$ and $x_2 = 5-\sqrt{-15}$ satisfied the equation. But then, what is $\sqrt{-15}$? He, and other mathematicians, tried to give this expression some meaning. Even while making mathematical models of real life solutions, the mathematicians of the 17th and 18th centuries were coming across more and more examples of equations which had no real roots. To overcome this shortcoming the concept of a complex number slowly came into being. It was the famous mathematician Gauss (1777-1855) who used and popularised the name 'complex number' for numbers of the type $5+\sqrt{-15}$.

In the early 1800s, a geometric representation of complex numbers was developed. This representation finally made complex numbers acceptable to all mathematicians. Since then complex numbers have seeped into all branches of mathematics. In fact, they have even been necessary for developing several areas in modern physics and engineering.

In this unit we aim to familiarise you with complex numbers and the different ways of representing them. We shall also discuss the basic algebraic operations on complex numbers. Finally, we shall acquaint you with a very useful result, namely, De Moivre's theorem. It has several applications. We shall discuss only two of them in some detail.

We would like to reiterate that whatever mathematics course you study, you will need the knowledge of the subject matter covered in this unit. So, please study it carefully and ensure that you have achieved the following

Objectives

After studying this unit, you should be able to

- define a complex number ;
- describe the geometrical, polar and exponential representations of a complex number ;
- apply the various algebraic operations on complex numbers ;
- prove and use De Moivre's theorem.

2.2 WHAT A COMPLEX NUMBER IS

When you consider the linear equation $2x + 3 = 0$, you know that it has a solution, namely, $x = \frac{-3}{2}$. But, can you always find a real solution of the equation $ax + b = 0$, where $a, b \in \mathbf{R}$ and $a \neq 0$? Is the required solution $x = \frac{-b}{a}$? It is, since $a\left(\frac{-b}{a}\right) + b = 0$.

Now, what happens if we try to look for real solutions of any quadratic equation over **R** ? Consider one such equation, namely, $x^2 + 1 = 0$, that is, $x^2 = -1$. This equation has no solution in **R** since the square of any real number must be non-negative.

From about 250 A.D. onwards, mathematicians have been coming across quadratic equations, arising from real life situations, which did not have any real solutions. It was in the 16th century that the Italian mathematicians Cardano and Bombelli started a serious discussion on extending the number system to include square roots of negative numbers. In the next two hundred years more and more instances were discovered in which the use of square roots of negative numbers helped in finding the solutions of real problems.

In 1777 the Swiss mathematician Euler introduced the "imaginary unit", which he denoted by the Greek letter iota, that is i. He defined $i = \sqrt{-1}$. Soon after, the great mathematician Carl Friedrich Gauss introduced the term complex numbers for numbers such as $1 + i\ (= 1 + \sqrt{-1})$ or $-2 + i\sqrt{5}\ (= -2 + \sqrt{-5})$.

Nowadays these numbers are accepted and used in every field of mathematics.

Let us define a complex number now.

Definition : A complex number is a number of the form $x + iy$, where x and y are real numbers and $i^2 = -1$.

We say that x is the **real part** and y is the **imaginary part** of the complex number $x + iy$. We write $x = \text{Re}\,(x + iy)$ and $y = \text{Im}\,(x + iy)$.

Caution : i) i is not a real number.
ii) Im (x + iy) is the real number y, and not iy.

We denote the set of all complex numbers by **C**.

So, $\mathbf{C} = \{x + iy \mid x, y \in \mathbf{R}\}$.

By convention, we will usually denote an element of **C** by z. So, whenever we will talk of a complex number z, we will mean $z = x + iy$ for some $x, y \in \mathbf{R}$. In fact, $z = \text{Re}\, z + i\, \text{Im}\, z$.

There is another convention that we follow while writing complex numbers, which we give in the following remark.

Remark 1 : When you go through Sec. 2.4.2, you will see that $iy = yi\ \forall\ y \in \mathbf{R}$. That is why we can write the complex number $x + iy$ as $x + yi$ also.

By convention, we write any complex number $x + iy$ for which $y \in \mathbf{Q}$, as $x + yi$. For example, we prefer to write $2 + i$, $2 + \frac{3}{2}i$ and $2 + \frac{5}{9}i$ for $2 + i1$, $2 + i\frac{3}{2}$ and $2 + i\frac{5}{9}$, respectively.

But, if $z \in C$ is of the form $z = a + i\sqrt{b}$, $b \in R$, then we prefer to write z in this form and not as $z = a + \sqrt{b}\,i$.

Now that you know what a complex number is, would you agree that the following numbers belong to **C** ?

$5 + \sqrt{-15}$, $3i$, $\sqrt{2}$, $\sqrt{-2}$

Each of them is a complex number because

$5 + \sqrt{-15} = 5 + i\sqrt{15}$

$3i = 0 + i3$

$\sqrt{2} = \sqrt{2} + i\,0$

$\sqrt{-2} = 0 + i\sqrt{2}$

$\sqrt{-a} = i\sqrt{a}\ \forall\ a \geq 0.$

From these examples you may have realised that some complex numbers can have their real part or their imaginary part equal to zero. We have names for such numbers.

Definition : Consider a complex number $z = x + iy$.

If $y = 0$, we say z is **purely real.**

If $x = 0$, we say z is **purely imaginary.**

We usually write the purely real number $x + 0i$ as x only, and write the purely imaginary number $0 + iy$ as iy only.

Try these exercises now.

E1) Complete the following table :

z	Re z	Im z
$\frac{1+\sqrt{-23}}{2}$		
i		
	0	0
$\frac{-1+\sqrt{3}}{5}$		

E2) Is $R \subseteq C$? Why ?

Now, given any complex number, we can define a related complex number in a very natural way, as follows.

Definition : Let $z = x + iy \in C$. We define the **complex conjugate** (or simply the **conjugate**) of z to be the complex number

$\bar{z} = x - iy.$

Thus, $Re\,\bar{z} = Re\,z$ and $Im\,\bar{z} = -Im\,z$.

For example, if $z = 15 + i$ then $\bar{z} = 15 - i$.

Try this simple exercise now.

E3) Obtain the conjugates of
$2 + 3i$, $2 - 3i$, 2, $3i$.

In Section 2.4.2 you will see one important use of the complex conjugate.

So far we have shown you an algebraic method of representing complex numbers. Now let us consider a geometrical way of doing so.

2.3 GEOMETRICAL REPRESENTATION

You know that we can geometrically represent real numbers on the number line. In fact, there is a one-one correspondence between real numbers and points on the number line. You have also seen that a complex number is determined by two real numbers, namely, its real and imaginary parts. This observation led the mathematicians Wessel and Gauss to think of representing complex numbers as points in a plane. This geometric representation was given in the early 1800s. It is called an **Argand diagram**, after the Swiss mathematician J. R. Argand, who propagated this idea.

Let us see what an Argand diagram is.

Take a rectangular set of axes OX and OY in the XOY plane. Any point in the plane is determined by its Cartesian coordinates. Now we consider any complex number x + iy. We represent it by the point in the plane with Cartesian coordinates (x, y). This representation is an Argand diagram. For example, in Figure 1, P represents the complex number 2 + 3i, whose real part is 2 and imaginary part is 3. And what number does P' represent ? P' corresponds to 2 – 3i.

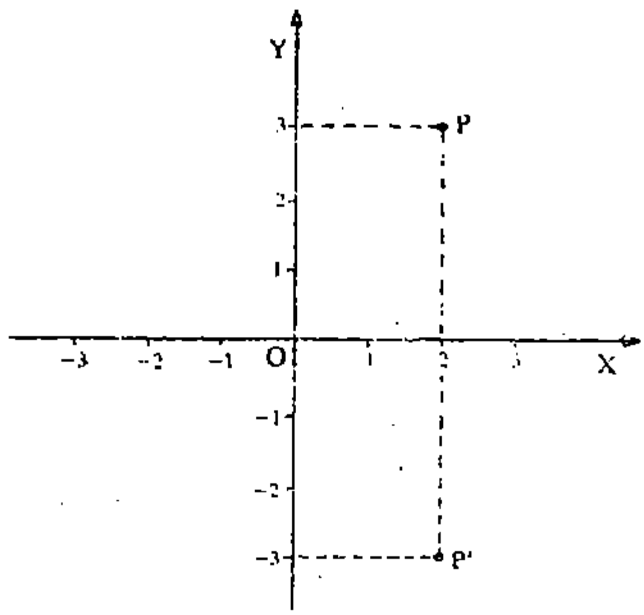


Figure 1 : An Argand diagram

You may have realised that in an Argand diagram the purely real numbers lie along the x-axis and the purely imaginary numbers lie along the y-axis.

So, you have seen that, given $x + iy \in \mathbf{C}$ we associate with it the unique point $(x, y) \in \mathbf{R}^2$. The converse is also true. That is, given $(x, y) \in \mathbf{R}^2$, we can associate with it the unique complex number x + iy. This means that the following definition of a complex number is equivalent to our previous definition.

Definition : A complex number is an ordered pair of real numbers. In the language of sets, we can say that $\mathbf{C} = \mathbf{R} \times \mathbf{R}$.

With the help of this definition can you say when two complex numbers are equal ?

Definition : We say that two complex numbers (x_1, y_1) and (x_2, y_2) are **equal** iff $x_1 = x_2$ and $y_1 = y_2$.

In other words, $x_1 + iy_1 = x_2 + iy_2$ iff $x_1 = x_2$ and $y_1 = y_2$.

Thus, two elements of C are equal iff their real parts are equal and their imaginary parts are equal.

So, for example, $\frac{-1+\sqrt{-3}}{2} = \frac{-1}{2} + i\frac{\sqrt{3}}{2}$, but
$\frac{-1+\sqrt{-3}}{2} = \frac{-1}{2} + i\frac{1}{2}$.

Try these exercises now.

E4) a) Plot the following elements of C in an Argand diagram :
$3, -1+i, \overline{-1+i}, i$

b) Plot the sets $S_1 = \{2 + iy \mid y \in R\}$, $S_2 = \{x + 3i \mid x \in R\}$ and
$S_3 = \{x + ix \mid x \in R\}$ in an Argand diagram.

E5) Write down the elements of C represented by the points $\left(\frac{-1}{2}, \frac{1}{3}\right)$, (2, 0) and (0, −2) in the plane.

E6) For what values of k and m is $k + 3i = \frac{1}{2} + i\,m$?

While solving E4 you may have observed that in an Argand diagram the point that represents $\bar{z}$ is the reflection in the x-axis of the point that represents z, for any $z \in C$.

Here are two more exercises about complex conjugates.

E7) For which $z \in C$ will $z = \bar{z}$?

E8) For any $z \in C$, show that $\bar{\bar{z}} = z$, that is, the conjugate of the conjugate of z is z.

Now consider any non-zero complex number $z = x + iy$. We represent it by P in the Argand diagram in Figure 2. We call the distance OP the modulus of z, and denote it by $|z|$.

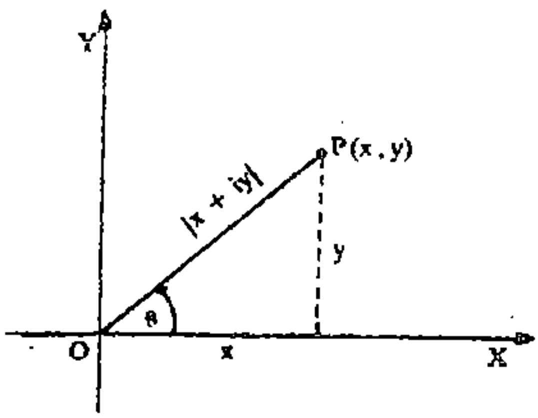


Figure 2: Modulus and argument

Using the Pythagoras theorem, we see that

$|z| = \sqrt{x^2 + y^2}$.

If z is real, what is $|z|$? It is just the absolute value of z.

Here's another important remark on the modulus.

Remark 2 : $z \in \mathbf{C}$, but $|z| \in \mathbf{R}$.

Now, if you see Figure 2 again, you will see that $\angle XOP = \theta$. We call θ **an argument** of $z = x + iy$.

For $z = 0$, $|z| = 0$ and its argument is not defined.

Now, if $z \in \mathbf{C}$, $z \neq 0$, will it have a unique argument? If θ is an argument, so are $2\pi + \theta$, $4\pi + \theta$, etc. If we insist that θ lie in the range $-\pi < \theta \leq \pi$, then we get a unique argument. We call this value of θ **the argument** of z, and denote it by **Arg z**.

If, in Figure 2, we write $|z| = r$. Then you can see that $\sin\theta = \frac{y}{r}$ and $\cos\theta = \frac{x}{r}$.

$\therefore x = r\cos\theta, y = r\sin\theta$. ...... (1)

So, we can also write z as

$z = r(\cos\theta + i\sin\theta)$, where $r = |z|$ and $\theta = \text{Arg } z$.

This is called the **polar form** of z.

Note that, given $z = x + iy$ we can use (1) to obtain $\text{Arg } z = \tan^{-1}\left(\frac{y}{x}\right)$. However, as more than one angle between $-\pi$ and π have the same tan value, we must draw an Argand diagram to find the right value of Arg z.

Let us look at an example.

Example 1 : (a) Obtain the modulus and argument of $1 + i$.

(b) Obtain z, if $|z| = 2$ and $\text{Arg } z = \frac{\pi}{3}$.

Solution : (a) Let $z = 1 + i$.

Then $\text{Re } z = 1$, $\text{Im } z = 1$. Thus, $1 + i$ corresponds to $(1, 1)$, which lies in the first quadrant. We find that

$$|z| = \sqrt{(\text{Re } z)^2 + (\text{Im } z)^2} = \sqrt{1^2 + 1^2} = \sqrt{2}, \text{ and}$$

$$\text{Arg } z = \tan^{-1}\left(\frac{\text{Im } z}{\text{Re } z}\right) = \tan^{-1}(1) = \frac{\pi}{4} \text{ or } \frac{3\pi}{4}.$$

Since z lies in the first quadrant, Arg z must be between 0 and $\frac{\pi}{2}$.

Thus, $\text{Arg } z = \frac{\pi}{4}$.

b) $z = |z|(\cos(\text{Arg } z) + i\sin(\text{Arg } z))$

$$= 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = 2\left(\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)$$

$$= 1 + i\sqrt{3}$$

Try the following exercises now.

E9) Write down the polar forms of the complex numbers listed in E4 (a).

E10) Show that $\{z \in \mathbf{C} \mid |z| = 1\}$ is the equation of the circle $x^2 + y^2 = 1$ in $\mathbf{R}^2$, and vice versa.

There is yet another way of representing a complex number. In fact this method is closely related to the polar representation. It uses the expression e^z, where $z \in \mathbf{C}$. Let us define this expression.

Definition : For any $z = x + iy \in \mathbf{C}$, we define

$e^z = e^x(\cos y + i \sin y)$.

In particular, if $z = iy$, a purely imaginary number, then we get

Euler's formula : $e^{iy} = \cos y + i \sin y \;\forall\; y \in \mathbf{R}$.

This formula is due to the famous Swiss mathematician Leonhard Euler. You will be using it quite often while dealing with complex numbers.

Figure 3 : Euler (1707 - 1783)

Now consider any $z \in \mathbf{C}$. We write it in its polar form,

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

Now, using Euler's formula we find that

$z = re^{i\theta}$.

This is the exponential form of the complex number z.

For example, the exponential form of $z = \frac{3\sqrt{3}}{2} + \frac{3i}{2}$

is $3e^{i\pi/6}$, since $|z| = 3$ and $\text{Arg } z = \frac{\pi}{6}$.

Try this exercise now.

E11) Write the following complex numbers in polar form and exponential form :

$$\sqrt{\frac{5}{2}} + i\sqrt{\frac{5}{2}}, 1 + i, -1, i.$$

By now you must be thoroughly familiar with the various ways of representing a complex number. Let us now discuss some operations on complex numbers.

2.4 ALGEBRAIC OPERATIONS

In this section we will discuss the addition, subtraction, multiplication and division of complex numbers. Let us first consider '+' and '−' in **C**.

2.4.1 Addition And Subtraction

We will now define addition in **C** using the definition of addition in **R**.

Definition : The sum of two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ is the complex number $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$, that is

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

Let us look at an example.

Example 2 : Find the sum of

i) $3 + i$ and $-2 + 4i$,

ii) -5 and $5 - i$.

Solution : i) $(3 + i) + (-2 + 4i) = (3 + (-2)) + i(1 + 4)$

$= 1 + 5i$

ii) $(-5)+(5-i) = (-5+0i)+(5-i) = (-5+5)+i(0-1)$
$= 0 + i(-1)$
$= -i.$

Have you observed that any complex number is the sum of a purely real and a purely imaginary number ? This is because $x + iy = (x + 0i) + (0 + iy)$.

In the following exercises we ask you to verify some very important properties of addition in **C**.

E12) a) Find the sum of $2 + 3i$ and $\overline{2+3i}$.

b) Show that $z + \bar{z} = 2\text{Re } z$ for any $z \in \mathbf{C}$.

E13) Show that $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \ \forall \ z_1, z_2 \in \mathbf{C}$.

E14) a) Show that $z_1 + z_2 = z_2 + z_1$ for any $z_1, z_2 \in \mathbf{C}$.

b) Show that $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ for any $z_1, z_2, z_3 \in \mathbf{C}$.

+ on C is commutative and associative.

E15) Find an element $a + ib \in \mathbf{C}$ such that

$z + (a + ib) = z \ \forall \ z \in \mathbf{C}$.

If you have solved these exercises, you must have realised that the addition in C satisfies most of the properties that addition in R satisfies. Also, because of what you proved in E15, we say that $0 + i0 \ (= 0)$ is the additive identity in **C**.

Now, can you define subtraction in C? We give you a very natural definition.

Definition : The difference $z_1 - z_2$ of two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ is $z_1 + (-z_2)$,

where $-z_2 = (-x_2) + i(-y_2)$.

Thus, $z_1 - z_2 = z_1 + (-z_2)$

$= (x_1 + iy_1) + [(-x_2) + i(-y_2)]$

$= (x_1 - x_2) + i(y_1 - y_2)$.

So, what do you think $z - z$ is, for any $z \in \mathbf{C}$?

Let's see. Take $z = x + iy$. Then

$z - z = (x - x) + i(y - y) = 0$, the additive identity in **C**.

z ∈ C, (−z) is the additive inverse of z.

Try the following exercises now.

E16) Find $(-6 + 3i) - (-3 - 2i)$.

E17) Find $z - \bar{z}$, for any $z \in \mathbf{C}$.

E18) Find the relationship between

a) $|z|$ and $|-z|$,

b) Arg z and Arg $(-z)$,

for any $z \in \mathbf{C}$. (See Figure 4.)

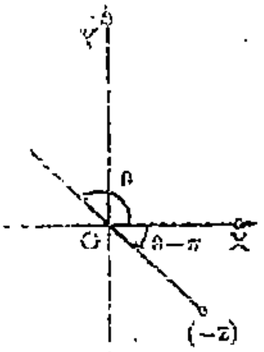


Figure 4: z and −z

We will now make a brief remark on the graphical representation of the sum of complex numbers.

Remark 3: The addition of two complex numbers has an important geometrical representation. Consider an Argand diagram (Figure 5) in which we represent two complex numbers (x_1, y_1) and (x_2, y_2) by the points P and Q.

If we complete the parallelogram whose adjacent sides are OP and OQ, the fourth vertex R is the sum $(x_1, y_1) + (x_2, y_2)$.

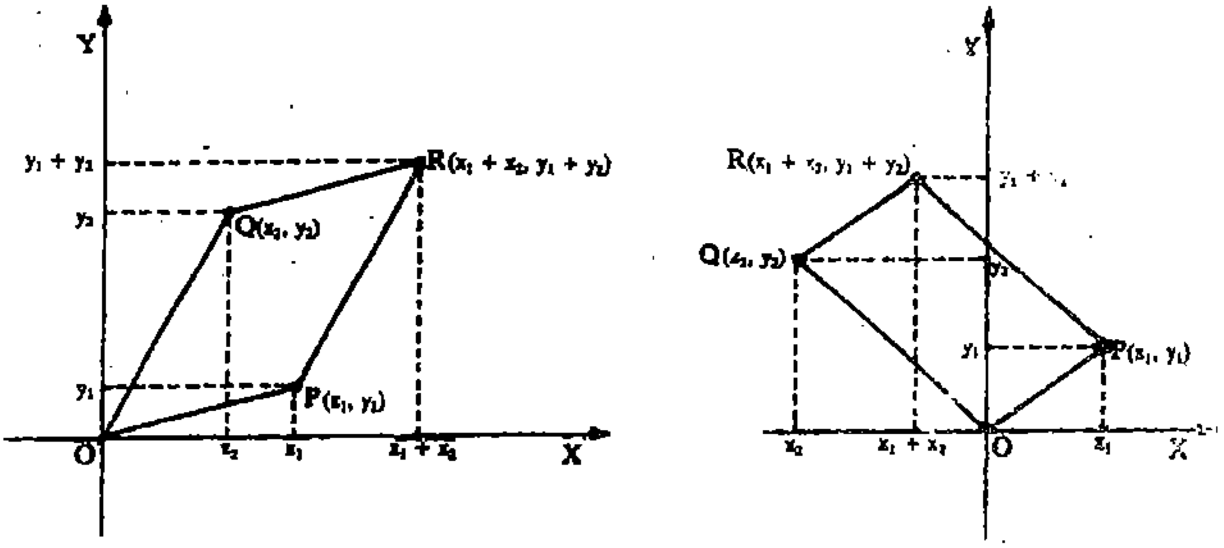


Figure 5: Addition in C

In vector algebra you will come across a similar parallelogram law of addition.

So far you have seen how naturally we have defined addition (and subtraction) in C by using addition (and subtraction) in R. Let us see if we can do the same for multiplication.

2.4.2 Multiplication And Division

We will now use multiplication in R to define multiplication in C. But the route is slightly circuitous. Consider the following product of two linear polynomials $a + bx$ and $c + dx$, where $a, b, c, d \in R$.

$$(a + bx)(c + dx) = ac + (ad + bc)x + bd\,x^2.$$

Now, if we put $x = i$ in this, we get

$$(a + ib)(c + id) = (ac - bd) + i(ad + bc), \text{ since } i^2 = -1.$$

This is the way we shall define a product in C.

Definition: The product $z_1 z_2$ of two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ is the complex number

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Or, in the language of ordered pairs,

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

For example,

$$(1, 2)(-3, 2) = (1.(-3) - 2.2, 1.2 + (-3)2)$$

$$= (-7, -4).$$

Let us check and see what i^2 is according to this definition.

$i^2 = i.i = (0 + i)(0 + i) = (0 - 1) + i(0 - 0) = -1$, which is as it should be.

Multiplication has several properties, which you will discover if you try the following exercises.

E19) Obtain $(x, y)(1, 0)$, $(x, y)(0, 1)$, $(x, y)(0, 0)$, $(x, 0)(y, 0)$ and $(x, y)(1, 1)$ " $(x, y) \in C$.

Multiplication in C is commutative and associative.

E20) Prove that

a) $z_1 z_2 = z_2 z_1 \ \forall z_1, z_2 \in \mathbb{C}$.

b) $(z_1 z_2) z_3 = z_1 (z_2 z_3) \ \forall z_1, z_2, z_3 \in \mathbb{C}$.

c) $(x, y)\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right) = (1, 0) \ \forall (x, y) \in \mathbb{C} \setminus \{0\}$

(Note that $x^2 + y^2 \neq 0$, since $(x, y) \neq (0, 0)$.)

d) $z\bar{z} = |z|^2 \ \forall z \in \mathbb{C}$.

If you've solved these exercises, you must have realised that $z \cdot 1 = z \ \forall z \in \mathbb{C}$.

This means that 1 is the **multiplicative identity** of C. E19 also says that

$z \cdot 0 = 0 \ \forall z \in \mathbb{C}$,

$i(x + iy) = -y + ix \ \forall x, y \in \mathbb{R}$,

and that in case z_1 and z_2 are purely real numbers, our definition of multiplication coincides with the usual one for R.

Also, from E20 (a) you can see that multiplication is commutative, and from E20 (b) that multiplication is associative.

And, what does E20 (c) say? It says that for any non-zero element z of C, $\exists z' \in \mathbb{C}$, such that $zz' = 1$. In this case we say that z' is the **multiplicative inverse** of z. So $z' = \frac{1}{z}$.

Using E20 let us see how to obtain the standard form of the quotient of a complex number by a non-zero complex number. We will use a process similar to the one you must have used for rationalising the denominator in expressions like $\frac{a + b\sqrt{3}}{c + d\sqrt{2}}$. Consider an example.

Example 3 : Obtain $\frac{2 + 3i}{1 - i}$ in the form $a + ib$, $a, b \in \mathbb{R}$.

Solution : E20 (d) gives us a clue to a method for making the denominator a real number. Let us multiply and divide $\frac{2 + 3i}{1 - i}$ by $\overline{1 - i}$. What do we get ?

$z\bar{z} \in \mathbb{R} \ \forall z \in \mathbb{C}$.

$$\left(\frac{2 + 3i}{1 - i}\right)\left(\frac{1 + i}{1 + i}\right) = \frac{(2 + 3i)(1 + i)}{(1 - i)(1 + i)} = \frac{-1 + 5i}{1 + 1} = \frac{-1}{2} + \frac{5}{2}i$$

So, $\frac{2 + 3i}{1 - i} = \frac{-1}{2} + \frac{5}{2}i$.

If you've understood the way we have solved the example, you will have no problem in doing the following exercises.

E21) Obtain $\frac{-2 + i}{\sqrt{-3} + i\sqrt{-4}}$

E22) For $a, b, c, d \in \mathbb{R}$ and $c^2 + d^2 \neq 0$, write $\frac{a + ib}{c + id}$ as an element of C.

E23) a) Show that $\frac{1}{z} = \frac{1}{|z|^2}\bar{z} \ \forall z \in \mathbb{C} \setminus \{0\}$,

b) Show that $\left|\frac{1}{z}\right| = \frac{1}{|z|} \ \forall z \in \mathbb{C} \setminus \{0\}$.

If you have done E22, then you know how to write the quotient of one complex number by a non-zero complex number in standard form.

Now, in Section 2.3 we introduced you to the polar form of a complex number. This form is very handy when it comes to multiplying or dividing complex numbers. Let us see why.

Suppose we know $z_1, z_2 \in \mathbb{C}$ in their polar forms, say,
$z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$. Then

$$z_1 z_2 = r_1 r_2 (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2)$$

$$= r_1 r_2 \{(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)\}$$

$$= r_1 r_2 (\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)) \quad \text{..... (2)}$$

So, $|z_1 z_2| = r_1 r_2 = |z_1||z_2|$ and
$\text{Arg}(z_1 z_2) = (\theta_1 + \theta_2) + 2k\pi$, where we choose $k \in \mathbb{Z}$ so that
$-\pi < (\theta_1 + \theta_2) + 2k\pi \le \pi$.

In Figure 6 we give a graphic illustration of what we have just said.

$\cos(A + B)$
$= \cos A \cos B - \sin A \sin B$
$\sin(A + B) =$
$\sin A \cos B + \cos A \sin B$

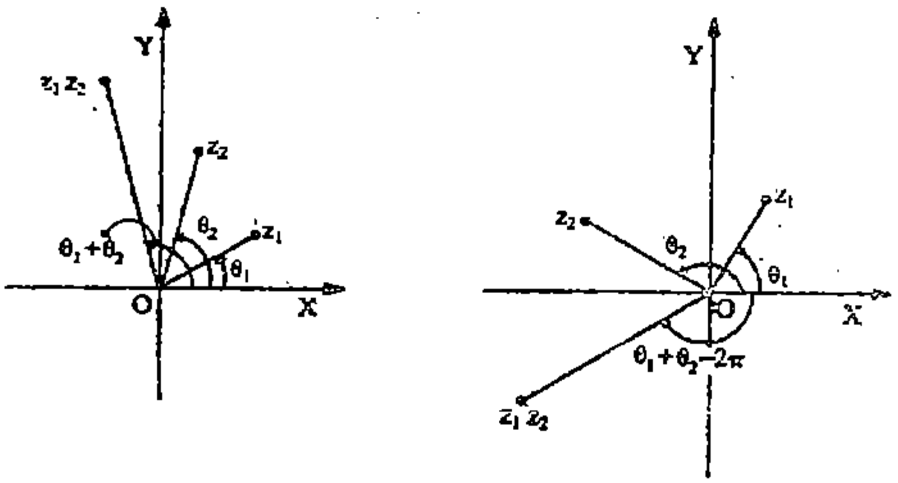


Figure 6 : Product in polar form

Let us consider an example.

Example 4 : Obtain the product of $z_1 = 2(\cos 1 + i\sin 1)$ and $z_2 = \cos 3 + i\sin 3$.

Solution : Here $|z_1| = 2$, $\text{Arg}\, z_1 = 1$, $|z_2| = 1$, $\text{Arg}\, z_2 = 3$.

Therefore, $z_1 z_2 = 2(\cos(1 + 3) + i\sin(1 + 3))$

$= 2(\cos 4 + i\sin 4)$.

Note that $\text{Arg}(z_1 z_2) \neq 4$, since $4 > \pi$. We need to choose an integer k such that
$-\pi < 4 + 2k\pi \le \pi$, $k = -1$ serves the purpose. Thus,

$\text{Arg}(z_1 z_2) = 4 - 2\pi$.

Hence $(z_1 z_2) = 2(\cos(4 - 2\pi) + i\sin(4 - 2\pi))$ is the polar form of $z_1 z_2$.

We have a very nice method of finding the multiplicative inverse of a non-zero complex number in an Argand diagram. Let us see what it is.

Let $z \in \mathbb{C} \setminus \{0\}$ be represented by a point P (see Figure 7). Let Q represent the real number $|z|^2$. Let R be the reflection of P in the x- axis, so that R represents $\bar{z}$.

Now, through (1,0) draw a line parallel to QR. Let it intersect the line OR in S. Then S represents $\frac{1}{z}$.

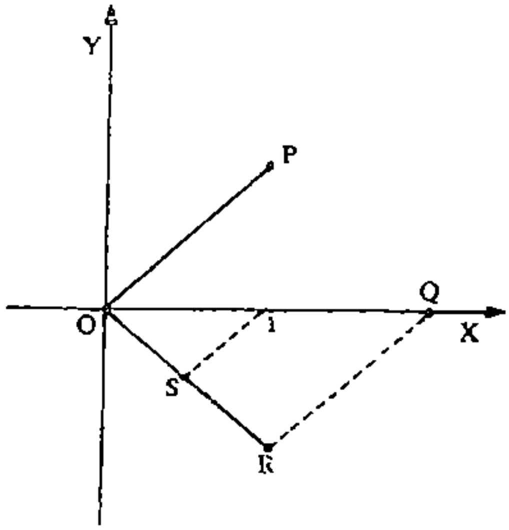


Figure 7 : Finding the multiplicative inverse

Try the following exercises now.

E24) Find the polar forms of z_1 and z_2 where $z_1 = -6$ and $z_2 = 1 + i$. Hence obtain $|z_1 z_2|$ and $\text{Arg}(z_1 z_2)$.

E25) Knowing the polar forms of $z_1, z_2 \in C, z_2 \neq 0$, obtain the polar form of $\frac{z_1}{z_2}$.

E26) Obtain $\frac{z_1}{z_2}$ in the polar form, where z_1 and z_2 are as in E24. Represent $z_1, z_2, \bar{z}_2, \frac{1}{z_2}$ and $\frac{z_1}{z_2}$ in an Argand diagram.

We will use multiplication and division in the polar form a great deal in the next section. Before going to it, let us give you a rule that relates '+' and '×' in C. Do you know of such a law in R ? You must have used the distributive law often enough. It says that $a(b + c) = ab + ac \; \forall \; a, b, c \in R$. The same law holds for C. Why don't you try and prove it ?

E27) a) Check that

$(1 + i) \quad \{(\sqrt{2} - 3i) + (5 + i)\}$

$= (1 + i)(\sqrt{2} - 3i) + (1 + i)(5 + i)$

b) Prove that $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3 \; \forall \; z_1, z_2, z_3 \in C$.

Now let us discuss a very useful theorem.

2.5 DE MOIVRE'S THEOREM

In the previous section, we proved that if

$z_1 = r_1(\cos \theta_1 + i \sin \theta_1)$ and $z_2 = r_2(\cos \theta_2 + i \sin \theta_2)$,

then $z_1 z_2 = r_1 r_2 \{\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)\}$.

In particular, if $z_1 = z_2$, then $r_1 = r_2$, $\theta_1 = \theta_2$, and we find that

$z_1^2 = r_1^2 (\cos 2\theta_1 + i \sin 2\theta_1)$.

In fact, this is a particular case of a very nice formula, namely, that if $z = r(\cos\theta + i\sin\theta)$, then $z^n = r^n(\cos n\theta + i\sin n\theta)$ for any integer n. To prove this result we need **De Moivre's theorem, named after the French mathematician Abraham De Moivre (1667-1754).** It may amuse you to know that De Moivre never explicitly stated this result. But he seems to have known it and used it in his writings of 1730. It was Euler who explicitly stated and proved this result in 1748.

Theorem 2 (De Moivre's theorem) : $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$, for any $n \in \mathbb{Z}$ and any angle θ.

Proof : Let us first prove it for $n > 0$. We will prove this by using the following important principle.

Principle of Induction : Let P(n) be a statement about a positive integer n, such that

i) P(1) is true, and

ii) if P(m) is true for some $m \in \mathbb{N}$, then P(m + 1) is true.

Then, P(n) is true $\forall\, n \in \mathbb{N}$.

You can study induction in greater detail in the course "Abstract Algebra".

How will we use this principle ? For any, $n \in \mathbb{N}$, we will take P(n) to be the statement "$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$". We will first prove that it holds for $n = 1$, that is, P(1) is true. Then, we will assume that it holds for $n = m$ for some $m \in \mathbb{N}$, and prove that it is true for $n = m + 1$. This will show that if P(m) is true, then so is P(m + 1).

Now, for $n = 1$,

$(\cos\theta + i\sin\theta)^1 = \cos\theta + i\sin\theta = \cos 1 \cdot \theta + i\sin 1 \cdot \theta$.

So the result is true for $n = 1$.

Assume that it is true for $n = m$, that is,

$(\cos\theta + i\sin\theta)^m = \cos m\theta + i\sin m\theta$. .... (3)

Now $(\cos\theta + i\sin\theta)^{m+1}$

$= (\cos\theta + i\sin\theta)^m (\cos\theta + i\sin\theta)$

$= (\cos m\theta + i\sin m\theta)(\cos\theta + i\sin\theta)$, by (3)

$= \cos(m\theta + \theta) + i\sin(m\theta + \theta)$, by the formula (2) for products.

$= \cos(m + 1)\theta + i\sin(m + 1)\theta$.

Hence, the result is true for $n = m + 1$.

Thus, by the principle of induction, the result is true $\forall\, n \in \mathbb{N}$.

Now let us see what happens if $n = 0$.

We define $z^0 = 1$, for any $z \in \mathbb{C} \setminus \{0\}$. (As in the case of $\mathbb{R}$, 0^0 is not defined.)

Therefore, $(\cos\theta + i\sin\theta)^0 = 1$.

Also, $\cos 0 \cdot \theta + i\sin 0 \cdot \theta = \cos 0 + i\sin 0 = 1$.

Thus, the result is also true for $n = 0$.

Now, what happens if $n < 0$? Maybe, you can prove this case. You can do the following exercises, which will lead you to the result.

E28) Prove that $\dfrac{1}{\cos\theta + i\sin\theta} = \cos\theta - i\sin\theta$, for any angle θ.

E29) Let $n < 0$, say $n = -m$, where $m > 0$. Then

$$(\cos\theta + i\sin\theta)^n = \frac{1}{(\cos\theta + i\sin\theta)^m}.$$

Use this fact and De Moivre's theorem for positive integers to prove that $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$.

So, De Moivre's theorem is true $\forall\, n \in \mathbb{Z}$.

Now, if $z = r(\cos\theta + i\sin\theta) \in \mathbb{C}$, then $\forall\, n \in \mathbb{Z}$

$z^n = r^n(\cos\theta + i\sin\theta)^n$

$= r^n(\cos n\theta + i\sin n\theta)$, using De Moivre's theorem.

What we have shown is that

$$[r(\cos\theta + i\sin\theta)]^n = r^n(\cos n\theta + i\sin n\theta) \text{ for } r \in \mathbb{R},\ \theta \in \mathbb{R},\ n \in \mathbb{Z}.$$

This result has several applications in mathematics and physics. We shall discuss two of them here.

2.5.1 Trigonometric Identities

One of the most useful applications of Theorem 1 is in proving identities that involve trigonometric functions like $\sin\theta$, $\cos\theta$, *etc.* Let us look at an example.

Example 5 : Find a formula for $\cos 4\theta$ in terms of $\cos\theta$ and $\sin\theta$.

Solution : By De Moivre's theorem

$$(\cos\theta + i\sin\theta)^4 = \cos 4\theta + i\sin 4\theta \qquad \text{.........} (4)$$

We can also expand the left hand side of (4) by using the binomial expansion. Then

$${}^nC_m = \frac{n!}{m!\,(n-m)!}\ \forall\, n, m \in \mathbb{N}.$$

$$(\cos\theta + i\sin\theta)^4 = (\cos\theta)^4 + {}^4C_1(\cos\theta)^3(i\sin\theta) + {}^4C_2(\cos\theta)^2(i\sin\theta)^2 + {}^4C_3\cos\theta(i\sin\theta)^3 + (i\sin\theta)^4$$

$$= \cos^4\theta + 4i\sin\theta\cos^3\theta - 6\sin^2\theta\cos^2\theta - 4i\sin^3\theta\cos\theta + \sin^4\theta \qquad \text{.........} (5)$$

Thus, comparing the real parts in (4) and (5), we get
$\cos 4\theta = \cos^4\theta - 6\sin^2\theta\cos^2\theta + \sin^4\theta$.

You can try the following exercise on similar lines.

E30) Find formulae for $\cos 3\theta$ in terms of $\cos\theta$ and $\sin 3\theta$ in terms of $\sin\theta$.

Now, for any $m \in \mathbb{N}$ let us look at z^m, where $z \in \mathbb{C}$ such that $|z| = 1$. Then, by De Moivre's theorem

$z^m = \cos m\theta + i\sin m\theta$

and $z^{-m} = \cos m\theta - i\sin m\theta$, since $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$ for any angle θ.

$$\left.\begin{aligned} \text{Thus } z^m + z^{-m} &= 2\cos m\theta, \text{ and} \\ z^m - z^{-m} &= 2i\sin m\theta. \end{aligned}\right] \qquad \text{.........} (6)$$

We can use these relations to express $\cos^m\theta$ and $\sin^m\theta$ in terms of $\cos m\theta$ and $\sin m\theta$ for $m = \pm1, \pm2, \ldots\ldots$ Let us consider an example.

Example 6 : Expand $2^{4n-2}(\cos^{4n}\theta + \sin^{4n}\theta)$ in terms of the cosines or sines of multiples of θ.

Solution: Putting $m = 1$ in the equations (6), we get

$2\cos\theta = z + \frac{1}{z}$ and $2i\sin\theta = z - \frac{1}{z}$

$\therefore 2^{4n}\cos^{4n}\theta = \left(z + \frac{1}{z}\right)^{4n}$

$= z^{4n} + 4n\, z^{4n-1}\frac{1}{z} + {}^{4n}C_2 z^{4n-2}\frac{1}{z^2} + \dots + {}^{4n}C_{2n} z^{2n}\frac{1}{z^{2n}} + \dots + 4n\, z\frac{1}{z^{4n-1}} + \frac{1}{z^{4n}}$,

by the binomial expansion.

$= \left(z^{4n} + \frac{1}{z^{4n}}\right) + 4n\left(z^{4n-2} + \frac{1}{z^{4n-2}}\right) + \dots + {}^{4n}C_{2n}$

Also, $2^{4n}\sin^{4n}\theta = \left(z - \frac{1}{z}\right)^{4n}$, since $i^{4n} = (i^4)^n = 1$.

$= \left(z^{4n} + \frac{1}{z^{4n}}\right) - 4n\left(z^{4n-2} + \frac{1}{z^{4n-2}}\right) + \dots + {}^{4n}C_{2n}$

$\therefore 2^{4n}\left(\cos^{4n}\theta + \sin^{4n}\theta\right) = 2\left(z^{4n} + \frac{1}{z^{4n}}\right) + 2\left({}^{4n}C_2\right)\left(z^{4n-4} + \frac{1}{z^{4n-4}}\right) + \dots + 2\left({}^{4n}C_{2n}\right)$

$= 2\{2\cos 4n\theta + 2\left({}^{4n}C_2\right)\cos(4n-4)\theta + \dots\} + 2\left({}^{4n}C_{2n}\right)$, using (6).

$\therefore 2^{4n-2}(\cos^{4n}\theta + \sin^{4n}\theta) = \cos 4n\theta + {}^{4n}C_2\cos(4n-4)\theta + \dots + \frac{1}{2}\,{}^{4n}C_{2n}$.

The procedure we have shown in Example 6 is very useful for solving differential equations involving trigonometric functions. It is also useful for finding the Laplace transform of such functions.

Why don't you try this exercise now ?

E31) Apply De Moivre's formula to prove that

i) $\cos 2\theta = \cos^2\theta - \sin^2\theta$

ii) $\sin 2\theta = 2\sin\theta\cos\theta$

E32) Expand $\cos^6\theta - \sin^6\theta$ in terms of the cosines of multiples of θ.

Let us now look at another area in which we can apply De Moivre's theorem with great success.

2.5.2 Roots Of A Complex Number

In Section 2.2 we told you that the whole subject of complex numbers first arose in an attempt to find the square roots of -1. By now you know that we can always find two distinct complex square roots of any non-zero real number.

That is, given $a \in \mathbb{R}\setminus\{0\}$, $\exists$ distinct $z_1, z_2 \in \mathbb{C}$ such that $z_1^2 = a$, $z_2^2 = a$.

In fact, the set of complex numbers has a much stronger property, which is a major reason for its importance in mathematics. This property is :

given any $n \in \mathbb{N}$ and $z \in \mathbb{C}$, $z \neq 0$, we can find distinct $z_1, \dots, z_n \in \mathbb{C}$ such that $z_k^n = z$ $\forall k = 1, \dots, n$.

Each of these z_k's is called an nth root of z.

To extract all the nth roots of a complex number, we need De Moivre's theorem as well as the following result that we ask you to prove.

E33) Let x be a positive real number and $n \in N$. Show that there is one and only one positive real number b such that $b^n = x$.

(Hint : Let $r, s > 0$ be such that $r^n = x = s^n$. Suppose $r \neq s$. Then $r^n - s^n = 0$ and $r - s \neq 0$. Then you should reach a contradiction.)

We denote the unique positive nth root obtained in E33 by $x^{1/n}$.

Now let us consider an example of extraction of roots of a complex number.

Example 7 : Obtain all the fifth roots of i in C.

Solution : Let $z = r(\cos\theta + i\sin\theta)$ be any 5th root of i. Then $z^5 = i$. The polar form of i is $i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$. Therefore

$z^5 = i$.

$\Rightarrow r^5(\cos\theta + i\sin\theta)^5 = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$

$$\Rightarrow r^5(\cos 5\theta + i\sin 5\theta) = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} \quad \text{......... (7)}$$

by De Moivre's theorem.

Comparing the moduli (plural of 'modulus') and arguments of the complex numbers on both sides of (7), we get

$r^5 = 1$ and $5\theta = \frac{\pi}{2} + 2k\pi$, where $k = 0, \pm1, \pm2, \ldots$.

r is the unique positive real fifth root of 1 (see E33). Since $1 \in R$, $r = 1$, that is $|z| = 1$. The possible values of θ are

$\theta = \frac{1}{5}\left(\frac{\pi}{2} + 2k\pi\right), k = 0, \pm1, \pm2, \ldots$.

Thus, the possible 5th roots of i are

$z = \cos\left(\frac{\pi}{10} + 2k\frac{\pi}{5}\right) + i\sin\left(\frac{\pi}{10} + 2k\frac{\pi}{5}\right), k = 0, \pm1, \pm2, \ldots$

From this it seems that i has infinitely many 5th roots, one for each $k \in Z$. But this is not true. There are only 5 distinct ones among these. They will be the values of z for $k = -2, -1, 0, 1, 2$. Let us see why.

When $k = -2$, $z = \cos\left(\frac{\pi}{10} - \frac{4\pi}{5}\right) + i\sin\left(\frac{\pi}{10} - \frac{4\pi}{5}\right)$

$= \cos\frac{7\pi}{10} - i\sin\frac{7\pi}{10} = z_{-2}$, say.

When $k = -1$, $z = \cos\frac{3\pi}{10} - i\sin\frac{3\pi}{10} = z_{-1}$, say.

When $k = 0$, $z = \cos\frac{\pi}{10} + i\sin\frac{\pi}{10} = z_0$, say.

When $k = 1$, $z = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = z_1$, say.

When $k = 2$, $z = \cos\frac{9\pi}{10} + i\sin\frac{9\pi}{10} = z_2$, say.

When $k = 3$, $z = \cos\frac{13\pi}{10} + i\sin\frac{13\pi}{10} = \cos\left(2\pi - \frac{7\pi}{10}\right) + i\sin\left(2\pi - \frac{7\pi}{10}\right) = z_{-2}$.

When $k = 4$, $z = \cos\frac{17\pi}{10} + i\sin\frac{17\pi}{10} = \cos\left(2\pi - \frac{3\pi}{10}\right) + i\sin\left(2\pi - \frac{3\pi}{10}\right) = z_{-1}$.

Similarly, when $k = 5$, you will get z_0, and so on.

Thus, $k = 5, 6, 7, \ldots$ don't give us new values of z.

Now, if we put $k = -3$, we get $z = \cos\left(\frac{-11\pi}{10}\right) + i\sin\left(\frac{-11\pi}{10}\right) = z_2$.

Similarly, $k = -4, -5, \ldots$ will not give us new values of z.

Therefore, the only 5th roots of i are

$$\cos\left(\frac{\pi}{10} + 2k\frac{\pi}{5}\right) + i\sin\left(\frac{\pi}{10} + 2k\frac{\pi}{5}\right) \text{ for } k = 0, \pm 1, \pm 2.$$

Remark 4 : We also get the 5th roots of i by taking $k = 0, 1, 2, 3, 4$ in

$\cos\left(\frac{\pi}{10} + \frac{2k\pi}{5}\right) + i\sin\left(\frac{\pi}{10} + \frac{2k\pi}{5}\right)$, as you have seen. Only note that for $k = 3$ and $k = 4$, the angles θ will not lie in the range $-\pi < \theta \leq \pi$. That's why we had taken $k = 0, \pm 1, \pm 2$.

Now, look at all the fifth roots of i. How are their moduli related ? They have the same modulus, namely, $|i|^{1/5} (=1)$. Thus, they all lie on the circle with centre (0,0) and radius 1. These points will be equally spaced on the circle, since the arguments of consecutive points differ by $\frac{2\pi}{5}$, a constant. We plot them in the Argand diagram in Figure 8.

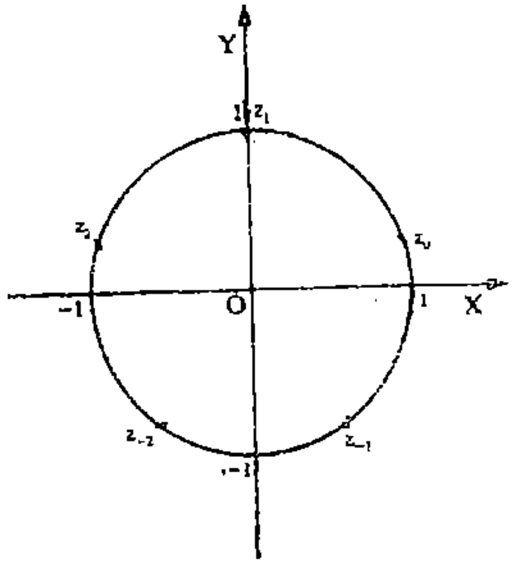


Figure 8 : The fifth roots of i

Using the same procedure as above we can obtain the distinct nth roots of any non-zero complex number, for any $n \in \mathbb{N}$. Thus given, any non-zero complex number w, we write it in its polar form

$w = a(\cos\alpha + i\sin\alpha)$, where $a = |w|$ and $\alpha = \text{Arg } w$.

By E33, there is a unique $r \in \mathbb{R}$, $r > 0$, such that $r^n = a$, that is, $r = a^{1/n}$. Then the distinct nth roots of w are

[illegible]

Geometrically, they lie on a circle of radius [illegible] spaced on it.

Note that

> a non-zero complex number has exactly n distinct nth roots for any $n \in N$. If z is one root, then the others are $z\alpha_1, z\alpha_2, \ldots, z\alpha_{n-1}$, where $\alpha_1, \ldots, \alpha_{n-1}$ are the nth roots of unity.

Now you can do some exercises.

E34) Find the complex cube roots of unity, that is, those $z \in C$ such that $z^3 = 1$. Plot them in an Argand diagram.

E35) Solve the equation $z^4 - 4z^2 + 4 - 2i = 0$.

(**Hint**: The equation can be rewritten as $(z^2-2)^2 = (1+i)^2$.)

The cube roots of unity that you obtained in E34 are very important. We usually denote the cube root $\frac{-1+i\sqrt{3}}{2}$ by the Greek letter ω (omega).

Note that $\omega^2 = \left(\frac{-1+i\sqrt{3}}{2}\right)^2 = \frac{-1-i\sqrt{3}}{2}$, the other non-real cube root of unity. Thus,

> the three cube roots of unity are $1, \omega, \omega^2$, where $\omega = \frac{-1+i\sqrt{3}}{2}$.

Also note that

$$1 + \omega + \omega^2 = 0. \qquad \ldots\ldots (8)$$

We will often use ω and (8) in Unit 3.

We will equally often use the following results, that we ask you to prove.

E36) a) Let $a \in R$. Show that a has a real cube root r, and the cube roots of a are $r, r\omega, r\omega^2$.

b) Show that if $a \in R$, $a < 0$ and n is an even positive integer, then a will not have a real nth root.

c) Let $z \in C \setminus R$. Then show that z has three cube roots, and if any one of them is γ, the other two are $\gamma\omega, \gamma\omega^2$.

With this we come to the end of our discussion on complex numbers. This does't mean that you won't be dealing with them any more. In fact, you will often use whatever we have covered in this unit, while studying this course as well as other mathematics courses.

Let us take a brief look at the points covered in this unit.

2.6 SUMMARY

In this unit on complex numbers you have studied the following points.

1) The definition of a complex number:

A complex number is a number of the form $x + iy$, where $x, y \in R$ and $i = \sqrt{-1}$. Equivalently, it is a pair $(x, y) \in R \times R$.

2) x is the real part and y is the imaginary part of $x + iy$.

3) $x_1 + iy_1 = x_2 + iy_2$ iff $x_1 = x_2$ and $y_1 = y_2$.

4) The conjugate of $z = x + iy$ is $\overline{z} = x - iy$.

5) The geometric representation of complex numbers in Argand diagrams.

6) The polar form of $z = x + iy$ is $z = r(\cos\theta + i\sin\theta)$, where $r = |z| = \sqrt{x^2 + y^2}$ and $\theta = \text{Arg}\, z = \tan^{-1}\left(\frac{y}{x}\right)$, where we choose the θ that corresponds to the position of z in an Argand diagram.

7) Euler's formula : $e^{i\theta} = \cos\theta + i\sin\theta \ \forall\ \theta \in \mathbf{R}$.

8) The exponential form of $z = x + iy$ is $z = re^{i\theta}$, where $r = |z|$ and $\theta = \text{Arg}\, z$.

9) Addition, subtraction, multiplication and division in $\mathbf{C}$: $\forall\ a, b, c, d \in \mathbf{R}$

$(a + ib) \pm (c + id) = (a \pm c) + i(b \pm d)$,

$(a + ib).(c + id) = (ac - bd) + i(ad + bc)$,

$\frac{1}{a + ib} = \frac{a}{a^2 + b^2} - \frac{ib}{a^2 + b^2}$, when $a + ib = 0$,

$\frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{c^2 + d^2}$, for $c + id = 0$.

10) For $z_1, z_2 \in \mathbf{C}$,

$|z_1 z_2| = |z_1||z_2|$, $\text{Arg}(z_1 z_2) = \text{Arg}\, z_1 + \text{Arg}\, z_2 + 2k\pi$,

$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$, $\text{Arg}\left(\frac{z_1}{z_2}\right) = \text{Arg}\, z_1 - \text{Arg}\, z_2 + 2m\pi$, (for $z_2 = 0$)

where $k, m \in \mathbf{Z}$ are chosen so that

$-\pi < \text{Arg}(z_1 z_2) \leq x$ and $-\pi < \text{Arg}\left(\frac{z_1}{z_2}\right) \leq \pi$.

11) De Moivre's theorem : $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta \ \forall\ n \in \mathbf{Z}$ and any angle θ.

12) The use of De Moivre's theorem in proving trigonometric identities and for obtaining nth roots of complex numbers, where $n \in \mathbf{N}$.

13) The cube roots of unity are $1, \omega, \omega^2$, where $\omega = \frac{-1 + i\sqrt{3}}{2}$.

Now that you have gone through this unit, please go back to the objectives listed in Section 2.1. Do you think you have achieved them? One way of finding out is to solve all the exercises that we have given you in this unit. If you would like to verify your solutions or answers, you can see what we have written in the following section.

2.7 SOLUTIONS / ANSWERS

E1)

z	Re z	Im z
$\frac{1 + \sqrt{-23}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{23}}{2}$
i	0	1
0	0	0
$\frac{-1 + \sqrt{3}}{5}$	$\frac{-1 + \sqrt{3}}{5}$	0

E2) Yes, because every real number x is the complex number x + 0i.

E3) 2 – 3i, 2 + 3i, 2, –3i.

E4) a)

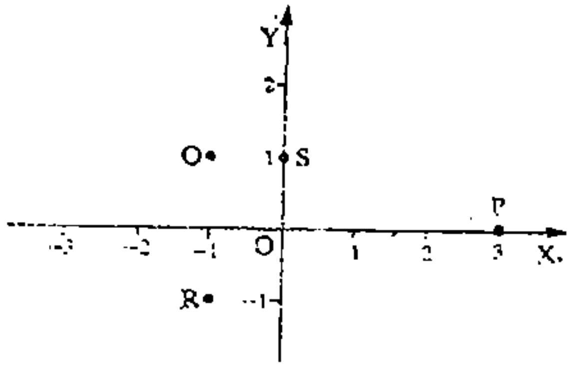


Figure 9

P, Q, R and S represent 3, – 1 + i, $\overline{-1+i}$ and i, respectively.

b)

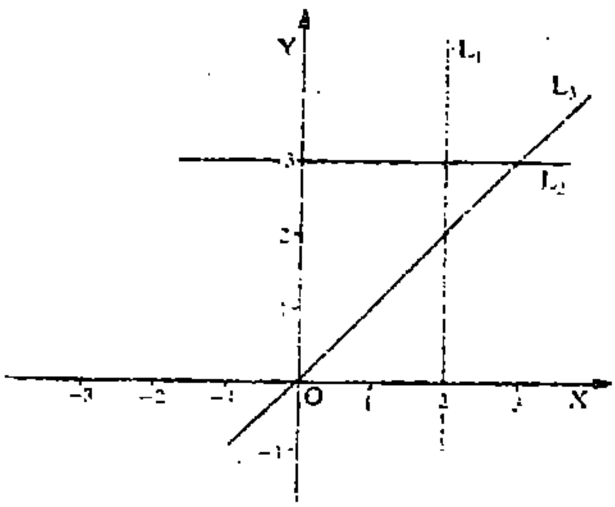


Figure 10

L_1, L_2 and L_3 represent the sets S_1, S_2 and S_3, respectively.

E5) $\frac{-1}{2}+\frac{i}{3}$, 2, – 2i.

E6) $k = \frac{1}{2}, m = 3$.

E7) Let $z = x + iy$. Then $\bar{z} = x - iy$.

$\therefore z = \bar{z} \Rightarrow x + iy = x - iy \Rightarrow y = -y \Rightarrow y = 0$.

$\therefore \forall z \in \mathbb{R}, z = \bar{z}$.

E8) Let $z = x + iy$. Then $\bar{z} = x - iy$.

$\therefore \bar{\bar{z}} = \overline{x - iy} = x + iy = z$.

E9) $3 = 3(\cos 0 + i \sin 0)$

Now $|-1+i| = \sqrt{1+1} = \sqrt{2}$, and

$\text{Arg}(-1+i) = \tan^{-1}(-1) = -\pi/4$ or $3\pi/4$.

Since $-1+i$ corresponds to $(-1,1)$, which lies in the 2nd quadrant, $\text{Arg}(-1+i) = \frac{3\pi}{4}$.

$$\therefore (-1+i) = \sqrt{2}\left(\cos\left(\frac{3\pi}{4}\right) + i\sin\left(\frac{3\pi}{4}\right)\right).$$

$$\overline{-1+i} = -1-i = \sqrt{2}\left(\cos\left(\frac{-3\pi}{4}\right) + i\sin\left(\frac{-3\pi}{4}\right)\right).$$

$$= \sqrt{2}\left(\cos\left(\frac{3\pi}{4}\right) - i\sin\left(\frac{3\pi}{4}\right)\right)$$

$$i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}.$$

E10) For any $z = x + iy \in \mathbb{C}$, $|z| = 1 \Leftrightarrow \sqrt{x^2+y^2} = 1 \Leftrightarrow x^2 + y^2 = 1$.

E11) $\sqrt{\frac{5}{2}} + i\sqrt{\frac{5}{2}} = \sqrt{5}\,(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$ (polar form)

$= \sqrt{5}\, e^{i\pi/4}$ (exponential form)

$1 + i = \sqrt{2}\,(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$ (polar form)

$= \sqrt{2}\, e^{i\pi/4}$ (exponential form)

$-1 = \cos\pi + i\sin\pi$ (polar form)

$= e^{i\pi}$ (exponential form)

$i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$ (polar form)

$= e^{i\pi/2}$ (exponential form)

E12) a) $2+3i+\overline{2+3i} = 2+3i+2-3i = 4+0i = 4$.

b) Let $z = x+iy$. Then

$z+\bar{z} = (x+iy)+(x-iy) = 2x = 2\text{Re}\, z$.

E13) Let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$. Then

$z_1 + z_2 = (x_1 + x_2, y_1 + y_2)$

$\therefore \overline{z_1+z_2} = (x_1 + x_2, -(y_1 + y_2))$

$= (x_1 + x_2, -y_1 - y_2)$

$= (x_1, -y_1) + (x_2, -y_2)$

$= \bar{z}_1 + \bar{z}_2$.

E14) a) Let $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$.

Then $z_1 + z_2 = (x_1, y_1) + (x_2, y_2)$

$= (x_1 + x_2, y_1 + y_2)$

$= (x_2 + x_1, y_2 + y_1)$, since $a+b = b+a\ \forall a, b \in \mathbb{R}$.

$= (x_2, y_2) + (x_1, y_1)$

$= z_2 + z_1$

b) Let $z_1 = (x_1, y_1), z_2 = (x_2, y_2), z_3 = (x_3, y_3)$.

Then, use the fact that $(a+b)+c = a+(b+c)\ \forall a, b, c \in \mathbb{R}$, to prove the result.

E15) Let $z = x + iy$.

Then $z + (a + ib) = z$

$\Rightarrow (x+iy) + (a+ib) = x+iy$

$\Rightarrow (x+a) + i(y+b) = x+iy$

$\Rightarrow x + a = x$ and $y + b = y$

$\Rightarrow a = 0, b = 0$

$\therefore a + ib = 0 + i0 = 0.$

E16) $(-6-(-3)) + i(3-(-2)) = -3+5i.$

E17) Let $z = x + iy$. Then

$z - \bar{z} = (x+iy) - (x-iy) = (x-x) + i(y+y) = 2iy.$
$= i(2\,\text{Im}\, z).$

E18) Let $z = x + iy$. Then $-z = (-x) + i(-y)$. Thus,

a) $|z| = \sqrt{x^2+y^2}$, and

$|-z| = \sqrt{(-x)^2+(-y)^2} = \sqrt{x^2+y^2} = |z|.$

b) $\text{Arg}\, z = \tan^{-1}\left(\frac{y}{x}\right).$

$\text{Arg}(-z) = \tan^{-1}\left(\frac{-y}{-x}\right) = \tan^{-1}\left(\frac{y}{x}\right) = \text{Arg}\, z - \pi$, because $(-z)$ is the reflection of z in the origin.

E19) $(x, y)(1, 0) = (x, y)$

$(x, y)(0, 1) = (-y, x)$

$(x, y)(0, 0) = (0, 0)$

$(x, 0)(y, 0) = (xy, 0)$

$(x, y)(1, 1) = (x-y, x+y).$

$z.1 = z \;\forall\, z \in \mathbb{C}$
$z.0 = 0 \;\forall\, z \in \mathbb{C}.$

E20) a) Let $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$. Then

$z_1 z_2 = (x_1, y_1)(x_2, y_2)$
$= (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)$
$= (x_2 x_1 - y_2 y_1, x_2 y_1 + x_1 y_2)$, since $ab = ba \;\forall\, a, b, c \in \mathbb{R}$
$= (x_2, y_2)(x_1, y_1)$
$= z_2 z_1.$

b) If $z_1 = (x_1, y_1), z_2 = (x_2, y_2), z_3 = (x_3, y_3)$, then

$z_1 z_2 = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$, and

$z_2 z_3 = (x_2 x_3 - y_2y_3, x_2 y_3 + x_3y_2)$

Therefore, $(z_1 z_2) z_3$

$= ((x_1x_2 - y_1y_2)x_3 - (x_1y_2 + x_2y_1)y_3, (x_1x_2 - y_1y_2)y_3 + x_3(x_1y_2 + x_2y_1))$
$= (x_1(x_2x_3 - y_2y_3) - y_1(x_2y_3 + x_3y_2), x_1(x_2y_3 + x_3y_2) + (x_2x_3 - y_2y_3)y_1)$
$= z_1(z_2 z_3)$

c) $(x, y)\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right) = \left(\frac{x^2+y^2}{x^2+y^2}, \frac{-xy+xy}{x^2+y^2}\right) = (1, 0).$

d) Let $z = x + iy \in \mathbb{C}$. Then

$z\bar{z} = (x+iy)(x-iy) = x^2+y^2 = \left(\sqrt{x^2+y^2}\right)^2 = |z|^2$

E21) $\frac{-2+i}{i\sqrt{3}+i(2i)} = \frac{-2+i}{-2+i\sqrt{3}}$, since $i^2 = -1$.

$= \frac{(-2+i)(-2-i\sqrt{3})}{(-2)^2+(\sqrt{3})^2}$

$= \frac{4+\sqrt{3}}{7} + \frac{2}{7}(\sqrt{3}-1)\,i$

E22) $c^2+d^2 \neq 0$ means that $c \neq 0$ or $d \neq 0$. Thus, $c + id \neq 0$. Hence $\frac{a+ib}{c+id}$ is meaningful.

$$\frac{a+ib}{c+id} = \frac{(a+ib)(c-id)}{(c+id)(c-id)} = \frac{(ac+bd)+i(bc-ad)}{c^2+d^2} = \left(\frac{ac+bd}{c^2+d^2}\right) + i\left(\frac{bc-ad}{c^2+d^2}\right).$$

E23) a) Let $z = x + iy \neq 0$. Then, from E20(d) we know that $z\bar{z} = |z|^2$.

Therefore, $z\left(\frac{1}{|z|^2}\bar{z}\right) = 1$. Thus, $\frac{1}{|z|^2}\bar{z}$ is the multiplicative inverse of z.

b) For $z \neq 0, z \cdot \frac{1}{z} = 1. \therefore |z| \cdot \left|\frac{1}{z}\right| = |1| = 1. \therefore \left|\frac{1}{z}\right| = \frac{1}{|z|}$.

E24) $z_1 = 6(\cos\pi + i\sin\pi), z_2 = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$

$\therefore |z_1z_2| = 6\sqrt{2}$ and

$\text{Arg}(z_1z_2) = \left(\pi + \frac{\pi}{4}\right) + 2k\pi$, where $k \in \mathbb{Z}$ such that $-\pi < \text{Arg}(z_1z_2) \le \pi$.

$\therefore \text{Arg}(z_1z_2) = \frac{-3\pi}{4}$.

E25) If $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$,

then $\frac{z_1}{z_2} = \frac{r_1(\cos\theta_1 + i\sin\theta_1)}{r_2(\cos\theta_2 + i\sin\theta_2)}$

$= \frac{r_1}{r_2}(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 - i\sin\theta_2)$, multiplying and dividing by $(\cos\theta_2 - i\sin\theta_2)$.

$= \frac{r_1}{r_2}(\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2))$

$= \frac{r_1}{r_2}(\cos(\theta_1 - \theta_2 + 2k\pi) + i\sin(\theta_1 - \theta_2 + 2k\pi))$, where $k \in \mathbb{R}$ such that $-\pi < \theta_1 - \theta_2 + 2k\pi \le \pi$.

E26) $\frac{z_1}{z_2} = \frac{6}{\sqrt{2}}\left(\cos\left(\pi - \frac{\pi}{4}\right) + i\sin\left(\pi - \frac{\pi}{4}\right)\right)$

$= \frac{6}{\sqrt{2}}\left(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}\right)$

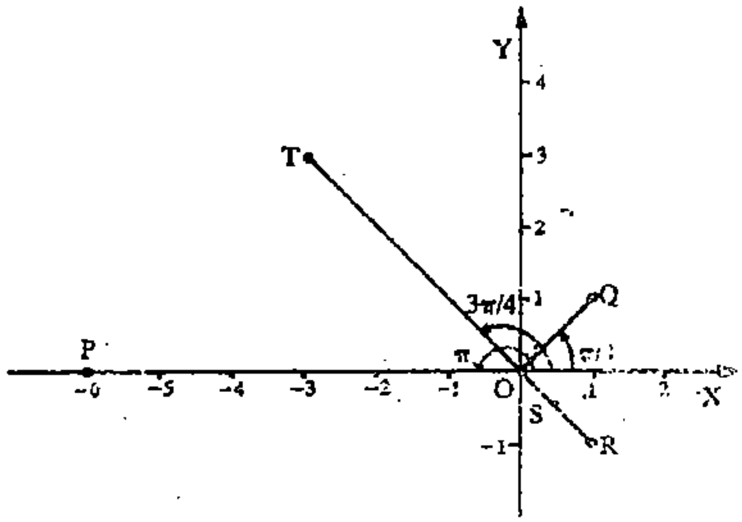


Figure 11

The points P, Q, R, S and T in Figure 11 represent $z_1, z_2, \bar{z}_2, \frac{1}{z_2}$ and $\frac{z_1}{z_2}$, respectively.

Here $OT = \frac{OP}{OQ}$ and $\angle XOT = \angle XOP - \angle XOQ$.

E27) a) LHS $= (1 + i)[(\sqrt{2} + 5) - 2i] = (7 + \sqrt{2}) + i(3 + \sqrt{2})$

RHS $= [(\sqrt{2} + 3) + i(\sqrt{2} - 3)] + (4 + 6i) = (\sqrt{2} + 7) + i(\sqrt{2} + 3)$.

Thus, LHS = RHS.

LHS stands for left hand side and RHS stands for right hand side.

b) Let $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, $z_3 = x_3 + iy_3$.

Then $z_1(z_2 + z_3) = (x_1 + iy_1)[(x_2 + x_3) + i(y_2 + y_3)]$

$= [x_1(x_2 + x_3) - y_1(y_2 + y_3)] + i[x_1(y_2 + y_3) + y_1(x_2 + x_3)]$

$= (x_1x_2 - y_1y_2) + (x_1x_3 - y_1y_3) + i[(x_1y_2 + x_2y_1) + (x_1y_3 + x_3y_1)]$

$= [(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)] + [(x_1x_3 - y_1y_3) + i(x_1y_3 + x_3y_1)]$

$= z_1z_2 + z_1z_3$.

You can also solve this by writing z_1, z_2 and z_3 in polar form. If you do, you must remember to be careful about $z_i = 0$ for any i.

E28) $\dfrac{1}{\cos\theta + i\sin\theta} = \dfrac{\cos\theta - i\sin\theta}{\cos^2\theta + \sin^2\theta} = \cos\theta - i\sin\theta$.

E29) For $n < 0$, say $n = -m$, $m > 0$,

$$(\cos\theta - i\sin\theta)^n = \frac{1}{(\cos\theta + i\sin\theta)^m} = \left(\frac{1}{\cos\theta + i\sin\theta}\right)^m$$

$= (\cos\theta - i\sin\theta)^m$

$= [\cos(-\theta) + i\sin(-\theta)]^m$

$= \cos(-m\theta) + i\sin(-m\theta)$, since $m > 0$.

$= \cos n\theta + i\sin n\theta$.

E30) $(\cos\theta + i\sin\theta)^3 = \cos 3\theta + i\sin 3\theta$.

Also, $(\cos\theta + i\sin\theta)^3 = \cos^3\theta + 3\cos^2\theta(i\sin\theta) + 3\cos\theta(i\sin\theta)^2 + (i\sin\theta)^3$

$= (\cos^3\theta - 3\cos\theta\sin^2\theta) + i(3\sin\theta\cos^2\theta - \sin^3\theta)$.

Thus, comparing real parts of the two equalities, we get

$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta = \cos^3\theta - 3\cos\theta(1 - \cos^2\theta)$

$= 4\cos^3\theta - 3\cos\theta$.

Similarly, comparing the imaginary parts we get

$\sin 3\theta = 3\sin\theta(1 - \sin^2\theta) - \sin^3\theta = 3\sin\theta - 4\sin^3\theta$.

E31) $(\cos\theta + i\sin\theta)^2 = \cos 2\theta + i\sin 2\theta$, and

$(\cos\theta + i\sin\theta)^2 = \cos^2\theta + 2i\cos\theta\sin\theta - \sin^2\theta$.

$\therefore \cos 2\theta = \cos^2\theta - \sin^2\theta$, and

$\sin 2\theta = 2\sin\theta\cos\theta$.

E32) Let $z = \cos\theta + i\sin\theta$. Then, using (6),

$$(2\cos\theta)^6 = \left(z + \frac{1}{z}\right)^6 = \left(z^6 + \frac{1}{z^6}\right) + 6\left(z^4 + \frac{1}{z^4}\right) + 15\left(z^2 + \frac{1}{z^2}\right) + 20,$$

$$(2i\sin\theta)^6 = \left(z^6 + \frac{1}{z^6}\right) - 6\left(z^4 + \frac{1}{z^4}\right) + 15\left(z^2 + \frac{1}{z^2}\right) - 20.$$

$$\therefore 2^6\left(\cos^6\theta - \sin^6\theta\right) = 2\left(z^6 + \frac{1}{z^6}\right) + 30\left(z^2 + \frac{1}{z^2}\right)$$

$$= 4\cos 6\theta + 60\cos 2\theta$$

$$\Rightarrow \cos^6\theta - \sin^6\theta = \frac{1}{16}(\cos 6\theta + 15\cos 2\theta)$$

E33) Let, $r, s \in \mathbb{R}$, $r, s > 0$ and $r^n = x = s^n$. We will prove the result by contradiction (see appendix to this block). Suppose $r \neq s$. Then

$r^n - s^n = (r - s)(r^{n-1} + r^{n-2}s + \ldots + rs^{n-2} + s^{n-1}) = 0$.

Since $r > 0$, $s > 0$, $r^{n-1} + r^{n-2}s + \ldots + rs^{n-2} + s^{n-1} > 0$.

Also, $r - s \neq 0$.

But then how can the product of two non-zero numbers be zero? So we reach a contradiction. Therefore, our assumption must be false. Thus, $r = a$.

E34) Let $z = r(\cos\theta + i\sin\theta)$ be a cube root of $1 = \cos 0 + i \sin 0$.

Then $r = 1^{1/3} = 1$, $\theta = \frac{0 + 2k\pi}{3} = \frac{2k\pi}{3}$ for $k = 0, 1, -1$.

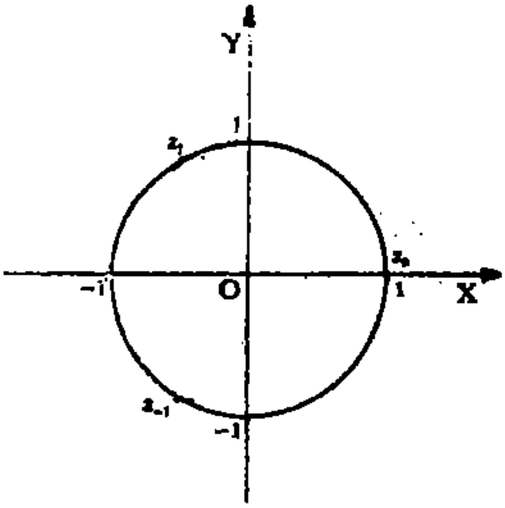


Figure 12 : Cube roots of unity

Thus, the roots are $1, \frac{-1}{2} + i\frac{\sqrt{3}}{2}$ and $\frac{-1}{2} - i\frac{\sqrt{3}}{2}$.

E35) We want to obtain those $z \in C$ for which

$(z^2 - 2) = \pm(1 + i)$, that is,

$z^2 - 2 = 1 + i$ and $z^2 - 2 = -(1 + i)$, that is,

$z^2 = 3 + i$ and $z^2 = 1 - i$.

Thus, we want to find the square roots of $3 + i$ and $1 - i$.

Now, $3 + i = \sqrt{10}\left\{\cos(\tan^{-1}\frac{1}{3}) + i \sin\left(\tan^{-1}\frac{1}{3}\right)\right\}$.

Thus, the square roots of $3 + i$ are

$10^{1/4}\left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right)$ and $10^{1/4}\left\{\cos\left(\frac{\theta}{2} + \pi\right) + i\sin\left(\frac{\theta}{2} + \pi\right)\right\}$,

where $\theta = \tan^{-1}\frac{1}{3}$.

Also $1 - i = \sqrt{2}\left(\cos\left(\frac{-\pi}{4}\right) + i\sin\left(\frac{-\pi}{4}\right)\right)$, so that the square roots of $1 - i$ are

$2^{1/4}\left(\cos\frac{\pi}{8} - i\sin\frac{\pi}{8}\right)$ and $2^{1/4}\left(\cos\frac{7\pi}{8} - i\sin\frac{7\pi}{8}\right)$.

These 4 square roots are the 4 roots of the given equation.

E36) a) If $a \geq 0$, then by E33, a has a real cube root, $a^{1/3}$. Now, $a = a(\cos 0 + i\sin 0)$.

Thus, the cube roots of a are

$a^{1/3}\left(\cos\frac{2k\pi}{3} + i\sin\frac{2k\pi}{3}\right)$, $k = 0, 1, 2$,

that is, $a^{1/3}, a^{1/3}\omega, a^{1/3}\omega^2$.

If $a < 0$, then $-a > 0$. Thus, $-a$ has a real cube root, say b. Then $r = -b$ is a real cube root of a. And $|r| = |a|^{1/3}$, that is, $r = -|a|^{1/3}$ (since r is negative).

Now $a = |a|(\cos\pi + i\sin\pi)$. Therefore, the cube roots of a are

$$|a|^{1/3}\left(\cos\frac{(2k+1)\pi}{3} + i\sin\frac{(2k+1)\pi}{3}\right), k = 0, 1, 2.$$

$$= r(\cos\pi + i\sin\pi)\left(\cos\frac{(2k+1)\pi}{3} + i\sin\frac{(2k+1)\pi}{3}\right), k = 0, 1, 2.$$

(since $-1 = \cos\pi + i\sin\pi$.)

$$= r\left(\cos\frac{(2k+4)\pi}{3} + i\sin\frac{(2k+4)\pi}{3}\right), k = 0, 1, 2.$$

Thus, the cube roots of a are r, rω, rω².

b) Let $n = 2m$, $m \in N$. Then, for any $b \in \mathbf{R}$,

$b^n = b^{2m} = (b^2)^m \geq 0$.

Thus, $b^n \neq a$ for any $b \in \mathbf{R}$. Hence, a can't have a real nth root.

c) Let $z = r(\cos\theta + i\sin\theta)$, in polar form.

Then its cube roots are $r^{1/3}\left(\cos\frac{\theta + 2k\pi}{3} + i\sin\frac{\theta + 2k\pi}{3}\right)$, $k = 0, 1, 2$.

Thus, if $\gamma = r^{1/3}\left(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\right)$, then the other roots are

$$r^{1/3}\left(\cos\frac{\theta + 2\pi}{3} + i\sin\frac{\theta + 2\pi}{3}\right) = \gamma\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right) = \gamma\omega, \text{ and}$$

$$r^{1/3}\left(\cos\frac{\theta + 4\pi}{3} + i\sin\frac{\theta + 4\pi}{3}\right) = \gamma\omega^2.$$

UNIT 3 CUBIC AND BIQUADRATIC EQUATIONS

Structure

3.1 INTRODUCTION

In this unit we will look at an aspect of algebra that has exercised the minds of several mathematicians through the ages. We are talking about the solution of polynomial equations over R. The ancient Hindu, Arabic and Babylonian mathematicians had discovered methods of solving linear and quadratic equations. The ancient Babylonians and Greeks had also discovered methods of solving some cubic equations. But, as we have said in Unit 2, they had not thought of complex numbers. So, for them, a lot of quadratic and cubic equations had no solutions.

In the 16th century various Italian mathematicians were looking into the geometrical problem of trisecting an angle by straight edge and compass. In the process they discovered a method for solving the general cubic equation. This method was divulged by Girolamo Cardano, and hence, is named after him. This is the same Cardano who was the first to introduce complex numbers into algebra. Cardano also published a method developed by his contemporary, Ferrari, for solving quartic equations. Later, in the 17th century, the French mathematician Descartes developed another method for solving 4th degree equations.

In this unit we will acquaint you with the solutions due to Cardano, Ferrari and Descartes. But first we will quickly cover methods for solving linear and quadratic equations. In the process we will also touch upon some general theory of equations.

There are several reasons, apart from a mathematician's natural curiosity, for looking at cubic and biquadratic equations. The material covered in this unit is also useful for mathematicians, physicists, chemists and social scientists.

After going through the unit, please check to see if you have achieved the following objectives.

Objectives

After studying this unit, you should be able to

solve a linear equation;

- solve a quadratic equation;
- apply Cardano's method for solving a cubic equation;
- apply Ferrari's or Descartes' method for solving a quartic equation;
- use the relation between roots and the coefficients of a polynomial equation for obtaining solutions.

3.2 LET US RECALL

You may be familiar with expressions of the form $2x + 5$, $-5x^2 + \frac{7}{2}$, $\sqrt{2}\,x^3 + x^2 + 1$, etc.

All these expressions are polynomials in one variable with coefficients in R. In general, we have the following definitions.

Definitions : An expression of the form $a_0 x^0 + a_1 x^1 + a_2 x^2 + \ldots\ldots + a_n x^n$, where $n \in N$ and $a_i \in C\ \forall i = 1, \ldots, n$, is called a polynomial over C in the variable x.

$a_0, a_1, \ldots, a_n$ are the **coefficients** of the polynomial.

If $a_n \neq 0$, we say that the degree of the polynomial is n and the leading term is $a_n x^n$.

While discussing polynomials we will observe the following conventions.

Conventions : We will

i) write x^0 as 1, so that we will write a_0 for $a_0 x^0$,

ii) write x^1 as x,

iii) write x^m instead of $1.x^m$ (i.e., when $a_m = 1$),

iv) omit terms of the type $0.x^m$.

Thus, the polynomial $2 + 3x^2 - x^3$ is $2x^0 + 0.\,x^1 + 3x^2 + (-1)\,x^3$.

We usually denote polynomials in x by f (x), g (x), etc. If the variable x is understood, then we often write f instead of f (x). We denote the degree of a polynomial f (x) by deg f (x) or deg f.

Note that the degree of f (x) is the highest power of x occurring in f (x). For example,

i) $3x + 6x^2 + \frac{5}{2}ix^3$ is a polynomial of degree 3,

ii) x^5 is a polynomial of degree 5, and

iii) $2 + i$ is a polynomial of degree 0, since $2 + i = (2 + i)\,x^0$.

Remark 1 : If f(x) and g(x) are two polynomials, then

$\deg\,(f(x) + g(x)) \leq \max\,(\deg f(x), \deg g(x))$

$\deg\,(f(x)\,.\,g(x)) \leq \deg f(x) + \deg g(x)$.

We say that f (x) is a polynomial over **R** if its coefficients are real numbers, and f (x) is over Q if its coefficients are rational numbers. For example, $2x + 3$ and $x^2 + 3$ are polynomials over Q as well as R (of degrees 1 and 2, respectively). On the other hand, $\sqrt{5}$ is a polynomial (of degree 0) over R but not over Q. In this course we shall almost always be dealing with polynomials over R.

Note that any non-zero element of R is a polynomial of degree 0 over R.

We define the degree of 0 to be $-\infty$.

Now, if we put a polynomial of degree n equal to zero, we get a polynomial equation of degree n, or an nth degree equation.

For example,

(i) $2x + 3 = 0$ is a polynomial equation of degree 1, and

(ii) $3x^2 + \sqrt{2}\,x - 1 = 0$ is a polynomial equation of degree 2.

If $f(x) = a_0 + a_1 x + \ldots\ldots + a_n x^n$ is a polynomial and $a \in C$, we can substitute x by a to get f(a), the value of the polynomial at x = a. Thus, $f(a) = a_0 + a_1 a + a_2 a^2 + \ldots\ldots + a_n a^n$.

For example, if $f(x) = 2x + 3$, then $f(1) = 2.1 + 3 = 5$, $f(i) = 2i + 3$, and $f\left(\frac{-3}{2}\right) = 2\left(\frac{-3}{2}\right) + 3 = 0$.

Since $f\left(\frac{-3}{2}\right) = 0$, we say that $\frac{-3}{2}$ is a root of f (x).

Definition : Let f (x) be a non-zero polynomial. $\alpha \in C$ is called a root (or a zero) of f (x) if $f(\alpha) = 0$.

In this case we also say that α is a **solution** (or a **root**) of the equation f (x) = 0.

A polynomial equation can have several solutions. For example, the equation $x^2 - 1 = 0$ has the two solutions x = 1 and x = – 1.

The set of solutions of an equation is called its **solution set**. Thus, the solution set of $x^2 + 1 = 0$ is $\{i, -i\}$.

Another definition that you will need quite often is the following.

Definition : Two polynomials $a_0 + a_1 x + \ldots + a_n x^n$ and $b_0 + b_1 x \ldots + b_m x^m$ are called equal if n = m and $a_i = b_i \;\forall\; i = 0, 1, \ldots, n$.

Thus, two polynomials are equal if they have the same degree and their corresponding coefficients are equal. Thus, $2x^3 + 3 = ax^3 + bx^2 + cx + d$ iff $a = 2,\ b = 0,\ c = 0,\ d = 3$

Let us now take a brief look at polynomials over **R** whose degrees are 1 or 2, and their solutions sets. We start with degree 1 equations.

3.2.1 Linear Equations

Consider any polynomial ax + b with $a, b \in R$ and $a \neq 0$. We call such a polynomial a **linear polynomial**. If we put it equal to zero, we get a **linear equation**.
Thus,

$ax + b = 0,\ a, b \in R,\ a \neq 0,$

is the most general form of a linear equation.

You know that this equation has a solution in **R**, namely, $x = \frac{-b}{a}$; and that this is the only solution.

Sometimes you may come across equations that don't appear to be linear, but, after simplification they become linear.

Let us look at some examples.

Example 1 : Solve $\frac{3p-1}{3} - \frac{2p}{p-1} = p$. (Here, we must assume $p \neq 1$.)

Solution : At first glance, this equation in p does not appear to be linear. But, by cross-multiplying, we get the following equivalent equation :

Two equations are equivalent if their solution sets are equal.

$(3p-1)(p-1) - 3(2p) = 3(p-1)p.$

On simplifying this we get

$3p^2-4p+1-6p = 3p^2-3p,$

that is, $7p-1=0$.

The solution set of this equation is $\left\{\frac{1}{7}\right\}$. Thus, this is the solution set of the equation we started with.

Example 2: Suppose I buy two plots of land for Rs. 1,20,000, and then sell them. Also suppose that I have made a profit of 15% on the first plot and a loss of 10% on the second plot. If my total profit is Rs. 5500, how much did I pay for each piece of land ?

Solution : Suppose the first piece of land cost Rs. x. Then the second piece cost Rs. $(1{,}20{,}000-x)$. Thus, my profit is Rs. $\frac{15}{100}$ x and my loss is Rs. $\frac{10}{100}(1,20,000-x)$.

$\therefore \frac{15}{100}x - \frac{10}{100}(1,20,000-x) = 5500$

$\Leftrightarrow 25x - 1,750,000 = 0$

$\Leftrightarrow x = 70,000.$

Thus, the first piece cost Rs. 70,000 and the second plot cost Rs. 50,000.

You may like to try these exercises now.

E1) Solve each of the following equations for the variable indicated. Assume that all denominators are non-zero.

a) $J\left(\frac{x}{k}+a\right) = x$ for x, where J, k and a are constants.

b) $\frac{1}{R} = \frac{1}{r_1}+\frac{1}{r_2}$ for R, keeping r_1 and r_2 constant.

c) $C = \frac{5}{9}(F-32)$ for F, keeping C constant.

E2) An isosceles triangle has a perimeter of 30 cm. Its equal sides are twice as long as the third side. Find the lengths of the three sides.

E3) A student cycles from her home to the study centre in 20 minutes. The return journey is uphill and takes her half an hour. If her rate is 8 km per hour slower on the return trip, how far does she live from the study centre?

E4) Simple interest is directly proportional to the principal amount as well as the time for which the amount is invested. If Rs. 1000, left at interest for 2 years, earns Rs. 110, find the amount of interest earned by Rs. 5000 for 3 years.

(Hint : $S = kPt$, where k is the constant of proportionality, S is the simple interest, P is the principal and t is the time.)

Now that we have looked at first degree equations, let us consider second degree equations, that is, equations of degree 2.

3.2.2 Quadratic Equations

Consider the general polynomial in x over R of degree 2 :

ax^2+bx+c, where $a, b, c \in \mathbb{R}$, $a \neq 0$.

We call this polynomial a quadratic polynomial. On equating a quadratic polynomial to zero, we get a quadratic equation in standard form.

The word 'quadratic' comes from the Latin word 'quadratum' meaning 'square'.

Can you think of an example of a quadratic equation? One is $x^2 = 5$, which is the same as $x^2-5 = 0$. Another is the equation Cardano tried to solve, namely, $x^2-10x+40 = 0$ (see Sec. 2.1). We are sure you can think of several others.

Various methods for solving such equations have been known since Babylonian times (2000 B.C.). Brahmagupta, in 628 A.D. approximately, also gave a rule for solving quadratic equations. The method that can be used for any quadratic equation is "completing the square". Using it we get the quadratic formula. Let us state this formula.

> **Quadratic Formula :** The solutions of the quadratic equation $ax^2 + bx + c = 0$, where $a, b, c \in \mathbb{R}$ and $a \neq 0$, are
> $$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

The expression $b^2 - 4ac$ is called the **discriminant** of $ax^2 + bx + c = 0$.

Note that this formula tells us that a quadratic equation has only two roots. These roots may be equal or they may be distinct; they may be real or complex.

Convention : We call a root that lies in $\mathbb{C} \setminus \mathbb{R}$ a complex root, that is, a root of the form $a + ib$, $a, b \in \mathbb{R}$, $b \neq 0$, is a complex root.

By a complex root we mean a root in $\mathbb{C} \setminus \mathbb{R}$.

Let us consider some examples.

Example 3 : Solve

i) $x^2 - 4x + 1 = 0$

ii) $4x^2 + 25 = 20x$

iii) $x^2 - 10x + 40 = 0$

Solution : i) This equation is in standard form. So we can apply the quadratic formula immediately. Here $a = 1$, $b = -4$, $c = 1$. Substituting these values in the quadratic formula, we get the two roots of the equation to be

$$x = \frac{-(-4) + \sqrt{(-4)^2 - 4(1)(1)}}{2(1)} = \frac{4 + \sqrt{12}}{2} = 2 + \sqrt{3}, \text{ and}$$

$$x = \frac{-(-4) - \sqrt{(-4)^2 - 4(1)(1)}}{2(1)} = 2 - \sqrt{3}.$$

Thus, the solutions are $2 + \sqrt{3}$ and $2 - \sqrt{3}$, two distinct elements of $\mathbb{R}$.

Note that in this case the discriminant was positive.

ii) In this case let us first rewrite the equation in standard form as

$4x^2 - 20x + 25 = 0$.

Now, putting $a = 4$, $b = -20$, $c = 25$ in the quadratic formula, we find that

$$x = \frac{20 + \sqrt{400 - 4(4)(25)}}{2(4)} = \frac{20 + \sqrt{0}}{8} = \frac{5}{2}, \text{ and}$$

$$x = \frac{20 - \sqrt{400 - 4(4)(25)}}{2(4)} = \frac{5}{2}.$$

Here we find that both the roots coincide and are real.

Note that in this case the discriminant is 0.

iii) Using the quadratic formula, we find that the solutions are

$$x = \frac{10 \pm \sqrt{100 - 160}}{2} = 5 \pm \frac{\sqrt{-60}}{2} = 5 \pm \frac{\sqrt{4(-15)}}{2}$$

$$= 5 \pm \sqrt{-15}$$

$$= 5 \pm i\sqrt{15}.$$

Thus, in this case we get two distinct complex roots, $5 + i\sqrt{15}$ and $5 - i\sqrt{15}$.

Note that in this case the discriminant is negative.

In the example above do you see a relationship between the types of roots of a quadratic equation and the value of its discriminant? There is such a relationship, which we now state.

> The equation $ax^2 + bx + c = 0$, $a \neq 0$, $a, b, c \in R$ has two roots. They are
> i) real and distinct if $b^2 - 4ac > 0$;
> ii) real and equal if $b^2 - 4ac = 0$;
> iii) complex and distinct if $b^2 - 4ac < 0$.

Now, is there a difference in the character of the roots of $ax^2 + bx + c = 0$ and $dax^2 + dbx + dc = 0$, where d is a non-zero real number? For example, if $b^2 - 4ac > 0$, what is the sign of $(db)^2 - 4(da)(dc)$? It will also be positive. In fact, **the character of the roots of equivalent quadratic equations is the same.**

Now let us consider some important remarks which will be useful to you while solving quadratic equations.

Remark 2 : α and β are roots of a quadratic equation $ax^2 + bx + c = 0$ if any only if

$ax^2 + bx + c = a(x - \alpha)(x - \beta)$.

Thus, $\alpha \in C$ is a root of $ax^2 + bx + c = 0$ if and only if $(x - \alpha) \mid (ax^2 + bx + c)$.

Remark 3 : From the quadratic formula you can see that if $b^2 - 4ac < 0$, then the quadratic equation $ax^2 + bx + c = 0$ has 2 complex roots which are each other's conjugates.

Remark 4 : Sometimes a quadratic equation can be solved without resorting to the quadratic formula. For example, the equation $x^2 = 9$ clearly has 3 and -3 as its roots. Similarly, the equation $(x - 1)^2 = 0$ clearly has two coincident roots, both equal to 1 (see Remark 2).

Using what we have said so far, try and solve the following exercises.

E5) A quadratic equation over R can have complex roots while a linear equation over R can only have a real root. True or false ? Why?

E6) Solve the following equations:

a) $x^2 + 5 = 0$

b) $(x + 9)(x - 1) = 0$

c) $x^2 - \sqrt{5}x = 1$

d) $pm^2 - 8qm + \frac{1}{r} = 0$ for m, where $p, q, r \in R$ and $p, r \neq 0$.

E7) For what values of k will the equation

$kx^2 + (2k + 6)x + 16 = 0$ have coincident roots ?

E8) Show that the quadratic equation $ax^2 + bx + c = 0$ has equal roots if $(2ax + b) \mid (ax^2 + bx + c)$.

E9) Find the values of b and c for which the polynomial $x^2 + bx + c$ has $1 + i$ and $1 - i$ as its roots.

E10) If α and β are roots of $ax^2 + bx + c = 0$, then show that $\alpha + \beta = -\frac{b}{a}$ and $\alpha\beta = \frac{c}{a}$.

E11) Let $\alpha, \beta \in C$ such that $\alpha + \beta = p \in R$ and $\alpha\beta = q \in R$. Show that α and β are the roots of $x^2 - px + q = 0$.

E11 is the converse of E10. We will use it in the next section.

Let us now consider some equations which are not quadratic, but whose solutions can be obtained from related quadratic equations. Look at the following example.

Example 4 : Solve

i) $2x^4 + x^2 + 1 = 0$, and

ii) $x = \sqrt{15 - 2x}$.

Solution : i) $2x^4 + x^2 + 1 = 0$ can be written as $2y^2 + y + 1 = 0$, where $y = x^2$. Then, solving this for y, we get $y = \frac{-1 \pm i\sqrt{7}}{4}$, that is, $x^2 = \frac{-1 \pm i\sqrt{7}}{4}$, two polynomials over **C**.

Thus, the four solutions of the original equation are

$$\sqrt{\frac{-1+i\sqrt{7}}{4}},\ -\sqrt{\frac{-1+i\sqrt{7}}{4}},\ \sqrt{\frac{-1-i\sqrt{7}}{4}},\ -\sqrt{\frac{-1-i\sqrt{7}}{4}}$$

ii) $x = \sqrt{15-2x}$ is not a polynomial equation. We square both sides to obtain the polynomial equation $x^2 = 15 - 2x$.

Now, any root of $x = \sqrt{15-2x}$ is also a root of the equation $x^2 = 15 - 2x$.

But the converse need not be true, since $x^2 = 15 - 2x$ can also mean $x = -\sqrt{15-2x}$.

So we will obtain the roots of $x^2 = 15 - 2x$ and see which of these satisfy $x = \sqrt{15-2x}$.

Now, the roots of the quadratic equation $x^2 = 15 - 2x$ are $x = -5$ and $x = 3$. We must put these values in the original equation to see if they satisfy it.

Now, for $x = -5$,

$x - \sqrt{15-2x} = (-5) - \sqrt{15+10} = (-5) - 5 = -10 \neq 0$.

So $x = -5$ is not a solution of the given equation. But it is a solution of $x^2 = 15 - 2x$. We call it an **extraneous solution.**

What happens when we put $x = 3$ in the given equation? We get $3 = \sqrt{15-6}$ i.e., $3 = 3$, which is true. Thus, $x = 3$ is the solution of the given equation.

Now you may like to solve the following exercises. Remember that you must check if the solutions you have obtained satisfy the given equations. This will help you

i) to get rid of extraneous solutions, if any, and

ii) to ensure that your calculations are alright.

E12) Reduce the following to quadratic equations and hence, solve them.

a) $4p^4 - 16p^2 + 5 = 0$

b) $(5x^2 - 6)^{1/4} = x$

c) $\sqrt{2x+3} - \sqrt{x+1} = 1$

E13) Ameena walks 1 km per hour faster than Alka. Both walked from their village to the nearest library, a distance of 24 km. Alka took 2 hours more than Ameena. What was Alka's average speed ?

In this section our aim was to help you recall the methods of solving linear and quadratic equations. Let us now see how to solve equations of degree 3.

3.3 CUBIC EQUATIONS

In this section we are going to discuss some mathematics to which the great 11th century Persian poet Omar Khayyam gave a great deal of thought. He, and Greek mathematicians before him, obtained solutions for third degree equations by considering geometric methods that involved the intersection of conics. But we will only discuss algebraic methods of obtaining solutions of such equations, that is, solutions obtained by using the basic algebraic operations and by radicals. Let us first see what an equation of degree 3, or a cubic equation, is.

Definition : An equation of the form

$ax^3 + bx^2 + cx + d = 0$ with $a, b, c, d \in R$, $a \neq 0$,

is the most general form of a **cubic equation** (or a **third degree equation**) over **R**.

For example $2x^3 = 0$, $\sqrt{3}\,x^3 + 5x^2 = 0$, $-2x = 5x^3 - 1$ and $x^3 + 5x^2 + 2x = -7$ are all cubic equations, since each of them can be written in the form $ax^3 + bx^2 + cx + d = 0$, with $a \neq 0$. On the other hand, $x^4 + 1 = 0$, $x^3 + 2x^2 = x^3 - x$ and $x^3 + \sqrt{x} = 0$ are not cubic equations.

There are several situations in which one needs to solve cubic equations. For example, many problems in the social, physical and biological sciences reduce to obtaining the eigenvalues of a 3 × 3 matrix (which you can study about in the Linear Algebra course). And for this you need to know how to obtain the solutions of a cubic equation.

For obtaining solutions of a cubic equation, or any polynomial equation, we need some results about the roots of polynomial equations. We will briefly discuss them one by one. We give the first one without proof.

Theorem 1 : The polynomial equation of degree n,

$a_0 + a_1 x + \ldots + a_n x^n = 0$, where $a_0, a_1, \ldots, a_n \in R$ and $a_n \neq 0$, has n roots, which are real or non-real complex numbers.

If $x_1, \ldots, x_n$ are the n roots of the equation in Theorem 1, then
$a_0 + a_1 x + \ldots + a_n x^n = a_n (x - x_1)(x - x_2) \ldots (x - x_n)$.

(Note that the roots need not be distinct. For example, $1 + 2x + x^2 = (x + 1)^2$)

We will not prove this fact here; but we will now state a very important result which is used in the proof.

Theorem 2 (Division algorithm) : Given polynomials f (x) and g (x) (≠ 0) over R, ∃ polynomials q(x) and r(x) over **R** such that
$f(x) = g(x)\,q(x) + r(x)$ and $\deg r(x) < \deg g(x)$.

We will also use this theorem to prove the following result which tells us something about complex roots, that is, roots that are non-real complex numbers.

Theorem 3 : If a polynomial equation over R has complex roots, they occur in pairs. In fact, if a + ib ∈ C is a root, then a – ib is also a root.

Proof : Let $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ be a polynomial over **R** of degree n. Suppose $a + ib \in C$ is a root of $f(x) = 0$, that is $(x - (a + ib)) \mid f(x)$. We want to show that $(x - (a - ib)) \mid f(x)$ too.
Now, $(x - (a + ib))(x - (a - ib)) = (x - a)^2 + b^2$

Also, by the division algorithm, ∃ polynomials g (x) and r (x) over R such that

$f(x) \mid g(x)$
$\Rightarrow \deg f(x) \leq \deg g(x)$
(see Remark 1).

$f(x) = \{(x - a)^2 + b^2\}\,g(x) + r(x)$, where $\deg r(x) < 2$.

Since $x - (a + ib)$ divides $f(x)$ and $(x - a)^2 + b^2$, it divides $f(x) - \{(x - a)^2 + b^2\}\,g(x)$, that is, r (x).

But r(x) is linear over R or a constant in R. So (x – (a + ib)) can only divide r (x) if r (x) = 0.

So we find that

$$f(x) = \{(x-a)^2 + b^2\}\, g(x).$$

Since $x - (a - ib)$ divides the right hand side of this equation, it must divide $f(x)$.

Thus, $a - ib$ is a root of $f(x) = 0$ also.

Note that Theorem 3 does not say that $f(x) = 0$ must have a complex root. It only says that if it has a complex root, then the conjugate of the root is also a root.

Why don't you try the following exercises now ? In them we are just recalling some facts that you are already aware of.

E14) How many complex roots can a linear equation over **R** have?

E15) Under what circumstances does the quadratic equation over **R**, $x^2 + px + q = 0$, have complex roots? If it has complex roots, how many are they and how are they related?

Now let us look at Theorems 1 and 3 in the context of cubic equations. Consider the general cubic equation over **R**,

$$ax^3 + bx^2 + cx + d = 0,\ a \neq 0.$$

Any solution of this is also a solution of

$$x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} = 0,$$

and vice versa.

Thus, we can always assume that the most general equation of degree 3 over **R** is

$$x^3 + px^2 + qx + r = 0 \text{ with } p, q, r \in \mathbf{R}.$$

Theorem 1 says that this equation has 3 roots. Theorem 3 says that either all 3 roots are real or one is real and two are complex. Let us find these roots algebraically.

3.3.1 Cardano's Solution

Figure 1 : Card

The algebraic method of solving cubic equations is supposed to be due to the Italian, del Ferro (1465-1526). But it is called Cardano's method because it became known to people after the Italian, Girolamo Cardano, published it in 1545 in his 'Ars Magna'.

Let us see what the method is. We will first look at a particular case.

Example 5 : Solve $2x^3 + 3x^2 + 4x + 1 = 0$.

Solution : We first remove the second degree term by completing the cube in the following way.

$$2x^3 + 3x^2 + 4x + 1 = 0$$

$$\Leftrightarrow x^3 + \frac{3}{2}x^2 + 2x + \frac{1}{2} = 0$$

$$\Leftrightarrow \left[\left(x + \frac{1}{2}\right)^3 - \frac{3}{4}x - \frac{1}{8}\right] + 2x + \frac{1}{2} = 0$$

$$\Leftrightarrow \left(x + \frac{1}{2}\right)^3 + \frac{5}{4}x + \frac{3}{8} = 0$$

Put $x + \frac{1}{2} = y$. Then the equation becomes

$$y^3 + \frac{5}{4}y - \frac{1}{4} = 0.$$

Assume that the solution is $y = m + n$, where $m, n \in \mathbf{C}$. Then

$$(m+n)^3+\frac{5}{4}(m+n)-\frac{1}{4}=0$$

$$\Leftrightarrow m^3+3mn(m+n)+n^3+\frac{5}{4}(m+n)-\frac{1}{4}=0$$

$$\Leftrightarrow m^3+n^3+\left(3mn+\frac{5}{4}\right)(m+n)-\frac{1}{4}=0. \qquad \text{..........(1)}$$

Let us add a further condition on m and n, namely,

$$3mn+\frac{5}{4}=0, \text{ that is, } mn=-\frac{5}{12} \qquad \text{..........(2)}$$

Then (1) gives us $m^3+n^3=\frac{1}{4}$,

and (2) gives us $m^3n^3=-\frac{125}{1728}$.

Thus, using E11, we see that m^3 and n^3 are roots of

$$t^2-\frac{1}{4}t-\frac{125}{1728}=0.$$

Hence, by the quadratic formula we find that

$$m^3=\frac{1}{8}\left(1+\sqrt{\frac{152}{27}}\right)=\alpha, \text{ say,}$$

$$\text{and } n^3=\frac{1}{8}\left(1-\sqrt{\frac{152}{27}}\right)=\beta, \text{ say.}$$

From Unit 2 (E36) you know that α and β have real roots, say u and v, respectively. Thus, m can take the values u, ωu, $\omega^2 u$, and n can take the values v, ωv, $\omega^2 v$.

Now, ω and ω^2 are non-real complex numbers such that $\omega(\omega^2)=1$.

Also, from (2) we know that $mn=-\frac{5}{12}$, a real number.

Thus, if m = u, n must be v; if $m=\omega u$, n must be $\omega^2 v$; if $m=\omega^2 u$, n must be ωv.

Hence, the possible values of y are

$u+v$, $\omega u+\omega^2 v$, $\omega^2 u+\omega v$.

To get the three roots of the original equation, we simply put these values of y in the relation

$$x=y-\frac{1}{2}.$$

This example has probably given you some idea about Cardano's method for solving a general cubic equation. Let us outline this method for solving the general equation

$$x^3+px^2+qx+r=0,\ p, q, r \in \mathbb{R}. \qquad \text{......... (3)}$$

Step 1: We first write $x^3+px^2=\left(x+\frac{p}{3}\right)^3-\frac{p^2}{3}x-\frac{p^3}{27}$.

Then (3) becomes

$$\left(x+\frac{p}{3}\right)^3+qx+r-\left(\frac{p^2}{3}x+\frac{p^3}{27}\right)=0$$

$$\Leftrightarrow \left(x+\frac{p}{3}\right)^3+\left(q-\frac{p^2}{3}\right)x+\left(r-\frac{p^3}{27}\right)=0.$$

Now put $y=x+\frac{p}{3}$. Then $x=y-\frac{p}{3}$, and the equation becomes

$y^3 + \left(q - \frac{p^2}{3}\right)\left(y - \frac{p}{3}\right) + \left(r - \frac{p^3}{27}\right) = 0$, that is,

$$y^3 + Ay + B = 0, \qquad \text{.......... (4)}$$

where $A = q - \frac{p^2}{3}$ and $B = \frac{2p^3}{27} - \frac{pq}{3} + r$.

Step 2 : Now let us solve (4).

Let $y = \alpha + \beta$ be a solution. Putting this value of y in (4) we get

$(\alpha + \beta)^3 + A(\alpha + \beta) + B = 0$

$\Leftrightarrow \alpha^3 + 3\alpha\beta(\alpha + \beta) + \beta^3 + A(\alpha + \beta) + B = 0$

$$\Leftrightarrow \alpha^3 + \beta^3 + (3\alpha\beta + A)(\alpha + \beta) + B = 0 \qquad \text{.......... (5)}$$

Now, we choose α and β so that, $3\alpha\beta + A = 0$. Then we have the two equations

$$(\alpha\beta)^3 = (-\frac{A}{3})^3, \text{ that is, } \alpha^3\beta^3 = -\frac{A^3}{27}, \qquad \text{.......... (6)}$$

and from (5)

$$\alpha^3 + \beta^3 = -B. \qquad \text{.......... (7)}$$

Thus, using E11, we find that α^3 and β^3 are roots of the quadratic equation

$$t^2 + Bt - \frac{A^3}{27} = 0. \qquad \text{.......... (8)}$$

Hence, using the quadratic formula, we find that

$$\left.\begin{aligned} \alpha^3 &= -\frac{B}{2} + \sqrt{\frac{B^2}{4} + \frac{A^3}{27}} = u, \text{ say, and} \\ \beta^3 &= -\frac{B}{2} - \sqrt{\frac{B^2}{4} + \frac{A^3}{27}} = v, \text{ say.} \end{aligned}\right] \qquad \text{.......... (9)}$$

Now, from Unit 2 (E36) we know that any complex number has three cube roots. We also know that if γ is a cube root, then the three roots are γ, $\omega\gamma$ and $\omega^2\gamma$.

Therefore, if a and b denote a cube root each of u and v, respectively, then α can be a, aω or aω^2; and β can be b, bω or bω^2. Does this mean that $y = \alpha + \beta$ can take on 9 values ? Note that α and β also satisfy the relation $\alpha\beta = -\frac{A}{3} \in \mathbb{R}$.

Thus, since $\omega \in \mathbb{C}$, $\omega^2 \in \mathbb{C}$, $\omega^3 = 1 \in \mathbb{R}$, the only possibilities for y are $a + b$, $a\omega + b\omega^2$, $a\omega^2 + b\omega$.

Step 3 : The 3 solutions of (3) are given by substituting each of these values of y in the equation $x = y - \frac{p}{3}$.

So, what we have just shown is that

> the roots of $x^3 + px^2 + qx + r = 0$ are $\alpha + \beta - \frac{p}{3}$, $\alpha\omega + \beta\omega^2 - \frac{p}{3}$, $\alpha\omega^2 + \beta\omega - \frac{p}{3}$,
> where $\omega = \frac{-1 + i\sqrt{3}}{2}$, α is a cube root of $\left\{-\frac{B}{2} + \sqrt{\frac{B^2}{4} + \frac{A^3}{27}}\right\}$, β is a cube root of $\left\{-\frac{B}{2} - \sqrt{\frac{B^2}{4} + \frac{A^3}{27}}\right\}$, $A = q - \frac{p^2}{3}$, $B = \frac{2p^3}{27} - \frac{pq}{3} + r$.

The formula we have obtained is rather a complicated business. A calculator would certainly ease matters, as you may find while trying the following exercise.

E16) Solve the following cubic equations :

a) $2x^3 + 3x^2 + 3x + 1 = 0$

b) $x^3 + 21x + 342 = 0$

c) $x^3 + 6x^2 + 6x + 8 = 0$

d) $x^3 + 29x - 97 = 0$

e) $x^3 = 30x - 133$

In each of the equations in E16, you must have found that $\frac{B^2}{4} + \frac{A^3}{27} \geq 0$.

But what happens if $\frac{B^2}{4} + \frac{A^3}{27} < 0$.

This case is known as the irreducible case. In this case (9) tells us that α^3 and β^3 are complex numbers of the form $a + ib$ and $a - ib$, where $b \neq 0$. From Unit 2 you know that if the polar form of $a + ib$ is $r(\cos\theta + i\sin\theta)$, then its cube roots are

$$r^{1/3}\left(\cos\frac{\theta + 2k\pi}{3} + i\sin\frac{\theta + 2k\pi}{3}\right), k = 0, 1, 2.$$

Similarly, the cube roots of $a - ib$ are

$$r^{1/3}\left(\cos\frac{\theta + 2k\pi}{3} - i\sin\frac{\theta + 2k\pi}{3}\right), k = 0, 1, 2.$$

Hence, the 3 values of y in (4) are

$$2r^{1/3}\cos\frac{\theta + 2k\pi}{3}, \quad \text{where } k = 0, 1, 2.$$

In the irreducible case α and β are not real, but x is real

All these are real numbers. Thus in this case all the roots of (3) are real, and are given by

$$2r^{1/3}\cos\frac{\theta}{3} - \frac{p}{3},\ 2r^{1/3}\cos\frac{\theta + 2\pi}{3} - \frac{p}{3},\ 2r^{1/3}\cos\frac{\theta + 4\pi}{3} - \frac{p}{3}.$$

This trigonometric form of the solution is due to François Viète (1550-1603).

Now try an exercise.

E17) Solve the equation $x^3 - 3x + 1 = 0$

So far, we have seen that a cubic equation has three roots. We also know that either all the roots are real, or one is real and two are complex conjugates. Can we tell the roots or the character of the roots by just inspecting the coefficients? We shall answer this question now.

3.3.2 Roots And Their Relation With Coefficients

In this sub-section we shall first look at the cubic analogue of E10 and E11. Over there we saw how closely the roots of a quadratic equation are linked with its coefficients. The same thing is true for a cubic equation. Why don't you try and prove the relationship that we give in the following exercise ?

E18) Show that α, β and γ are the roots of the cubic equation

$ax^3 + bx^2 + cx + d = 0$, $a \neq 0$, if and only if

$$\alpha + \beta + \gamma = -\frac{b}{a},$$

$$\alpha\beta + \beta\gamma + \alpha\gamma = \frac{c}{a},$$

$$\alpha\beta\gamma = -\frac{d}{a}.$$

(**Hint** : Note that the given cubic equation is equivalent to $a(x-\alpha)(x-\beta)(x-\gamma) = 0$.)

The relationship in E18 allows us to solve problems like the following.

Example 6 : If α, β, γ are the roots of the equation

$x^3 - 7x^2 + x - 5 = 0$,

find the equation whose roots are $\alpha+\beta$, $\beta+\gamma$, $\alpha+\gamma$.

Solution : By E18 we know that

$$\left.\begin{aligned} \alpha+\beta+\gamma &= 7 \\ \alpha\beta+\beta\gamma+\alpha\gamma &= 1 \\ \alpha\beta\gamma &= 5 \end{aligned}\right\} \quad \text{.......... (10)}$$

Therefore, $(\alpha+\beta)+(\beta+\gamma)+(\alpha+\gamma) = 2(\alpha+\beta+\gamma) = 14$. .... (11)

Also, $\alpha+\beta = 7-\gamma$, $\beta+\gamma = 7-\alpha$, $\gamma+\alpha = 7-\beta$, so that

$(\alpha+\beta)(\beta+\gamma)+(\beta+\gamma)(\alpha+\gamma)+(\alpha+\gamma)(\alpha+\beta)$

$= \{49-7(\gamma+\alpha)+\gamma\alpha\} + \{49-7(\alpha+\beta)+\alpha\beta\} + \{49-7(\beta+\gamma)+\beta\gamma\}$

$= 147 - 98 + 1$, using (10) and (11).

$= 50$, and .... (12)

$(\alpha+\beta)(\beta+\gamma)(\alpha+\gamma) = (7-\gamma)(7-\beta)(7-\alpha)$

To evaluate the expression on the right hand side, we can use (10) or we can use the fact that

$x^3 - 7x^2 + x - 5 = (x-\alpha)(x-\beta)(x-\gamma)$
$\Rightarrow 7^3 - 7.7^2 + 7 - 5 = (7-\alpha)(7-\beta)(7-\gamma)$

Therefore, $(\alpha+\beta)(\beta+\gamma)(\alpha+\gamma) = 2$. .... (13)

Now, E18, (11), (12) and (13) give us the required equation, which is

$x^3 - 14x^2 + 50x - 2 = 0$.

Why don't you try the following exercise now ?

E19) Find the sum of the cubes of the roots of the equation $x^3 - 6x^2 + 11x - 6 = 0$.

Hence find the sum of the fourth powers of the roots.

Let us now study the character of the roots of a cubic equation. For this purpose we need to introduce the notion of the discriminant. In the case of a quadratic equation $x^2 + bx + c = 0$, you know that the discriminant is $b^2 - 4c$. Also, if α and β are the two roots of the equation, then $\alpha+\beta = -b$, $\alpha\beta = c$. Therefore,

$(\alpha-\beta)^2 = (\alpha+\beta)^2 - 4\alpha\beta = b^2 - 4c$.

Thus, the discriminant $= (\alpha-\beta)^2$, where α and β are the roots of the quadratic equation.

Now consider the general quadratic equation, $ax^2 + bx + c = 0$.
Let its roots be α and β. Then its discriminant is $b^2 - 4ac = a^2(\alpha-\beta)^2$.

We use this relationship to define the discriminant of any polynomial equation.

Definition : The **discriminant** of the nth degree equation
$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$ is

$$a_n^{2(n-1)} \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2,$$

where $\alpha_1, \ldots, \alpha_n$ are the roots of the polynomial equation.

In particular, if we consider the case $n = 3$ and $a_n = 1$, we find that

> The discriminant of the cubic $x^3 + px^2 + qx + r = 0$ is
> $D = -(27B^2 + 4A^3)$, where $A = q - \frac{p^2}{3}$, $B = \frac{2p^3}{27} - \frac{pq}{3} + r$.

Now consider Cardano's solution of the cubic equation (3), namely,

$x^3 + px^2 + qx + r = 0$.

The expression under the square root sign is $\frac{B^2}{4} + \frac{A^3}{27} = \frac{-D}{108}$, where D is the discriminant.

Now, (9) tells us that the sign of the discriminant is closely related to the characters of the roots of the equation. Let us look at the different possibilities for the roots α, β and γ of (3).

1) The roots of (3) are all real and distinct. Then $(\alpha - \beta)^2 (\beta - \gamma)^2 (\alpha - \gamma)^2$, that is D, must be positive.

2) Only one root of (3) is real. Let this root be α. Then β and γ are complex conjugates.
   $\therefore$ $\beta - \gamma$ is purely imaginary $\therefore$ $(\beta - \gamma)^2 < 0$.
   Also, $\alpha - \beta$ and $\alpha - \gamma$ are conjugates.
   Therefore, their product is positive.
   Hence, in this case $D < 0$.

3) Suppose $\alpha = \beta$ and $\gamma \neq \alpha$. Since $\alpha - \beta = 0$, $D = 0$.
   Also, $B \neq 0$. Why? Because if $B = 0$, then $A = 0$ (since $D = 0$).
   But $A = 0 \Rightarrow q = \frac{p^2}{3}$, that is $\alpha(\alpha + 2\gamma) = \frac{(2\alpha + \gamma)^2}{3}$,
   [Over here we have used the relationship between the roots,
   since $p = -(\alpha + \beta + \gamma) = -(2\alpha + \gamma)$ and $q = \alpha\beta + \beta\gamma + \alpha\gamma = \alpha(\alpha + 2\gamma)$]
   On simplifying we get $\alpha = \gamma$, a contradiction.
   Thus, $B \neq 0$.
   So, if exactly two roots of (3) are equal, then $D = 0$ and $B \neq 0$, and hence, $A \neq 0$.

4) If all the roots of (3) are equal, then $D = 0$, $B = 0$, and hence $A = 0$.
   Let us summarize the different possibilities for the character of the roots now.

> Consider the cubic equation $x^3 + px^2 + qx + r = 0$, $p, q, r \in \mathbb{R}$,
> and let $B = \frac{2p^3}{27} - \frac{pq}{3} + r$ and $A = q - \frac{p^2}{3}$. Then
>
> 1) all its roots are real and distinct iff $\frac{B^2}{4} + \frac{A^3}{27} < 0$.
>
> 2) exactly one root is real iff $\frac{B^2}{4} + \frac{A^3}{27} > 0$
>
> 3) exactly two roots are equal iff $\frac{B^2}{4} + \frac{A^3}{27} = 0$ and $B \neq 0$.
>   In this case all the roots are real.
>
> 4) all three roots are equal iff $\frac{B^2}{4} + \frac{A^3}{27} = 0$ and $B = 0$.

You may now like to try the following problem to see if you've understood what we have just discussed.

E20) Under what conditions on the coefficients of

$ax^3 + 3bx^2 + 3cx + d = 0, \; a \neq 0,$

will the equation have complex roots?

E21) Will all the roots of $x^3 = 15x + 126$ be real ? Why ?

So far we have introduced you to a method of solving cubic equations and we have studied the solutions in some depth. We shall study them some more in Unit 6, as an application of the Cauchy- Schwarz inequality. Now let us go on to a discussion of polynomial equations of degree 4.

3.4 BIQUADRATIC EQUATIONS

As in the case of cubic equations, biquadratic equations have been studied for a long time. The ancient Arabs were known to have studied them from a geometrical point of view. In this section we will discuss two algebraic methods of solving such equations. Let us first see what a biquadratic equation is.

Definition : An equation of the form

$ax^4 + bx^3 + cx^2 + dx + e = 0$, where $a, b, c, d, e \in \mathbf{R}$ and $a \neq 0$,

is the most general form of a **biquadratic equation** (or a **quartic equation**, or a **fourth degree equation**) over **R**.

Can you think of examples of quartic equations over **R**? What about $x^4 + 5 = \sqrt{2}\,x - x^2$? This certainly is a quartic equation, as it is equivalent to $x^4 + x^2 - \sqrt{2}\,x + 5 = 0$.

What about $\sqrt{x} = x^4 + 1$? This isn't even a polynomial equation. So it can't be a quartic.

Let us now consider various ways in which we can solve an equation of degree 4. In some cases, as you have seen in Example 4, such an equation can be solved by solving related quadratic equations. But most biquadratic equations can't be solved in this manner. Two algebraic methods for obtaining the roots of such equations were developed in the 16th and 17th centuries. Both these methods depend on the solving of a cubic equation. Let us see what they are.

3.4.1 Ferrari's Solution

The first method for solving a biquadratic equation that we will discuss is due to the 16th century Italian mathematician Ferrari, who worked with Cardano. Let us see what the method is with the help of an example.

Example 7 : Solve the equation

$x^4 - 2x^3 - 5x^2 + 10x - 3 = 0.$

Solution : We will solve this in several steps.

Step 1 : Add the quadratic polynomial $(ax + b)^2 = a^2x^2 + 2abx + b^2$ to both sides. We get

$x^4 - 2x^3 + (a^2 - 5)x^2 + 2(ab + 5)x + b^2 - 3 = (ax + b)^2.$ .... (14)

Step 2 : Choose a and b in **R** so that the left hand side of (14) becomes a perfect square, say $(x^2 - x + k)^2$, where k is an unknown. Thus, we need to choose a and b so that

$$x^4 - 2x^3 + (a^2 - 5)x^2 + 2(ab + 5)x + b^2 - 3 = x^4 + x^2 + k^2 - 2x^3 - 2kx + 2kx^2.$$

Equating the coefficients of x^2, x and the constant term on both sides, we get

$a^2 - 5 = 2k + 1$ .... (15)

$2(ab + 5) = -2k$ .... (16)

$b^2 - 3 = k^2$ .... (17)

$(15) \Rightarrow a^2 = 2k + 6$

Also, $(16) \Rightarrow a = -\frac{1}{b}(k+5)$

Thus, $2k + 6 = \frac{1}{b^2}(k+5)^2$

Then $(17) \Rightarrow k^2 + 3 = \frac{(k+5)^2}{2k+6}$

$\Rightarrow 2k^3 + 5k^2 - 4k - 7 = 0$ .... (18)

This cubic equation is called the resolvent cubic of the given biquadratic equation. We have obtained it by eliminating a and b from the equations (15), (16) and (17).

We choose any one root of the cubic. One real solution of (18) is $k = -1$. (It is easy to see this by inspection. Otherwise you can apply Cardano's method.)

Then, from (15), (16) and (17) we get

$a^2 = 4, \; b^2 = 4, \; ab = -4.$

$a = 2$ and $b = -2$ (or $a = -2$ and $b = 2$) satisfy these equations. We need only one set of values of a and b. Either will do. Let us take $a = 2$ and $b = -2$.

Step 3 : Put these values of k, a and b in $(x^2 - x + k)^2 = (ax + b)^2$. On taking square roots, we get two quadratic equations, namely,

$x^2 - x - 1 = \pm (2x - 2)$, that is,

$x^2 - 3x + 1 = 0$ and $x^2 + x - 3 = 0$.

Applying the quadratic formula to these equations we get

$$x = \frac{3+\sqrt{5}}{2}, \frac{3-\sqrt{5}}{2}, \frac{-1+\sqrt{13}}{2}, \frac{-1-\sqrt{13}}{2}$$

Does Example 7 give you some idea of the general method developed by Ferrari? Let us see what it is.

We want to solve the general 4th degree equation over **R**, namely,

$x^4 + px^3 + qx^2 + rx + s = 0, \; p, q, r, s \in \mathbf{R}$ .... (19)

The idea is to express this equation as a difference of squares of two polynomials. Then this difference can be split into a product of two quadratic factors, and we can solve the two quadratic equations that we obtain this way. Let us write down the steps involved.

Step 1 : Add $(ax + b)^2$ to each side of (19), where a and b will be chosen so as to make the left hand side a perfect square. So (19) becomes

$x^4 + px^3 + (q + a^2)x^2 + (r + 2ab)x + s + b^2 = (ax + b)^2$ .... (20)

Step 2 : We want to choose a and b so that the left hand side is a perfect square, say $(x^2 + \frac{p}{2}x + k)^2$, where k is an unknown.

Note that the coefficient of x is necessarily $\frac{p}{2}$, since the coefficient of x^3 in (20) is p.

So we see that

$$x^4 + px^3 + (q + a^2)x^2 + (r + 2ab)x + s + b^2 = x^4 + px^3 + \left(\frac{p^2}{4}\right)x^2 + 2kx^2 + pkx + k^2.$$

Comparing coefficients of x^2, x and the constant term, we have

$\frac{p^2}{4} + 2k = q + a^2, \; pk = r + 2ab, \; k^2 = s + b^2.$

Eliminating a and b from these equations, we get the resolvent cubic

$(pk - r)^2 = 4\left(\frac{p^2}{4} + 2k - q\right)(k^2 - s)$, that is,

$8k^3 - 4qk^2 + 2(pr - 4s)k + (4qs - p^2s - r^2) = 0.$

From Sec. 3.3 you know that this cubic equation has at least one real root, say α.

Then, we can find a and b in terms of α.

Step 3 : Our assumption was that

$$(x^2+\frac{p}{2}x+k)^2 = (ax+b)^2.$$

Now, putting $k = \alpha$ and substituting the values of a and b, we get the quadratic equations

$x^2+\frac{p}{2}x+\alpha = \pm(ax+b)$, that is,

$x^2+(\frac{p}{2}-a)x+(\alpha-b) = 0$, and

$x^2+(\frac{p}{2}+a)x+(\alpha+b) = 0$.

Then, using the quadratic formula we can obtain the 4 roots of these equations, which will be the roots of (20), and hence of (19).

The following exercise gives you a chance to try out this method for yourself.

E22) Solve the following equations:

a) $x^4-3x^2-42x-40 = 0$

b) $4x^4-20x^3+33x^2-20x+4 = 0$

c) $x^4+12x = 5$

Let us now consider the other classical method for solving quartic equations.

3.4.2 Descartes' Solution

The second method for obtaining an algebraic solution for a quartic was given by the mathematician and philosopher René Descartes in 1637. In this method we write the biquadratic polynomial as a product of two quadratic polynomials. Then we solve the resultant quadratic equations to get the 4 roots of the original quartic.

Let us consider an example. In fact, let us solve the problem in Example 7 by this method. Thus, we want to solve

$$x^4-2x^3-5x^2+10x-3 = 0 \qquad \text{.........(21)}$$

Step 1: Remove the cube term. For this we rewrite x^4-2x^3 as

$$\left(x-\frac{1}{2}\right)^4-\frac{3}{2}x^2+\frac{1}{2}x-\frac{1}{16}.$$

Thus, the given equation becomes

$$\left(x-\frac{1}{2}\right)^4-\frac{13}{2}x^2+\frac{21}{2}x-\frac{49}{16} = 0$$

Now, put $x-\frac{1}{2} = y$. We get

$$y^4-\frac{13}{2}(y+\frac{1}{2})^2+\frac{21}{2}(y+\frac{1}{2})-\frac{49}{16} = 0$$

$$\Rightarrow y^4-\frac{13}{2}y^2+4y+\frac{9}{16} = 0. \qquad \text{......(22)}$$

Step 2 : Write the left hand side of (22) as a product of quadratic polynomials. For this, let us assume that

$$y^4-\frac{13}{2}y^2+4y+\frac{9}{16} = (y^2+ky+m)(y^2-ky+n).$$

(Note that the coefficients of y in each of these factors are k and $-k$, respectively, since the product does not contain any term with y^3.)

Equating coefficients, we get

$$m + n - k^2 = -\frac{13}{2},\ k(n-m) = 4,\ mn = \frac{9}{16} \qquad \text{......... (23)}$$

Eliminating m and n from these equations, we get

$$\left(k^2 - \frac{13}{2} - \frac{4}{k}\right)\left(k^2 - \frac{13}{2} + \frac{4}{k}\right) = \frac{9}{4}, \text{ that is,}$$

$$k^6 - 13k^4 + 40k^2 - 16 = 0.$$

If we put $k^2 = t$, then this becomes the resolvent cubic

$$t^3 - 13t^2 + 40t - 16 = 0.$$

This has one real root; in fact, it has a positive real root, because of the following result, that we give without proof.

> **Every polynomial equation, whose leading coefficient is 1 and degree is an odd number, has at least one real root whose sign is opposite to that of its last term.**

So, using this result, we see that we can expect to get one positive value of t. By trial, we see that $t = 4$ is a root, that is, $k^2 = 4$, that is, $k = \pm 2$. Any one of these values is sufficient for us. So let us take $k = 2$.

Then, from the equations in (23) we get

$$m = -\frac{9}{4},\ n = -\frac{1}{4}.$$

Thus, (22) is equivalent to

$$\left(y^2 + 2y - \frac{9}{4}\right)\left(y^2 - 2y - \frac{1}{4}\right) = 0.$$

Step 3 : Solve the quadratic equations

$$y^2 + 2y - \frac{9}{4} = 0 \text{ and } y^2 - 2y - \frac{1}{4} = 0.$$

By the quadratic formula we get

$$y = \frac{-2 \pm \sqrt{13}}{2} \text{ and } y = \frac{2 \pm \sqrt{5}}{2}.$$

Step 4 : Put these values in $x = y + \frac{1}{2}$ to get the four roots of (21).

Thus, the roots of (21) are :

$$\frac{-1+\sqrt{13}}{2},\ \frac{-1-\sqrt{13}}{2},\ \frac{3+\sqrt{5}}{2},\ \frac{3-\sqrt{5}}{2}.$$

Let us write down the steps in this method of solution for the general quartic equation

$$x^4 + ax^3 + bx^2 + cx + d = 0,\ a, b, c, d \in \mathbb{R}. \qquad \text{......... (24)}$$

Step 1: Reduce the equation to the form

$$x^4 + qx^2 + rx + s = 0. \qquad \text{......... (25)}$$

Step 2 : Assume that

$$x^4 + qx^2 + rx + s = (x^2 + kx + m)(x^2 - kx + n).$$

Then, on equating coefficients, we get

$$m + n - k^2 = q,\ k(n-m) = r,\ mn = s.$$

From these equations we get

$$m + n = k^2 + q,\ n - m = \frac{r}{k}.$$

Therefore, $2m = k^2 + q - \frac{r}{k}$, $2n = k^2 + q + \frac{r}{k}$.

Substituting in $mn = s$, we get

$(k^3 + qk - r)(k^3 + qk + r) = 4sk^2$, that is,

$k^6 + 2qk^4 + (q^2 - 4s)k^2 - r^2 = 0$, that is,

$t^3 + 2qt^2 + (q^2 - 4s)t - r^2 = 0$, putting $k^2 = t$.

This is a cubic with at least one positive real root. Then, with a known value of t, we can determine the values of k, m and n. So, (25) is equivalent to

$(x^2 + kx + m)(x^2 - kx + n) = 0$

Step 3 : Solve the quadratic equation:

$x^2 + kx + m = 0$ and $x^2 - kx + n = 0$.

This will give us the 4 roots of (25), and hence, the 4 roots of (24).

Now, why don't you try the following exercises to see if you have grasped Descartes' method ?

E23) Solve the following equations by Descartes' method:

a) $x^4 - 2x^2 + 8x - 3 = 0$

b) $x^4 + 8x^3 + 9x^2 - 8x = 10$

c) $x^4 - 3x^2 - 6x - 2 = 0$

d) $x^4 + 4x^3 - 7x^2 - 22x + 24 = 0$.

E24) Reduce the equation $2x^8 + 5x^6 - 5x^2 = 2$ to a biquadratic equation. Hence solve it.

While solving quartic equations you may have realised that the methods that we have discussed appear to be very easy to use; but, in practice, they can become quite cumbersome. This is because Cardano's method for solving a cubic often requires the use of a calculator.

Well, so far we have discussed methods of obtaining algebraic solutions for polynomial equations of degrees 1, 2, 3 and 4. You may think that we are going to do something similar for quintic equations, that is, equations of degree 5. But, in 1824 the Norwegian algebraist Abel (1802-1829) published a proof of the following result:

> There can be no general formula, expressed in explicit algebraic operations on the coefficients of a polynomial equation, for the roots of the equation, if the degree of the equation is greater than 4.

This result says that polynomial equations of degree > 4 do not have a general algebraic solution. But, there are methods that can give us the value of any real root to any required degree of accuracy. We will discuss these methods in our course on Numerical Analysis. There are, of course, special polynomial equations of degree ≥ 5 that can be solved (as in E24).

Let us now look a little closely at the roots of a biquadratic equation. We shall see how they are related to the coefficients of the equation, just as we did in the case of the cubic.

3.4.3 Roots And Their Relation With Coefficients

In the two previous sub-sections we have shown you how to explicitly obtain the 4 roots of a biquadratic equation. Let us go back to Theorems 1 and 3 for a moment. Theorem 1 tells us that a quartic has 4 roots, which may be real or complex. By Theorem 3, the possibilities are

i) all the roots are real, or

ii) two are real and two are complex conjugates of each other, or

iii) the roots are two pairs of complex conjugates, that is, $a + ib$, $a - ib$, $c + id$, $c - id$ for some $a, b, c, d \in \mathbb{R}$.

Now, if r_1, r_2, r_3, r_4 are the roots of the quartic $ax^4 + bx^3 + cx^2 + dx + e = 0$, then

$$ax^4 + bx^3 + cx^2 + dx + e = a(x - r_1)(x - r_2)(x - r_3)(x - r_4)$$
$$= x^4 - (r_1 + r_2 + r_3 + r_4)x^3 + (r_1 r_2 + r_1 r_3 + r_1 r_4 + r_2 r_3 + r_2 r_4 + r_3 r_4)x^2$$
$$- (r_1 r_2 r_3 + r_1 r_2 r_4 + r_1 r_3 r_4 + r_2 r_3 r_4)x + r_1 r_2 r_3 r_4.$$

Comparing the coefficients, we see that

$$r_1 + r_2 + r_3 + r_4 = -\frac{b}{a}$$

$$r_1 r_2 + r_1 r_3 + r_1 r_4 + r_2 r_3 + r_2 r_4 + r_3 r_4 = \frac{c}{a}$$

$$r_1 r_2 r_3 + r_1 r_2 r_4 + r_1 r_3 r_4 + r_2 r_3 r_4 = -\frac{d}{a}$$

$$r_1 r_2 r_3 r_4 = \frac{e}{a}$$

This means that

coeff. is short for coefficient.

$$\text{sum of the roots} = -\frac{\text{coeff. of } x^3}{\text{coeff. of } x^4}$$

$$\text{sum of the roots taken two at a time} = \frac{\text{coeff. of } x^2}{\text{coeff. of } x^4}$$

$$\text{sum of the roots taken three at a time} = -\frac{\text{coeff. of } x}{\text{coeff. of } x^4}$$

$$\text{product of the roots} = \frac{\text{coeff. of } x^0}{\text{coeff. of } x^4}, \text{ that is, } \frac{\text{constant term}}{\text{coeff. of } x^4}.$$

These four equations are a particular case of the following result that relates the roots of a polynomial equation with its coefficients.

Theorem 4: Let $\alpha_1, \ldots, \alpha_n$ be the n roots of the equation

$a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0$, $a_i \in \mathbf{R}\ \forall\ i = 0, 1, \ldots, n$, $a_0 \neq 0$. Then

$\sum_{i=1}^{n} A_i = A_1 + \ldots + A_n$

$$\sum_{i=1}^{n} \alpha_i = -\frac{a_1}{a_0}$$

$$\sum_{\substack{i,j=1 \\ i<j}}^{n} \alpha_i \alpha_j = \frac{a_2}{a_0}$$

$$\vdots$$

$$\sum_{i_1 < i_2 < \ldots < i_r}^{n} \alpha_{i_1} \alpha_{i_2} \ldots \alpha_{i_r} = (-1)^r \frac{a_r}{a_0}$$

$$\vdots$$

$\prod_{i=1}^{n} A_i = A_1 A_2 \ldots A_n$

$$\prod_{i=1}^{n} \alpha_i = (-1)^n \frac{a_n}{a_0}$$

In E10 and E18 you have already seen that this result is true for n = 2 and 3.

Theorem 4 is very useful in several ways. Let us consider an application in the case n = 4.

Example 8: If the sum of two roots of the equation

$$4x^4 - 24x^3 + 31x^2 + 6x - 8 = 0$$

is zero, find all the roots of the equation.

Solution : Let the roots be a, b, c, d, where a+b = 0.

$$\text{Then } a + b + c + d = \frac{24}{4} = 6 \qquad \ldots\ldots (26)$$

$$\therefore\ c + d = 6$$

Also $ab + ac + ad + bc + bd + cd = (a + b)(c + d) + ab + cd = \frac{31}{4}$

$\therefore \; ab + cd = \frac{31}{4}$ ....(27)

Further, $(a + b)\,cd + ab\,(c + d) = acd + bcd + abc + abd =$

$\therefore$ (26) $\Rightarrow ab = -\frac{1}{4}$ ....(28)

Finally, $abcd = -2$

$\therefore$ (28) $\Rightarrow cd = 8$ ....(29)

Now using E11, (26) and (29) tell us that c and d are roots of $x^2 - 6x + 8 = 0$.

Thus, by the quadratic formula, $c = 2$, $d = 4$.

Similarly, a and b are roots of $x^2 - \frac{1}{4} = 0$. $\therefore$ $a = \frac{1}{2}$, $b = -\frac{1}{2}$.

Thus, the roots of the given quartic are

$\frac{1}{2}$, $-\frac{1}{2}$, 2, 4.

Try the following problems now.

E25) Solve the equation

$x^4 + 15x^3 + 70x^2 + 120x + 64 = 0,$

given that the roots are in G.P., i.e., geometrical progression.

(**Hint:** If four numbers a,b,c,d are in G.P., then ad = bc.)

E26) Show that if the sum of two roots of $x^4 - px^3 + qx^2 - rx + s = 0$ (where $p, q, r, s \in \mathbb{R}$) equals the sum of the other two, then $p^3 - 4pq + 8r = 0$.

We have touched upon relations between roots and coefficients for $n = 2, 3, 4$. But you can apply Theorem 4 for any $n \in N$. So, in future whenever you need to, you can refer to this theorem and use its result for equations of degree ≥ 5.

Let us now wind up this unit with a summary of what we have done in it.

3.5 SUMMARY

In this unit we have introduced you to the theory of lower degree equations. Specifically, we have covered the following points:

1) The linear equation $ax + b = 0$ has one root, namely, $x = \frac{-b}{a}$.

2) The quadratic equation $ax^2 + bx + c = 0$ has 2 roots given by the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

3) Every polynomial equation of degree n over $\mathbb{R}$ has n roots in C.

4) If $a + ib \in C$ is a root of a real polynomial, then so is $a - ib$.

5) Cardano's method for solving a cubic equation.

6) A cubic equation can have:

   i) three distinct real roots, or

   ii) one real root and two complex roots, which are conjugates, or

   iii) three real roots, of which exactly two are equal, or

   iv) three real roots, all of which are equal.

7) Methods due to Ferrari and Descartes for solving a quartic equation. Both these methods require the solving of one cubic and two quadratic equations.

8) A quartic equation can have four real roots, or two real and two complex roots, or 4 complex roots.

9) If the n roots of the nth degree equation $a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n = 0$, are $p_1, p_2, \dots, p_n$, then

$$\sum_{i=1}^{n} p_i = -\frac{a_1}{a_0}$$

$$\sum_{\substack{i<j \\ i,j=1,\dots,n}} p_i p_j = \frac{a_2}{a_0}$$

$$\vdots \qquad \vdots$$

$$\prod_{i=1}^{n} p_i = (-1)^n \frac{a_n}{a_0}$$

That is, the sum of the product of the roots taken k at a time is

$$(-1)^k \frac{a_k}{a_0} \forall k = 1, \dots, n.$$

As in our other units, we have given our solutions and/or answers to the exercises in the unit in the following section. You can go through them if you like. After that please go back to Section 3.1 and see if you have achieved the objectives.

3.6 SOLUTIONS/ANSWERS

E1) a) This has a solution provided $J \neq k$.

$$J\left(\frac{x}{k}+a\right) = x \Leftrightarrow x\left(\frac{J}{k}-1\right)+Ja = 0 \Leftrightarrow x = \frac{Ja}{\left(\frac{J}{k}-1\right)} = \frac{Jak}{J-k}$$

b) $\frac{1}{R} = \frac{r_1 + r_2}{r_1 r_2} \Leftrightarrow R = \frac{r_1 r_2}{r_1 + r_2}$

c) $F = \frac{9}{5}C + 32$

E2) Let the third side be x cm.

Then the other two sides are each 2x cm long.

Therefore, $x + 2x + 2x = 30 \Rightarrow x = 6$.

Thus, the lengths of the sides are 6 cm, 12 cm and 12 cm.

E3) Let her rate of travel to the study centre be x km per hour. Thus, the distance from her home to the study centre is $\frac{x}{3}$ km. While returning, her rate is $(x-8)$ km/hr.

$$\therefore \frac{1}{2}(x-8) = \frac{x}{3} \Rightarrow x = 24.$$

Thus, the distance is $\frac{24}{3}$ km $= 8$ km.

E4) $S = k \cdot P t$.

We know that $110 = k \times 1000 \times 2 \Rightarrow k = \frac{11}{200}$.

$$\therefore S = \frac{11}{200} Pt.$$

So, the required interest is

$\frac{11}{200} \times 5000 \times 3$, that is, Rs. 825/-.

E5) True. For example, $x^2 + 1 = 0$ has complex roots. Any linear equation $ax + b = 0$ over $\mathbf{R}$ has only one root, namely, $\frac{-b}{a} \in \mathbf{R}$.

E6) a) $x^2 = -5 \Rightarrow x = i\sqrt{5}$ and $-i\sqrt{5}$.

b) This is $(x-(-9))(x-1) = 0$. Thus, by Remark 2, -9 and 1 are the roots.

c) We rewrite the given equation in standard form as

$$x^2 - \sqrt{5}\,x - 1 = 0$$

$$\Rightarrow x = \frac{\sqrt{5} \pm \sqrt{5+4}}{2} = \frac{\sqrt{5} \pm 3}{2}$$

$$\therefore x = \frac{\sqrt{5}+3}{2} \text{ and } \frac{\sqrt{5}-3}{2}.$$

d) $m = \dfrac{8q \pm \sqrt{64q^2 - \frac{4p}{r}}}{2p} = \dfrac{4q}{p} \pm \dfrac{1}{p}\sqrt{16q^2 - \dfrac{p}{r}}$

E7) The roots are

$$x = \frac{-(2k+6) \pm \sqrt{(2k+6)^2 - 64k}}{2k}$$

The roots will coincide if the discriminant is zero, that is, $(2k+6)^2 - 64k = 0$.

This will happen when $k^2 - 10k + 9 = 0$, that is,

$k = 1$ or $k = 9$.

E8) $(2ax + b) \mid (ax^2 + bx + c)$

$\Rightarrow$ the root of $2ax + bx = 0$ is a root of $ax^2 + bx + c = 0$.

$\Rightarrow x = \frac{-b}{2a}$ is a root of $ax^2 + bx + c = 0$.

$$\Rightarrow a\left(\frac{-b}{2a}\right)^2 + b\left(\frac{-b}{2a}\right) + c = 0.$$

$\Rightarrow b^2 - 4ac = 0$

$\Rightarrow ax^2 + bx + c = 0$ has coincidental roots.

E9) By Remark 2, we must have

$$x^2 + bx + c = (x-(1+i))(x-(1-i))$$
$$= x^2 - 2x + 2.$$

Thus, comparing the coefficients of x^1 and x^0, we get

$b = -2$, $c = 2$.

E10) α and β are roots of $ax^2 + bx + c = 0$

$\Leftrightarrow ax^2 + bx + c = a(x-\alpha)(x-\beta)$

$\Leftrightarrow ax^2 + bx + c = a\{x^2 - (\alpha+\beta)x + \alpha\beta\}$

$\Leftrightarrow b = -a(\alpha+\beta)$ and $c = a\alpha\beta$

$\Leftrightarrow \alpha + \beta = -\frac{b}{a}$ and $\alpha\beta = \frac{c}{a}$.

E11) Substituting $x = \alpha$ in $x^2 - px + q$, we get

$\alpha^2 - p\alpha + q = \alpha^2 - (\alpha+\beta)\alpha + \alpha\beta$, since $\alpha + \beta = p$ and $\alpha\beta = q$,

$= 0$

$\therefore$ α is a root of $x^2 - px + q = 0$.

Similarly, β is a root of $x^2 - px + q = 0$.

E12) a) $4p^4 - 16p^2 + 5 = 0.$

Put $p^2 = x$. Then the equation becomes

$4x^2 - 16x + 5 = 0.$

Its roots are $2 + \frac{\sqrt{11}}{2}$ and $2 - \frac{\sqrt{11}}{2}$

Now, $p^2 = 2 + \frac{\sqrt{11}}{2} \Rightarrow p = \pm\sqrt{2 + \frac{\sqrt{11}}{2}}$

and $p^2 = 2 - \frac{\sqrt{11}}{2} \Rightarrow p = \pm\sqrt{2 - \frac{\sqrt{11}}{2}}.$

These 4 values of p are the required roots.

b) $(5x^2 - 6)^{\frac{1}{4}} = x.$

Every root of this is a root of

$5x^2 - 6 = x^4$

$\Rightarrow x^4 - 5x^2 + 6 = 0.$

Put $x^2 = y$. Then

$y^2 - 5y + 6 = 0.$

Its roots are $y = \frac{5 \pm \sqrt{25 - 24}}{2} = \frac{5 \pm 1}{2} = 3, 2.$

Now $x^2 = 3 \Rightarrow x = \sqrt{3}$ or $-\sqrt{3}$,

and $x^2 = 2 \Rightarrow x = \sqrt{2}$ or $-\sqrt{2}$.

Putting these 4 values of x in the given equation, we find that $\sqrt{3}$ and $\sqrt{2}$ are its solutions.

c) Separating the radicals, we get

$\sqrt{2x + 3} = 1 + \sqrt{x + 1}.$

Squaring both sides, we get

$2x + 3 = 1 + (x + 1) + 2\sqrt{x + 1}$

$\Leftrightarrow x + 1 = 2\sqrt{x + 1}.$

Again squaring both sides, we get

$x^2 - 2x - 3 = 0.$

Its roots are $x = 3$ and $x = -1$.

Substituting these values of x in the given equation, we get

$\sqrt{2(3) + 3} - \sqrt{3 + 1} = 1$, and $\sqrt{2(-1) + 3} - \sqrt{-1 + 1} = 1.$

Thus, both $x = 3$ and $x = -1$ are roots of the given equation.

E13) Let Alka's rate be x km per hour. Then Ameena's is $(x + 1)$ km per hour.

The time taken by Ameena to walk to the library $= \frac{24}{x + 1}$ hours. Thus, the time taken by Alka $= \left(\frac{24}{x + 1} + 2\right)$ hours.

$\therefore \frac{24}{x} = \frac{24}{x + 1} + 2$

$\Rightarrow x(x + 1) = 12$

$\Rightarrow x = -4$ or $x = 3.$

Since (-4) can't be the rate, it is an extraneous solution. Thus, the required speed must be 3 km per hour.

E14) None, since $ax + b = 0, a, b \in \mathbf{R} \Rightarrow x = -\frac{b}{a} \in \mathbf{R}$.

E15) If $p^2 - 4q < 0$.

There will be two such roots, and they will be conjugates.

E16) a) $2x^3 + 3x^2 + 3x + 1 = 0$

$\Leftrightarrow x^3 + \frac{3}{2}x^2 + \frac{3}{2}x + \frac{1}{2} = 0.$

Referring to Cardano's formula, we see that in this case

$p = \frac{3}{2}, \quad q = \frac{3}{2}, \quad r = \frac{1}{2}.$

$\therefore A = \frac{3}{4}, B = 0.$

$\therefore \alpha = \frac{1}{2}, \beta = -\frac{1}{2}.$

$\therefore$ the roots are $-\frac{1}{2}, \frac{\omega - \omega^2}{2} - \frac{1}{2}, \frac{\omega^2 - \omega}{2} - \frac{1}{2}$, that is,

$-\frac{1}{2}, \omega, \omega^2$ (since $1 + \omega + \omega^2 = 0$).

b) $x^3 + 21x + 342 = 0.$

Here we don't need to apply Step 1 of Cardano's method, since there is no term containing x^2. Now, with reference to Cardano's formula,

$A = 21 - 0 = 21, \ B = 0 - 0 + 342 = 342.$

$$\therefore \alpha = \left\{ \frac{-342}{2} + \sqrt{\frac{(342)^2}{2} + \frac{(21)^3}{27}} \right\}^{\frac{1}{3}} = (-171 + 172)^{\frac{1}{3}} = 1$$

and $\beta = (-171 - 172)^{\frac{1}{3}} = -7.$

Thus, the roots of the equation are

$1 - 7, \ \omega - 7\omega^2, \ \omega^2 - 7\omega$, that is, $-6, \ \omega - 7\omega^2, \ \omega^2 - 7\omega$.

c) $x^3 + 6x^2 + 6x + 8 = 0.$

Here $p = 6, \ q = 6, \ r = 8.$

$\therefore A = q - \frac{p^2}{3} = -6, \ B = \frac{2p^3}{27} - \frac{pq}{3} + r = 12.$

$\therefore \alpha = (-6 + \sqrt{36 - 8})^{\frac{1}{3}} = (-6 + 2\sqrt{7})^{\frac{1}{3}} = -0.891$

and $\beta = (-6 - 2\sqrt{7})^{\frac{1}{3}} = -2.243.$

(We have used a calculator to evaluate α and β to 3 decimal places.)

Then the required roots are

$\alpha + \beta - 2, \ \alpha\omega + \beta\omega^2 - 2, \ \alpha\omega^2 + \beta\omega - 2.$

d) $x^3 + 29x - 97 = 0.$

Here $p = 0, \ q = 29, \ r = -97.$

$$\therefore \alpha = \left\{ \frac{97}{2} + \sqrt{\frac{(97)^2}{4} + \frac{(29)^3}{27}} \right\}^{\frac{1}{3}} = (54.557)^{\frac{1}{3}} = 4.01, \text{ and}$$

$\beta = (-8.557)^{\frac{1}{3}} = -2.045.$

Then the roots are

$\alpha + \beta, \ \alpha\omega + \beta\omega^2, \ \alpha\omega^2 + \beta\omega.$

c) $x^3 - 30x + 133 = 0$.

Here $p = 0,\ q = -30,\ r = 133$.

$\therefore A = -30, B = 133$.

$$\therefore \alpha = \left\{ -\frac{133}{2} + \sqrt{\frac{(133)^2}{4} - \frac{(30)^3}{27}} \right\}^{\frac{1}{3}} = (-8)^{\frac{1}{3}} = -2, \text{ and}$$

$\beta = (-66.5 - 58.5)^{\frac{1}{3}} = -5$.

$\therefore$ the roots are $-7,\ -2\omega - 5\omega^2,\ -2\omega^2 - 5\omega$.

E17) $x^3 - 3x + 1 = 0$.

Here $p = 0,\ q = -3,\ r = 1$.

$\therefore A = -3,\ B = 1$.

$\therefore \dfrac{B^2}{4} + \dfrac{A^3}{27} = -\dfrac{3}{4} < 0$.

So we are in the irreducible case.

$$\text{Now, } \frac{-B}{2} + \sqrt{\frac{B^2}{4} + \frac{A^3}{27}} = \frac{-1}{2} + \frac{i\sqrt{3}}{2} = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}.$$

Thus, the solutions of the given equation are $2\cos\left(\dfrac{\frac{2\pi}{3} + 2k\pi}{3}\right)$, where $k = 0, 1, 2$, that is,

$2\cos\dfrac{2\pi}{9},\ 2\cos\dfrac{8\pi}{9},\ 2\cos\dfrac{14\pi}{9}$.

E18) $\alpha,\ \beta,\ \gamma$ are the roots iff

$$ax^3 + bx^2 + cx + d = a(x-\alpha)(x-\beta)(x-\gamma)$$
$$= a\,(x^3 - (\alpha+\beta+\gamma)x^2 + (\alpha\beta + \beta\gamma + \alpha\gamma)x + \alpha\beta\gamma).$$

On comparing coefficients, we get

$$\alpha + \beta + \gamma = -\frac{b}{a}$$

$$\alpha\beta + \beta\gamma + \alpha\gamma = \frac{c}{a}$$

$$\alpha\beta\gamma = -\frac{d}{a}$$

E19) Let the roots be α, β, γ.

Then $\alpha + \beta + \gamma = 6,\ \alpha\beta + \beta\gamma + \alpha\gamma = 11,\ \alpha\beta\gamma = 6$.

$$\therefore \alpha^2 + \beta^2 + \gamma^2 = (\alpha+\beta+\gamma)^2 - 2(\alpha\beta + \beta\gamma + \alpha\gamma)$$
$$= 36 - 22 = 14.$$

$$\therefore \alpha^3 + \beta^3 + \gamma^3 = (\alpha+\beta+\gamma)^3 - 3\alpha^2(\beta+\gamma) - 3\beta^2(\alpha+\gamma)$$
$$- 3\gamma^2(\alpha+\beta) - 6\alpha\beta\gamma$$
$$= 6^3 - 3\alpha^2(6-\alpha) - 3\beta^2(6-\beta) - 3\gamma^2(6-\gamma) - 6\times 6$$
$$= 180 - 18(\alpha^2 + \beta^2 + \gamma^2) + 3(\alpha^3 + \beta^3 + \gamma^3)$$

$$\therefore 4(\alpha^3 + \beta^3 + \gamma^3) = 180 - 18\times 14 = -72$$

$$\therefore \alpha^3 + \beta^3 + \gamma^3 = -18.$$

Now, each of α, β, γ satisfy

$x^3 - 6x^2 + 11x - 6 = 0$.

Thus, they will satisfy $x^4 - 6x^3 + 11x^2 - 6x = 0$ also.

$\therefore\ \alpha^4 - 6\alpha^3 + 11\alpha^2 - 6\alpha = 0$

$\beta^4 - 6\beta^3 + 11\beta^2 - 6\beta = 0$

$\gamma^4 - 6\gamma^3 + 11\gamma^2 - 6\gamma = 0.$

Adding these equations, we get

$$(\alpha^4 + \beta^4 + \gamma^4) - 6(\alpha^3 + \beta^3 + \gamma^3) + 11(\alpha^2 + \beta^2 + \gamma^2) - 6(\alpha + \beta + \gamma) = 0$$

$$\Rightarrow \alpha^4 + \beta^4 + \gamma^4 = 6(-18) - 11(14) + 6(6) = -226.$$

E20) Here $p = \frac{3b}{a}$, $q = \frac{3c}{a}$, $r = \frac{d}{a}$.

$$\therefore\ A = \frac{3c}{a} - \frac{1}{3}\left(\frac{3b}{a}\right)^2 = \frac{3(ac - b^2)}{a^2}, \text{ and}$$

$$B = \frac{2}{27}\left(\frac{3b}{a}\right)^3 - \frac{1}{3}\left(\frac{3b}{a}\right)\left(\frac{3c}{a}\right) + \frac{d}{a} = \frac{2b^3 - 3abc + a^2 d}{a^3}.$$

Therefore, the equation has complex roots if

$\frac{B^2}{4} + \frac{A^3}{27} > 0$, that is,

$\frac{(2b^3 - 3abc + a^2 d)^2}{4a^6} + \frac{(ac - b^2)^3}{a^6} > 0$, that is,

$a^2 d^2 - 3b^2 c^2 - 6abcd + 4b^3 d + 4ac^2 > 0.$

E21) Here $B = -126$, $A = -15$.

$$\therefore\ \frac{B^2}{4} + \frac{A^3}{27} = 3844 > 0.$$

Thus, the equation has 1 real and 2 complex roots.

E22) a) $x^4 - 3x^2 - 42x - 40 = 0.$

Adding $(ax + b)^2$ to both sides we get

$x^4 + (a^2 - 3)x^2 + (2ab - 42)x + b^2 - 40 = (ax + b)^2$

Assume that the left hand side is $(x^2 + k)^2$.

(Note that the coefficient of x^3 in the given equation is 0.)

Then $x^4 + (a^2 - 3)x^2 + (2ab - 42)x + b^2 - 40 = x^4 + k^2 + 2kx^2.$

Comparing coefficients, we get

$$\left.\begin{array}{l} a^2 - 3 = 2k \\ 2ab - 42 = 0 \\ b^2 - 40 = k^2 \end{array}\right\} \Rightarrow \begin{array}{l} a^2 = 2k + 3 \\ ab = 21 \\ b^2 = k^2 + 40. \end{array}$$

Eliminating a and b we get

$(21)^2 = (2k + 3)(k^2 + 40) = 2k^3 + 3k^2 + 60k + 120$

$\therefore\ 2k^3 + 3k^2 + 80k - 321 = 0.$

3 is a root of this equation. With this value of k we get

$a^2 = 9$, $b^2 = 49$, $ab = 21$.

These equations are satisfied by $a = 3$, $b = 7$.

Thus, solving the given quartic reduces to solving the following quadratic equations :

$x^2 + 3 = 3x + 7$ and $x^2 + 3 = -(3x + 7)$, that is,

$x^2 - 3x - 4 = 0$ and $x^2 + 3x + 10 = 0$.

Thus, the required roots are $4, -1$ and $\frac{-3 \pm i\sqrt{31}}{2}$.

b) The given equation is equivalent to

$$x^4 - 5x^3 + \frac{33}{4}x^2 - 5x + 1 = 0.$$

The resolvent cubic is $8k^3 - 33k^2 + 42k - 17 = 0.$

One real root is 1.

With this value of k, we find that

$a = 0, \ b = 0.$

Thus, the given equation becomes

$$(x^2 - \frac{5}{2}x + 1)^2 = 0.$$

Therefore, the given equation has the roots

$2, \frac{1}{2}, 2, \frac{1}{2}$, that is, two pairs of equal roots.

c) $x^4 + 12x - 5 = 0$. The resolvent cubic is $k^3 + 5k - 18 = 0.$

A real root is $k = 2$.

Then, solving the given equation reduces to solving

$(x^2 + 2) = \pm(2x - 3)$, that is

$x^2 - 2x + 5 = 0$ and $x^2 + 2x - 1 = 0.$

Thus, the required roots are

$\frac{2 \pm \sqrt{4-20}}{2}$ and $\frac{-2 \pm \sqrt{4+4}}{2}$, that is,

$1 + 2i, \ 1 - 2i, \ -1 + \sqrt{2}, \ -1 - \sqrt{2}.$

E23) a) $x^4 - 2x^2 + 8x - 3 = 0.$

Since there is no x^3 term, we don't need to apply Step 1. Now assume

$x^4 - 2x^2 + 8x - 3 = (x^2 + kx + m)(x^2 - kx + n).$

Then $m + n - k^2 = -2, \ k(n - m) = 8, \ mn = -3.$

Thus, eliminating m and n, we get

$k^6 - 4k^4 + 16k^2 - 64 = 0.$

$k^2 = 4$ is a root of this cubic in k^2.

Thus, $k = 2$ is a solution. For this value of k, we get $n = 3, m = -1.$

Thus, the roots of the given equation are the roots of $x^2 + 2x - 1 = 0$ and $x^2 - 2x + 3 = 0$, that is, the roots are $-1 \pm \sqrt{2}$ and $-1 \pm i\sqrt{2}.$

b) This equation can be rewritten as

$(x + 2)^4 - 15x^2 - 40x - 26 = 0.$

Putting $x + 2 = y$, we get

$y^4 - 15y^2 + 20y - 6 = 0.$

Then the cubic in k^2 is

$k^6 - 30k^4 + 249k^2 - 400 = 0$, that is,

$t^3 - 30t^2 + 249t - 400 = 0$, putting $k^2 = t.$

One real positive root is $t = 16$. So we can take $k = 4$.

Then we need to solve the quadratic equations

$y^2 - 4y + 3 = 0$ and $y^2 + 4y - 2 = 0.$

Thus, $y = 3, \ 1, \ -2 \pm \sqrt{6}.$

Thus, the roots of the given equation are $(y-2)$, that is, $1, \ -1, -4 \pm \sqrt{6}.$

c) $x^4 - 3x^2 - 6x - 2 = 0$.

The cubic in k^2 is $k^6 - 6k^4 + 17k^2 - 36 = 0$.

$k^2 = 4$ is a root. So we can take $k = 2$.

Then we need to solve the equations

$x^2 + 2x + 2 = 0$, $x^2 - 2x - 1 = 0$.

$\therefore x = -1 \pm i, 1 \pm \sqrt{2}$.

d) 1, 2, –3, – 4.

E24) Putting $x^2 = y$ in the equation, we get

$2y^4 + 5y^3 - 5y - 2 = 0$.

Then, by either Ferrari's or Descartes' method, we can find the four values of y, which are $1, -1, -2, \frac{-1}{2}$.

Putting these values in $x^2 = y$, and solving, we get the 8 roots of the given equation. Thus, the required roots are

$\pm\sqrt{1}, \pm\sqrt{-1}, \pm\sqrt{-2}, \pm\sqrt{\frac{-1}{2}}$, that is,

$1, -1, i, -i, i\sqrt{2}, -i\sqrt{2}, \frac{i}{\sqrt{2}}, \frac{-i}{\sqrt{2}}$.

E25) Let the roots be a, b, c, d. Then ad = bc.

Now, we know that

i) $a + b + c + d = -15 \Rightarrow (a + d) + (b + c) = -15$ ……… (30)

ii) $ab + ac + ad + bc + bd + cd = 70$
$\Rightarrow (a + d)(b + c) + ad + bc = 70$ ……… (31)

iii) $abc + abd + acd + bcd = -120$ ……… (32)

iv) $abcd = 64$

Now, (32) $\Rightarrow ad(b + c) + bc(a + d) = -120 \Rightarrow ad(a + b + c + d) = -120$
$\Rightarrow -15\, ad = -120 \Rightarrow ad = 8$. Thus, $ad = 8 = bc$.

Then (31) $\Rightarrow (a + d)(b + c) = 70 - 16 = 54$.

This, with (30) tells us that a + d and b + c are roots of $x^2 + 15x + 54 = 0$.

Thus, by the quadratic formula,

$a + d = \frac{-15 + 3}{2} = -6$ and $b + c = \frac{-15 - 3}{2} = -9$.

Then, ad = 8 and bc = 8 tell us that a and d are zeros of $x^2 + 6x + 8 = 0$,

and b and c are zeros of $x^2 + 9x + 8 = 0$.

$\therefore a = \frac{-6 + 2}{2} = -2, d = -4, b = \frac{-9 + 7}{2} = -1, c = -8$.

E26) Let the roots be a, b, c, d, where

$a + b = c + d$. ……… (33)

We know that

$a + b + c + d = p$ ……… (34)

$(a + b)(c + d) + ab + cd = q$ ……… (35)

$ab(c + d) + (a + b)cd = r$ ……… (36)

$abcd = s$

(33) and (34) $\Rightarrow a + b = \frac{p}{2} = c + d.$

Then (36) $\Rightarrow \frac{p}{2}(ab + cd) = r \Rightarrow ab + cd = \frac{2r}{p}.$

Then (35) $\Rightarrow \frac{p}{2}.\frac{p}{2} + \frac{2r}{p} = q.$

$\Rightarrow p^3 + 8r = 4pq.$

$\Rightarrow p^3 - 4pq + 8r = 0.$

MISCELLANEOUS EXERCISES

This section is optional.

We have listed some problems related to the material covered in this block. You may like to do them to get more practice in solving problems. We have also given our solutions to these questions, because you may like to counter-check your answers.

1) Let 1, ω and ω^2 be the cube roots of unity. Evaluate

   a) $(1-\omega+\omega^2)(1+\omega-\omega^2)$

   b) $(1-\omega)(1-\omega^2)(1-\omega^4)(1-\omega^5)$

   c) $\prod_{i=1}^{5}(1-\omega^i)$

2) Give the equations, in standard form, whose roots are

   a) $2, -\frac{3}{2}, 9$

   b) $\sqrt{2}, -\sqrt{3}, \pm i$.

3) For what value of m ($\neq -1$) will the equation $\frac{x^2-bx}{ax-c} = \frac{m-1}{m+1}$ have roots equal in magnitude but opposite in sign? Here $a, b \in \mathbb{R}$ and $a+b \neq 0$.

4) Solve

   a) $2\sqrt{\frac{x}{a}} + 3\sqrt{\frac{a}{x}} = \frac{b}{a} + \frac{6a}{b}$, where $a, b \in \mathbb{R}$.

   b) $\sqrt{2}^x + \frac{1}{\sqrt{2}^x} = 2$.

5) $x^4 + 9x^3 + 12x^2 - 80x - 192 = 0$ has a pair of equal roots. Obtain all its roots.

6) If a, b, c are the roots of $x^3 - px^2 + r = 0$, find the equation whose roots are $\frac{b+c}{a}, \frac{c+a}{b}, \frac{a+b}{c}$

7) Solve $x^4 - 4x^2 + 8x + 35 = 0$, given that one root is $2 + \sqrt{-3}$.

8) Form the cubic whose roots are a, b, c, where

   $a + b + c = 3$,

   $a^2 + b^2 + c^2 = 5$, and

   $a^3 + b^3 + c^3 = 11$.

   Hence evaluate $a^4 + b^4 + c^4$.

9) Find the equation whose roots are 4 less in value than the roots of

   $x^4 - 5x^3 + 7x^2 - 17x + 11 = 0$.

   (Hint : Write the equation as an equation in $(x-4)$.)

10) Form the polynomial equation over R of lowest degree which is satisfied by $1-i$ and $3+2i$. Is it unique ?

11) Solve $x^4 + 9x^3 + 16x^2 + 9x + 1 = 0$.

   (Hint: Note that in this equation the coefficients of x^r and x^{4-r} are the same $\forall r = 0, 1, 2, 3, 4$. So we can divide throughout by x^2 and then write the equation as a quadratic in $x + \frac{1}{x} = y$, say. Now solve for y, and then for x.)

If you are interested in doing more exercises on the material covered in this block, please refer to the book 'Higher Algebra' by Hall and Knight. A copy is available in your study centre.

Solutions

1) a) We know that $1 + \omega + \omega^2 = 0$.

$\therefore \quad 1 - \omega + \omega^2 = -2\omega$ and $1 + \omega - \omega^2 = -2\omega^2$.

$\therefore \quad (1 - \omega + \omega^2)(1 + \omega - \omega^2) = (-2\omega)(-2\omega^2) = 4\omega^3 = 4$, since $\omega^3 = 1$.

b) $(1-\omega)(1-\omega^2)(1-\omega^4)(1-\omega^5)$

$= (1-\omega)^2(1-\omega^2)^2$, since $\omega^3 = 1$.

$= (1 - 2\omega + \omega^2)(1 - 2\omega^2 + \omega^4)$

$= (-3\omega)(-3\omega^2)$

$= 9.$

c) 0, since $1 - \omega^3 = 0$.

2) a) The equation is

$$(x-2)\left(x+\frac{3}{2}\right)(x-9) = 0$$

$\Leftrightarrow (x-2)(2x+3)(x-9) = 0$

$\Leftrightarrow 2x^3 - 19x^2 + 3x + 54 = 0$, in standard form.

b) $x^4 + (\sqrt{3} - \sqrt{2})x^3 + (1 - \sqrt{6})x^2 + (\sqrt{3} - \sqrt{2})x - \sqrt{6} = 0.$

3) The equation is equivalent to

$(1+m)x^2 + \{(a-b) - m(a+b)\}x + (m-1)c = 0.$

The roots are equal in magnitude and opposite in sign iff their sum is zero.

$\therefore \quad (a-b) - m(a+b) = 0$, that is,

$m = \frac{a-b}{a+b}.$

4) a) Let $y = \sqrt{\frac{x}{a}}$. Then the given equation is equivalent to

$2y + \frac{3}{y} = \frac{b^2 + 6a^2}{ab}$, that is,

$2y^2 - \frac{b^2+6a^2}{ab}y + 3 = 0.$

Its roots are $\frac{b}{2a}, \frac{3a}{b}$.

Thus, the roots of the given equation are

$\left(\sqrt{a}\,\frac{b}{2a}\right)^2$ and $\left(\sqrt{a}\,\frac{3a}{b}\right)^2$, that is, $\frac{b^2}{4a}$ and $\frac{9a^3}{b^2}$.

b) Put $y = \sqrt{2}^x$. Then our equation in y is

$y + \frac{1}{y} = 2 \quad \Leftrightarrow y^2 - 2y + 1 = 0 \quad \Leftrightarrow (y-1)^2 = 0.$

$\therefore \quad y = 1$

$\therefore \quad \sqrt{2}^x = 1$

$\therefore \quad x = 0$ is the required root.

5) Let its roots be a, a, b, c.

Then, the relations between the roots and the coefficients are

$2a + b + c = -9$ ...... (1)

$a^2 + 2ab + 2ac + bc = 12$ ...... (2)

$a^2 b + a^2 c + 2abc = 80$ ...... (3)

$a^2 bc = -192$ ...... (4)

Using (1), (2) and (3), we get

$4a^3 + 27a^2 + 24a = 80.$

$a = -4$ is a solution.

Then (1) $\Rightarrow b + c = -1$, and (4) $\Rightarrow bc = -12.$

Thus, b and c are roots of

$x^2 + x - 12 = 0.$

$\therefore b = 3,\ c = -4.$

Thus, the roots are $-4, -4, -4, 3$.

6) We know that

$a + b + c = p$ ...... (5)

$ab + ac + bc = 0$ ...... (6)

$abc = r$ ...... (7)

Now, $\frac{b+c}{a} + \frac{c+a}{b} + \frac{a+b}{c} = \frac{bc(p-a) + ac(p-b) + ab(p-c)}{abc}$, using (5)

$= -3$ ......(8)

$\left(\frac{b+c}{a}\right)\left(\frac{c+a}{b}\right) + \left(\frac{c+a}{b}\right)\left(\frac{a+b}{c}\right) + \left(\frac{a+b}{c}\right)\left(\frac{b+c}{a}\right) = \frac{a^3+b^3+c^3}{abc}$, using (6).

Now, a, b and c satisfy $x^3 - px^2 + r = 0.$

Thus, on substituting each of a, b and c in this equation, and summing up, we get

$a^3 + b^3 + c^3 - p(a^2 + b^2 + c^2) + 3r = 0.$

$\therefore a^3 + b^3 + c^3 = p\{(a+b+c)^2 - 2(ab + ac + bc)\} - 3r$

$= p^3 - 3r.$

$\therefore \left(\frac{b+c}{a}\right)\left(\frac{c+a}{b}\right) + \left(\frac{c+a}{b}\right)\left(\frac{a+b}{c}\right) + \left(\frac{a+b}{c}\right)\left(\frac{b+c}{a}\right) = \frac{p^3 - 3r}{r}$ ....(9)

Also $\left(\frac{b+c}{a}\right)\left(\frac{c+a}{b}\right)\left(\frac{a+b}{c}\right) = -\frac{r}{r}$, using (6) and (7)

$= -1$ .... (10)

Using (8), (9) and (10), we see that the required equation is

$x^3 + 3x^2 + \left(\frac{p^3 - 3r}{r}\right)x + 1 = 0.$

7) $2 + i\sqrt{3}$ is one root. So another must be $2 - i\sqrt{3}$.

Thus, $(x^2 - 4x + 7) \mid (x^4 - 4x^2 + 8x + 35)$.

By long division, or by inspection, we see that

$x^4 - 4x^2 + 8x + 35 = (x^2 - 4x + 7)(x^2 + 4x + 5).$

Thus, the other two roots of the given equation are those of $x^2 + 4x + 5 = 0$, that is, $-2 \pm i$.

8) Now, $a^2 + b^2 + c^2 = (a + b + c)^2 - 2(ab + ac + bc)$

$\therefore ab + ac + bc = \frac{9-5}{2} = 2.$

Also, $a^3 + b^3 + c^3 = (a + b + c)^3 - 3[(a^2 + b^2)c + (b^2 + c^2)a + (c^2 + a^2)b] - 6abc$

$\therefore$ $abc = \frac{2}{3}$.

Thus a, b, c are the roots of

$x^3 - 3x^2 + 2x - \frac{2}{3} = 0$, that is,

$3x^3 - 9x^2 + 6x - 2 = 0$.

Thus a, b and c satisfy this equation, as well as,

$3x^4 - 9x^3 + 6x^2 - 2x = 0$.

$\therefore$ $3(a^4 + b^4 + c^4) - 9(a^3 + b^3 + c^3) + 6(a^2 + b^2 + c^2) - 2(a + b + c) = 0$

$\Rightarrow$ $a^4 + b^4 + c^4 = 25$.

9) Put $y = x - 4$, that is, $x = y + 4$ in the given equation. Then the equation that we get in y will be the required equation. Thus, the required equation is

$(y+4)^4 - 5(y+4)^3 + 7(y+4)^2 - 17(y+4) + 11 = 0$

$\Leftrightarrow$ $y^4 + 11y^3 + 43y^2 + 55y - 9 = 0$.

10) If $1 - i$ is a root, so must $1 + i$ be. Similarly, $3+2i$ and $3-2i$ are roots of the equation. Thus, the equation of lowest degree is the quartic

$[x - (1 - i)]\,[x - (1 + i)]\,[x - (3 + 2i)]\,[x - (3 - 2i)] = 0$, that is,
$x^4 - 8x^3 + 27x^2 - 38x + 26 = 0$.

This is unique, up to equivalence. That is, any other polynomial equation that satisfies our requirements must be equivalent to this equation.

11) $x^4 + 9x^3 + 16x^2 + 9x + 1 = 0$

$\Leftrightarrow$ $\left(x^2 + \frac{1}{x^2}\right) + 9\left(x + \frac{1}{x}\right) + 16 = 0$, by dividing throughout by x^2.

$\Leftrightarrow$ $\left(x + \frac{1}{x}\right)^2 + 9\left(x + \frac{1}{x}\right) + 14 = 0$.

Putting $x + \frac{1}{x} = t$, we get

$t^2 + 9t + 14 = 0$

Thus, $t = \frac{-9 \pm \sqrt{81 - 56}}{2} = \frac{-9 \pm 5}{2} = -2, -7$.

Thus, we get two equations in x, namely,

$x + \frac{1}{x} = -2$ and $x + \frac{1}{x} = -7$, that is,

$x^2 + 2x + 1 = 0$ and $x^2 + 7x + 1 = 0$.

On solving these quadratic equations, we get the four solutions of the original equation, which are $1, 1, \frac{-7 \pm \sqrt{45}}{2}$

APPENDIX: SOME MATHEMATICAL SYMBOLS AND TECHNIQUES OF PROOF

To be able to do any mathematical study, you need to know the language of mathematics. In this appendix we shall introduce you to some symbols and their meaning. We shall also briefly discuss some paths that you will often take to reach conclusions.

Symbols

1) **Implication** (denoted by $\Rightarrow$) : We say that a statement A implies a statement B if B follows from A.

   We write this as the compound statement, 'A $\Rightarrow$ B' or 'If A, then B'.

   For example, consider A and B, where

   A: Triangles ABC and DEF are congruent.

   B: Triangles ABC and DEF have the same area.

   Then A $\Rightarrow$ B ......(1)

   In this case 'A $\Rightarrow$ B' is a true statement.

   Another way of saying A $\Rightarrow$ B is that 'A only if B', or that 'A is sufficient for B'.

   The converse of the statement 'if A, then B' is the statement 'if B, then A', that is, B $\Rightarrow$ A (which is the same as A $\Leftarrow$ B).

   For example, the converse of (1) is

   'if two triangles have the same area, then they are congruent.'

   While studying geometry you must have proved that this statement is false. (For example, the right-angled triangles with sides 2, 3, $\sqrt{13}$ cm., and 1, 6, $\sqrt{37}$ cm. have the same area; but they are incongruent.) Thus, (1) is true, but its converse is not.

   Another way of saying A $\Leftarrow$ B is that 'A is necessary for B'.

2) **Two-way implication** (denoted by $\Leftrightarrow$): Sometimes we find two statements A and B for which A $\Rightarrow$ B and B $\Rightarrow$ A. In this situation we save space and write A $\Leftrightarrow$ B. This statement is the same as:

   'A is equivalent to B' ; or

   'A if and only if B', which we abbreviate to 'A iff B'; or
   'A is necessary and sufficient for B'.

   For example, let

   A : $x + 2 = 3$ and

   B : $x = 1$.

   Then A $\Rightarrow$ B and B $\Rightarrow$ A. Therefore, A $\Leftrightarrow$ B.

   Note that for the composite statement 'A iff B' to be true, both A $\Rightarrow$ B and B $\Rightarrow$ A should be true. Hence the statement

   'Two triangles are congruent iff they have the same area' is a false statement.

3) **For all/for every** (denoted by $\forall$) : Sometimes a statement involving a variable x, say P(x), is true for every value that x takes. We write this statement as:

   '$\forall$ x, P (x)' or 'P (x) $\forall$ x',

   meaning that P (x) is a true statement for every value of x.

   For example, let P (x) be the statement 'x > 0 '.

   Then P (x) $\forall x \in$ N.

4) **There exist/there exists** (denoted by $\exists$) : If a statement depending on a variable x, say P (x), is true for some value of x, then we write

   $\exists$ x such that P(x).

   This says that there is some x for which P(x) is true.

For example, $\exists x \in R$ such that $x - 3 = 2$.

Now, consider the two statements

$\forall x \in R, \exists y \in R$ such that $x = 2y$, and ......(3)

$\exists y \in R$, such that $\forall x \in R$, $x = 2y$. ......(4)

Is there a difference in them ? What does (3) mean? It means that for any real number x, we can find a real number y for which $x = 2y$.

In fact, $y = \frac{x}{2}$ serves the purpose.

Now look at (4). It says that there is some real number y such that whatever real number x we take, $x = 2y$ is true. This is clearly a false statement.

This shows that we have to be very careful when dealing with mathematical symbols.

So far we have looked at the meaning and use of some common logical symbols. Let us now consider some common techniques of proof.

Methods of proof

In any mathematical theory, we assume certain facts called axioms. Using these axioms, we arrive at certain results (theorems) by a sequence of logical deductions. Each such sequence forms a proof of a theorem. We can give proofs in several ways.

1) **Direct proof:** A direct proof, or step in a proof, takes the following form:

   A is true and the statement '$A \Rightarrow B$' is true, therefore B is true.

   For example,

   Δ ABC is equilateral and (If a triangle is equilateral, then it is an isosceles triangle.), therefore, Δ ABC is an isosceles triangle.

   One kind of result that you will often meet in this course and other mathematics courses is a theorem that asserts the equivalence of a number of statements, says A, B, C. We can prove this by proving $A \Rightarrow B$, $B \Rightarrow A$, $A \Rightarrow C$, $C \Rightarrow A$ and $B \Rightarrow C$, $C \Rightarrow B$. But, if $A \Rightarrow B$ and $B \Rightarrow C$ are both true, then so is $A \Rightarrow C$. So, a shorter proof could consist of the steps $A \Rightarrow B$, $B \Rightarrow C$, $C \Rightarrow A$. We write this in short as $A \Rightarrow B \Rightarrow C \Rightarrow A$.
   Or, the proof could be $A \Rightarrow C \Rightarrow B \Rightarrow A$. Thus the order doesn't matter, as long as the path brings us back to the starting point, and all the statements are covered.

   Whenever you meet such a result in our courses, we shall indicate the path we shall follow.

2) **Contrapositive proof:** This is an indirect method of proof. It uses the fact that '$A \Rightarrow B$' is equivalent to its contrapositive, namely, 'not $B \Rightarrow$ not A', that is, if B does not hold, then A does not hold.

   (For example, $x = -2 \Rightarrow x^2 = 4$ is equivalent to its contrapositive,
   $x^2 \neq 4 \Rightarrow x \neq -2$.)

   Sometimes, it is easier to prove the contrapositive of a given result. In such situations we use this method of proof. So, how does this method work? To prove '$A \Rightarrow B$', we prove 'not $B \Rightarrow$ not A', that is, we assume that B does not hold, and then, through a sequence of logical steps, we conclude that A does not hold.

   Let us look at an example. Suppose we want to prove that 'If two triangles are not similar, then they are not congruent'. We prove its contrapositive, namely, 'If two triangles are congruent, then they are similar', which is easy to prove.

3) Proof by contradiction: This method is also called *reductio ad absurdum*, a Latin phrase. In this method, to prove that a statement A is true, we start by assuming that A is false. Then by logical steps we arrive at a known false statement. So we reach a contradiction. Thus, we are forced to conclude that A cannot be false. Hence, A is true.

   For example, to prove that $\sqrt{2} \notin Q$, we start by assuming that $\sqrt{2} \in Q$.

   Then $\sqrt{2} = \frac{p}{q}$ for some $p, q \in Z$, $q \neq 0$, and $(p, q) = 1$.

$\Rightarrow 2 = \frac{p^2}{q^2} \Rightarrow 2q^2 = p^2 \Rightarrow 2 \mid p^2 \Rightarrow 2 \mid p.$

Let $p = 2m$.

Then $2q^2 = p^2 = 4m^2$.

$\Rightarrow q^2 = 2m^2 \Rightarrow 2 \mid q^2 \Rightarrow 2 \mid q$, which is not possible because we assumed that $(p, q) = 1$.

Thus, we arrive at a contradiction. Hence, we conclude that $\sqrt{2} \notin Q$.

4) **Proof by counter-example:** Consider a statement P(x) depending on a variable x. Suppose we want to disprove it, that is, prove that it is false. One way is to produce an x for which P(x) is false. Such an x is called a **counter-example** to P(x).

For example, let P(x) be the statement

'Every natural number is a product of distinct primes'.

Then $x = 4$ is a counter-example, since $4 \in N$ and $4 = 2 \times 2$ is not a product of distinct primes. In fact, we have several counter-examples in this case.

This is not always the best method for disproving a statement. For example, suppose you want to check the truth of the statement

'Given $a, b, c \in Z$, $\exists\ n \in Z$ such that $an^2 + bn + c$ is not a prime number'.

If you try to look for counter-examples, then you're in trouble because you will have to find infinitely many— one for each triple $(a, b, c) \in R^3$. So, why not try a direct proof, as the one below.

Proof : For fixed $a, b, c \in Z$, take any $n \in Z$ and let $an^2 + bn + c = t$.

Then $a(n+t)^2 + b(n+t) + c = t(\dot{Z} + 2an + b)$, which is a proper multiple of t. Thus our statement is true.

NOTES

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

UGMM-04
Elementary Algebra

Block

2

EQUATIONS AND INEQUALITIES

BLOCK 2 EQUATIONS AND INEQUALITIES

Suppose you want to plan a meal for a growing child with three food types. Each food type has three vitamins in varying quantities. You will need to calculate the amount of each food type which will give the child the minimum required amount of the different vitamins per day. How will you do this ? Algebra gives you an easy way, which you will read about in Units 4 and 5. In these units we discuss three methods of finding common solutions of several linear equations. These methods aren't new; in fact, one of them has been used since ancient times. The other two methods were developed in the eighteenth century. All these methods are still the standard methods for solving systems of linear equations.

In Unit 6, the last unit of this course, we look at some algebraic inequalities. We will first discuss some inequalities which were even known to the ancient Greeks. Then we will study three inequalities due to some famous nineteenth century European mathematicians. All these inequalities are useful in mathematics and the other sciences, and they are very simple to prove and apply. That is why we have chosen to expose you to them in this elementary course on algebra.

Now, a few suggestions that may help you study the units in this block. Before starting the study of this block, we suggest that you glance through the notation given at the beginning of Block 1. Also, do try the exercises in the units in this block as and when you come to them. This will help you to confirm your understanding of the related study material. At the end of the block we have given a list of miscellaneous exercises for you to do, if you want to. Doing them will be useful to you. After finishing this block please try the assignment which is based on this course.

Equations and Inequalities

Equations and Inequalities

6

Equations and Inequalities

8

Equations and Inequalities

10

Equations and Inequalities

12

Equations and Inequalities

14

Equations and Inequalities

16

Equations and Inequalities

18

Equations and Inequalities

20

Equations and Inequalities

22

Equations and Inequalities

24

Equations and Inequalities

26

Equations and Inequalities

28

Equations and Inequalities

30

Equations and Inequalities

32

Equations and Inequalities

34

Equations and Inequalities

36

Equations and Inequalities

38

Equations and Inequalities

40

Equations and Inequalities

42

Equations and Inequalities

44

Equations and Inequalities

46

Equations and Inequalities

48

Equations and Inequalities

50

Equations and Inequalities

52

Equations and Inequalities

54

Equations and Inequalities

56

Equations and Inequalities

58

Equations and Inequalities

60

UNIT 4 SYSTEMS OF LINEAR EQUATIONS

Structure

4.1 INTRODUCTION

In the last unit we introduced you to polynomial equations in one variable. In this unit we will start by considering linear equations in one or more variables. After that we shall consider ways of obtaining common solutions for several such equations. We call a set of linear equations a system of linear equations. Such systems of equations can arise while studying many practical problems. These include studying oscillations, the flow of currents, migration patterns, chemical contents of various solutions, input-output models of industrial production, and so on. Therefore, it is important that you spend some time studying them.

The first definite trace of systems of linear equations is found in Chui-chang Suan-shu, that is, Nine Chapters on the Mathematical Art. This is an ancient Chinese mathematical text which was probably written in 1100 B.C. Much later, in the third century B.C., the Greeks used some methods for solving certain systems of equations. Further notable developments in this area of mathematics took place in the 17th century. The Japanese mathematician Seki Kowa (around 1683) contributed greatly to the theory of systems of linear equations. About the same time the European mathematician Leibniz also discovered a method for solving systems of linear equations. In the next century the mathematicians Gauss and Cramer published methods that use the concepts of matrices and determinants for solving simultaneous equations.

In this unit we will discuss two methods for solving systems of linear equations. We will do the method due to Cramer in the next unit.

Let us now list the objectives of this unit.

Objectives

After studying this unit, you should be able to

- obtain the solution set of a linear equation in one or more variables;
- define a system of m linear equations in n unknowns;
- apply the methods of substitution and elimination for solving simultaneous linear equations;
- choose the appropriate method, of the two methods discussed, for solving a given linear system.

Let us now start our discussion on linear equations.

4.2 LINEAR SYSTEMS

You know that the most general form of a **linear equation over R in one variable** x is $ax + b = 0$, a, b ∈ **R**, a ≠ 0. You also know that this has a unique solution, namely, $x = -\frac{b}{a}$.

Now, can you think of a linear equation in two variables ? What about $2x + 5y + 5 = 0$? According to the following definition, it is linear in two variables.

Definition: A linear equation in two variables x and y is an equation which can be written as

$$ax + by + c = 0,$$

where a,b,c ∈ **R** and a and b are not both zero.

For example,

$-x + \frac{1}{2}y = 0$, $x = 25$ and $2s - 4t = 2$ are linear equations in two variables.

What about $xy = 1$, the equation of a hyperbola ? Is it a linear equation in 2 variables ? It is not, since x and y are both variables; and hence, it is not of the form $ax + by + c = 0$, where a, b, c ∈ **R**.

Try the following exercises now.

E1) Which of the following equations are linear in 2 variables ? Can you explain why ?

a) $2x + 3xy - 4y = 10$

b) $x + y^2 = 6$

c) $\sqrt{u} + v = 2$, where u and v are variables.

d) $2x = \frac{5x - 2y}{4} + 1$

E2) "Every linear equation in one variable is also a linear equation in two variables." Is this statement true ? Why?

Now, what would any solution of the linear equation $2x+3y+1= 0$ look like ? It would consist of an ordered pair of real numbers say (a,b), such that $2a+3b+1= 0$. For example, (1, −1) is a solution, since $2(1) + 3(-1) + 1 = 0$.

You can check that $\left(\frac{1}{2}, -\frac{2}{3}\right)$ and $\left(-\frac{1}{2}, 0\right)$ are also solutions.

In fact, the given equation has infinitely many solutions given by $\left(x, \frac{-(2x+1)}{3}\right)$, as x varies in **R**. How do we get this general form of the solution ? We can rewrite the equation as $y = \frac{-(2x+1)}{3}$. Then, for any value that we give, x, say x=a , we get a corresponding value $\frac{-(2a+1)}{3}$ for y. Thus, $\left(a, \frac{-(2a+1)}{3}\right)$ is a solution ∀ a ∈ **R**. Note that the solution set is a subset of $\mathbf{R}^2$.

Now, we could also have rewritten $2x + 3y + 1 = 0$ as $x = \frac{-(3y+1)}{2}$. Then the solution set would have been $\left\{\left(\frac{-(3y+1)}{2}, y\right) \mid y \in \mathbf{R}\right\}$. Are the two solution sets different ? Not at all. if we write $-\left(\frac{3y+1}{2}\right) = x$, then $-\left(\frac{2x+1}{3}\right) = y$. Thus, the sets are the same.

This shows us that we can either obtain the solution set in terms of x or in terms of y.

Now, consider the equation $x - 2 = 0$ as a linear equation in two variables. What is its solution set? Whatever value y takes, x will always have to be 2. Thus, the solution set is $\{(2, y) \mid y \in \mathbf{R}\}$. It is the set of all points on the line $x - 2 = 0$ (Fig.1). In fact,

> any linear equation in 2 variables can be geometrically represented by a straight line in the xy-plane.

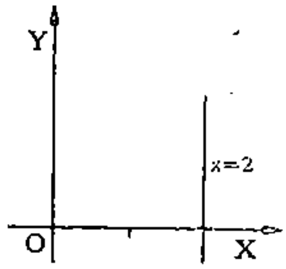


Fig. 1 : x = 2

Now let us define a linear equation in n variables, where $n \in \mathbf{N}$.

Definition: A linear equation over R in n variables $x_1, x_2, ..., x_n$ has the general form

$$a_1x_1 + a_2x_2 + ... + a_nx_n + b = 0,$$

where $a_1, a_2, ..., a_r, b \in \mathbf{R}$ and not all of $a_1, a_2, ...a_n$ equal zero.

Thus, $2x + 3y = 11z$ is a linear equation in 3 variables x,y and z. What does a solution of this look like ? It will be an ordered triple of real numbers that satisfies the equation. For example, (0, 0, 0) and (22, 0, 4) are solutions. But, (1, 1, 1) is not a solution.

Let us see what a solution of a general linear equation looks like.

Definition : An n-tuple $(b_1, b_2, ..., b_n)$ in $\mathbf{R}^n$ is called a **solution** of the linear equation $a_1x_1 + a_2x_2 + ... + a_nx_n = b$, if

$$a_1b_1 + a_2b_2 + ... + a_nb_n = b.$$

In this case we also say that $x_1 = b_1, x_2 = b_2, ..., x_n = b_n$ satisfy the linear equation.

Note that the first element of the n-tuple is substituted for the first variable, the second for the second variable, and so on.

Remember that **a linear equation in two or more variables has infinitely many solutions.**

Now why don't you see if you have absorbed what we have done so far.

E3) Which of the following are solutions of $3x - 2y + 5z = 80$?

a) (0, – 40), (b) (0, – 40, 0), (c) (2, 3, 15),

d) $(1, 1, \frac{79}{5})$.

E4) Find the solution set of $x = y$. Also give its geometrical representation.

Studying only one linear equation at a time has been found inadequate for interpreting and solving real-world problems mathematically. The mathematical models of many problems consist of a set of several linear equations which need to be solved at the same time. For example, suppose the Indian Government has to suddenly send supplies of blood, medical kits, food and water to a quake-hit area. It knows the volume and weight of each unit of these items. It also knows that each aeroplane can take a maximum capacity of 600 cubic metres and a maximum weight of 20,000Kg. These facts, put together, lead to the two equations

$$2x_1 + 3x_2 + 0.8x_3 + 0.6x_4 = 600$$

$$75x_1 + 50x_2 + 30x_3 + 35x_4 = 20{,}000,$$

where x_1, x_2, x_3, x_4 denote the number of containers of blood, medical kits, food and water, respectively. We need to find common solutions to both these equations so as to get the amounts that can be sent. In other words, we need to solve these equations simultaneously. That is why we call such a set of equations **simultaneous linear equations.**

Definition: Any finite set of linear equations is called a **system of linear equations,** or a **linear system,** or **simultaneous linear equations.**

You have just seen one example involving emergency airlifting. For another example, consider the three equations

$$\left.\begin{aligned} 2x+\frac{7}{2}y+3z &= 1200 \\ 3x+\frac{5}{2}y+2z &= 1150 \\ 4x+3y+2z &= 1400 \end{aligned}\right] \qquad \text{.........................(1)}$$

They form a linear system. This system is the mathematical formulation of the following problem:

A company produces 3 products, each of which must be processed through 3 divisions. A,B, and C. The number of hours taken by each unit of the product in each division, and the total number of hours available for production each week is given in Table 1.

Table 1

Division	Product			Total number of hours per week
	1	2	3	
A	2	3.5	3	1200
B	3	2.5	2	1150
C	4	3	2	1400

What is the number of units of each product that should be produced so as to exhaust the weekly capacities of the 3 divisions?

How is the system (1) obtained from this problem? Well, if x,y and z denote the number of units of each product, we get the system (1).

In the following exercises you can see some more examples of linear systems arising from practical problems.

E5) A dietitian is planning a noon meal for school children. It consists of 3 food types. He wants to ensure that the minimum daily requirements (MDR) for 4 vitamins are satisfied.
In Table 2 we summarise the vitamin content per unit of each food type in milligrams, and we give the MDR.

Table 2

Food Type	Vitamin content/unit (in mg.)			
	V_1	V_2	V_3	V_4
1	3	1	0	1
2	5	7	2	6
3	2	3	0	2
MDR	55	45	10	45

What is the mathematical formulation of this problem?

E6) Thirty litres of a 50% alcohol solution are to be made by mixing 70% solution and 20% solution. We want to know how many litres of each solution should be used. Translate the problem into a linear system.

Let us now discuss what the set of solutions of a system of linear equations looks like.

Consider the following linear system in one variable:

$$ax + b = 0$$

$$cx + d = 0,$$

where $a, b, c, d \in \mathbf{R}$, $a \neq 0$, $c \neq 0$.

This will have a solution if and only if the two equations have a common solution, that is , iff $-\frac{b}{a}=-\frac{d}{c}$. And then, $x=-\frac{b}{a}$ (or $-\frac{d}{c}$) is the unique solution.

For example, the system

$$x + 1 = 0$$
$$3x + 3 = 0$$

has the unique solution $x = -1$, while the system

$$3x = 0$$
$$2x + 5 = 0$$

has no solution.

Now consider the system

$$\left.\begin{aligned} x + 2y &= 5 \\ x + y &= 3 \end{aligned}\right] \qquad \text{......................(2)}$$

From the second equation we get $y = 3 - x$. Substituting this value in the first equation we get

$x + 2(3 - x) = 5$, that is, $x = 1$.

Then $y = 3 - 1 = 2$.

So, (2) has a solution, namely, $x = 1$ and $y = 2$, that is, the ordered pair (1, 2).

Now, recall that the solutions of a linear equation in two variables correspond to the points on the line representing the equation. Thus, the solutions of (2) would correspond to the points of intersection of the two lines representing the two equations. From Fig. 2(a) you can see that they intersect in only one point, namely, (1,2). Thus, (2) has a unique solution.

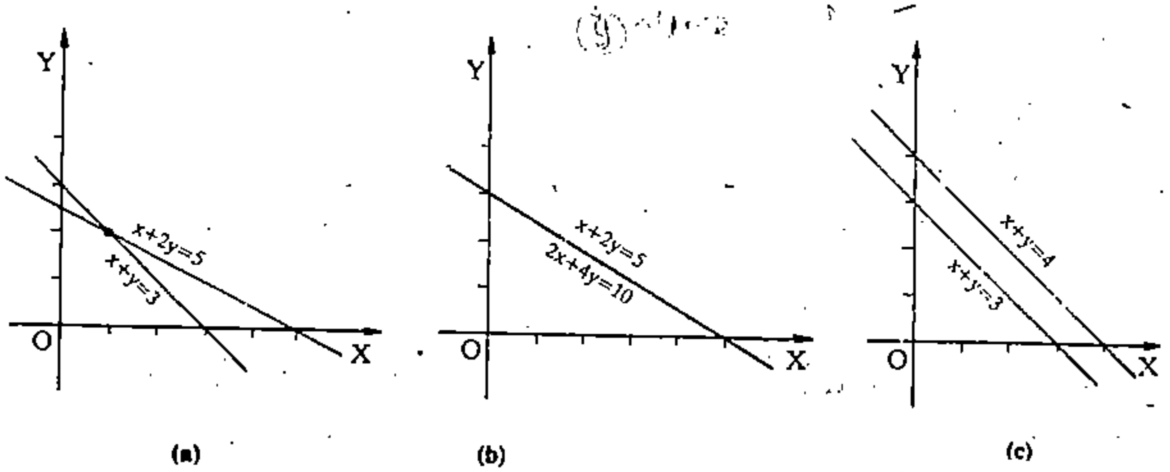


Fig. 2 : A linear system with (a) a unique solution; (b) infinitely many solutions; (c) no solution.

Now consider the system

$$\left.\begin{aligned} x + 2y &= 5 \\ 2x + 4y &= 10 \end{aligned}\right] \qquad \text{..............................(3)}$$

You can check that for any $y \in \mathbf{R}$, the ordered pair $(5 - 2y, y)$ is a solution of (3). Thus, this system has infinitely many solutions.

Geometrically, since both the equations of (3) are multiples of each other, they represent the same line in the plane (see Fig.2(b)). Thus, every point on the line is a common point. Hence the system (3) has infinitely many common points.

Finally, consider the system

$$\left.\begin{array}{l} y = 3 \\ x + y = 4 \end{array}\right] \quad \text{..............................(4)}$$

You can see that this system of equations has no solution, since any solution would lead to the false statement 3 = 4.

Geometrically, the two equations of (4) represent distinct parallel lines (see Fig. 2(c)). Thus, they have no point of intersection.

So you have seen three situations, namely,

> i) a linear system can have a unique solution, or
>
> ii) a linear system can have infinitely many solutions, or
>
> iii) a linear system can have no solution.

In fact, these are the only situations possible for any system of linear equations. We shall not prove this statement here.

Now let us go back to a general linear system. We give the following definition.

Definition: If a system of linear equations has a solution, we call it **consistent;** otherwise we call it **inconsistent.**

Thus, (2) and (3) are consistent systems, while (4) is not.

Why don't you try the following exercise now?

E7) Give the geometrical view of the following system of equations. Hence find out which of them are consistent.

a) $x + y = 3$

$x = 0$

$y = 0$

b) $x + y = 2$

$2x + 2y = 10$

$x = y$

c) $3x + y = 0$

$3x - y = 0$

$x - y = 0$

d) $x = 3$

$y = 4$

Now let us discuss a method of solving a system of linear equations.

4.3 SOLVING BY SUBSTITUTION

Let us consider the following system of linear equations in one variable:

$$\left.\begin{array}{l} 3x + 5 = 0 \\ 6x + 10 = 0 \end{array}\right] \quad \text{..............................(5)}$$

equation, we get

$-5 + 5 = 0$, a true statement.

Thus, the equations in (5) are consistent, and the unique solution is $x = -\frac{5}{3}$. The method we have just used for solving (5) is called the **substitution method.**

Let us see how this method can be used for solving linear systems in two variables. Consider the system

$$\left.\begin{array}{l} 2x + y = 7 \\ 5x + 3y = 18 \end{array}\right] \quad \text{.................................(6)}$$

We want to solve the equations in (6) simultaneously, that is, at the same time, by **substitution.** For this we first write one variable in terms of the other by using either of the equations. We will use the first one to write y in terms of x, as $y = 7 - 2x$.

Then we substitute this value of y in the second equation, to get $5x+3(7-2x) = 18$, that is, $21-x = 18$.

This gives $x = 3$.

Substituting this value of x in $y = 7-2x$, we get $y = 1$.

But, is (3,1) a solution? We must double check by substituting these values in (6). We get $2 \times 3 + 1 \times 1 = 7$, which is true, and $(5 \times 3) + (3 \times 1) = 18$, which is true. Thus, the system (6) has the unique solution (3,1).

We can also solve (6) by using the second equation to write $x = \frac{18-3y}{5}$. Then substituting in the first equation, we get $2\left(\frac{18-3y}{5}\right) + y = 7$, giving $y = 1$.

And then $x = \frac{18-3y}{5} = \frac{18-(3 \times 1)}{5} = 3$.

To get some practice in solving by substitution, try the following problems.

E8) Find solutions (if any) of the following sets of simultaneous equations by the substitution method.

a) $x + y = -2$

$y = 3$

b) $3a + 7b = 33$

$a + 3b = 13$

c) $2s + t = 20$

$2s - 5t = 30$

d) $x + y = 2$

$2x + 2y = 4$

e) $3x = y+5$

$9 + y = 3x$

The substitution method that we have employed for two equations in two unknowns can also be extended for solving several equations in several unknowns. But it becomes more and more difficult to apply as the number of equations and variables increases. In the next section we will discuss a better method of dealing with any number of equations in any number of variables.

4.4 SOLVING BY ELIMINATION

Fig. 3 : Gauss in 1803

This method of solving simultaneous linear equations is due to the great German mathematician Carl Friedrich Gauss (1777 — 1855). Because of his immense contribution to the development of mathematics, he is known as the 'prince of mathematicians'. The method of solution is called the **Gaussian elimination** (or **successive elimination**) method. In this method we use multiplication and addition to **eliminate the variables**, one by one, from the equations. At each stage we transform the system of equations into an equivalent one.

Two systems of equations are **equivalent** if they have the same solution set.

Any of the following transformations are allowed:

1) changing the order of the equations of the system;
2) multiplying both sides of any equation of the system by a non-zero real number;
3) replacing an equation by the sum of that equation and a non-zero multiple of another equation in the system.

Let's work out a simple example, using this method. Consider the system

$x + 2y + z = 4$ ....(7)
$3x - y - 4z = -9$ ....(8)
$x + y + z = 2$ ....(9)

Let us begin by eliminating y from (8) and (9), by adding them. We get

$4x - 3z = -7$ ....(10)

Now let us eliminate y from (7) and (8). For this we add (7) to 2 times (8). We get

$(x + 2y + z) + 2(3x - y - 4z) = 4 + 2(-9)$, that is,

$7x - 7z = -14$

Dividing throughout by 7, we get

$x - z = -2$ ....(11)

Now, we can eliminate x from (10) and (11) by adding (−4) times (11) to (10). We get

$(4x - 3z) - 4(x - z) = -7 - 4(-2)$, that is,
$z = 1$.

Substituting this value of z in (11) we get

$x = -2 + 1 = -1$.

Substituting $x = -1$, $z = 1$ in (9), we get

$y = 2$.

We must **verify** if the ordered triple (−1, 2, 1) satisfies all three equations.

On substituting this triple in each of the equations, we find that it is indeed the solution.

Whenever we use this method, or any method for solving a linear system we must keep the following remarks in mind.

Remark 1 : Whenever we solve an equation or a system of equations, we must always verify our solution.

Remark 2 : While solving a linear system, if we reach a false statement, it means that the system has no solution.

Now why don't you try to solve some linear systems.

E9) Solve the following systems by the Gaussian elimination method.

a) $2x + y + z = 9$

$-x-y+z=1$

$3x-y+z=9$

b) $-3x+4y+5z=6$

$6x+7y \quad = 8$

$2x-3y+z=1$

E10) Solve the system that you got in E5, by elimination.

E11) Determine, by elimination, the solution set of the system

$$\begin{aligned} -2x+y+3z &= 12 \\ x+2y+5z &= 10 \\ 6x-3y+9z &= 24 \\ 5x+5y+22 &= 0 \end{aligned}$$

In E10 and E11 you came across systems in which the number of equations was more than the number of variables. In such a situation also the system can have a unique solution, infinitely many solutions or no solution.

There can also be systems of equations with more variables than equations. Such a system will not have a unique solution. Thus, it will either be inconsistent, or it will have infinitely many solutions.

Let us consider the following example.

$4x-y+z=0$ ....(12)

$x+y+z=5$ ....(13)

We first eliminate x.

$(12)-4\times(13) \Rightarrow -5y-3z=-20$

$\Rightarrow 5y+3z=20$ ....(14)

We can't eliminate any more variables because playing around with (14) and the original system will only end in reintroducing x. Instead, we use (14) to write y in terms of z. We get

$y=\frac{20-3z}{5}$.

Then (13) $\Rightarrow x=5-\left(\frac{20-3z}{5}\right)-z=\frac{5-2z}{5}$.

Substitute the triple $\left(\frac{5-2z}{5}, \frac{20-3z}{5}, z\right)$ where $z \in \mathbf{R}$, in (12) and (13), to verify that it is a solution. What do you find? For any $z \in \mathbf{R}$, the triple is a solution of the given system.

For example, when $z = 0$ we get a solution (1, 4, 0); and when $z = 1$ we get a solution $\left(\frac{3}{5}, \frac{17}{5}, 1\right)$ and so on. Thus, the given linear system has infinitely many solutions.

We say that the solutions are $\left(\frac{5-2z}{5}, \frac{20-3z}{5}, z\right)$ where z is an **arbitrary** real number, or a **parameter**.

Now consider the system

$x+2y+z=1$ ....(15)

$x+2y+z=-1$ ....(16)

$(15)-(16) \Rightarrow 0=2$, a false statement.

Thus, the system is inconsistent.

Now why don't you practice the elimination method some more?

E12) Solve the system you got in E6, if it is consistent.

E13) Solve the system (1) that we gave at the beginning of Sec. 4.2.

E14) Solve the system

$$\begin{aligned} x + y + z &= 20 \\ 10x + y - 2z &= 5 \end{aligned}$$

Note the relationship between the systems in E15 (a) and E15 (b).

E15) Solve the systems

a)
$$\begin{aligned} x + y + z &= 0 \\ y + 2z &= 3 \end{aligned}$$

b)
$$\begin{aligned} x + y + z &= 0 \\ y + 2z &= 3 \\ z &= 4 \end{aligned}$$

So far we have discussed two methods of solving linear systems. In the next unit we will consider yet another method, which is specifically meant for a system of linear equations in which the number of equations is the same as the number of unknowns.

Let us now summarise what we have covered in this unit.

4.5 SUMMARY

In this unit we have discussed systems of linear equations. In particular you studied

1) what a linear system is and how it can arise from practical problems.

2) that a linear system can have a unique solution, infinitely many solutions or no solution.

3) the substitution method for solving "small" linear systems simultaneously.

4) the Gaussian elimination method, which is the method that is the most widely used.

We hope that you have tried all the exercises in the unit. You may like to see what our solutions to them are.

4.6 SOLUTIONS/ANSWERS

E1) (a) is not, since the quadratic term xy occurs in it.

(b) is not, since the quadratic term y^2 occurs in it.

(c) is not; in fact, it is not even a polynomial equation.

(d) is linear, since it is equivalent to the linear equation $3x + 2y - 4 = 0$.

E2) It is true because any linear equation in one variable is $ax + b = 0$, $a \neq 0$. This is equivalent to $ax + 0.\ y + b = 0$, $a \neq 0$, which is linear in two variables.

E3) (b) and (d) are. (a) $\in \mathbf{R}^2$, and hence can't be a solution. (c) is not, since $3(2) - 2(3) + 5(15) \neq 80$.

E4) $\{(x, x) \mid x \in \mathbf{R}\}$.

Its geometrical representation is given in Fig. 4.

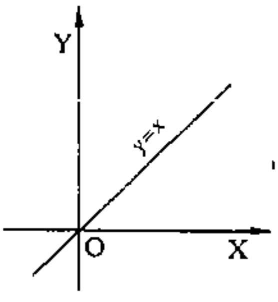


Fig. 4 : y = x

E5) Let, x, y, z denote the units of each food type. Then

$$\begin{aligned} 3x + 5y + 2z &= 55 \\ x + 7y + 3z &= 45 \\ 2y &= 10 \\ x + 6y + 2z &= 45 \end{aligned}$$

E6) Say, we take x litres of the 70% solution and y litres of the 20% solution to make 30 litres of the 50% solution.

Then $\frac{70}{100}x + \frac{20}{100}y = \frac{50}{100} \times 30$

Also x + y = 30.

Thus, the problem reduces to solving the linear system

$7x + 2y = 150$

$x + y = 30.$

E7) a) From Fig.5 you can see that there is no point common to all three lines. Hence the system is inconsistent.

b) We have given the geometrical representation of this system in Fig.6. Again, you can see that the system is inconsistent.

c) In Fig.7 you can see that the three lines have a unique point of intersection, namely, (0, 0). Thus, the system has the unique solution (0, 0).

d) From Fig.8 you can see that this system has the unique solution (3,4).

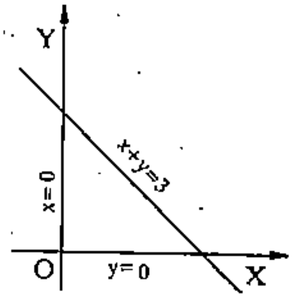


Fig. 5 : An inconsistent system

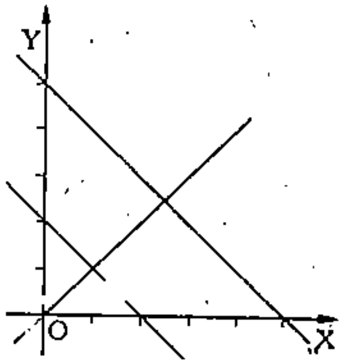


Fig. 6 : An inconsistent system

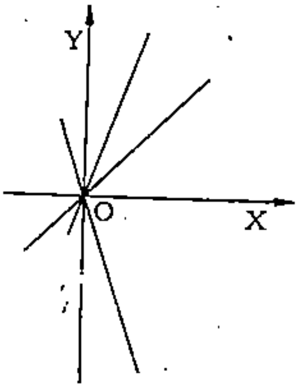


Fig. 7

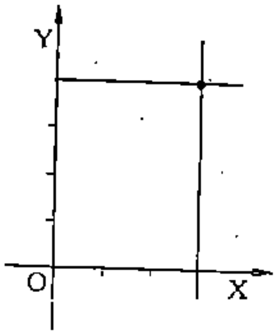


Fig. 8

E8) a) The second equation says y = 3. Substituting this value in the first equation, we get $x + 3 = -2 \Rightarrow x = -5$.

$\therefore$ (−5, 3) is the solution.

b) $a + 3b = 13 \Rightarrow a = 13 - 3b$.

$\therefore 3a + 7b = 33 \Rightarrow 3(13 - 3b) + 7b = 33 \Rightarrow 2b = 6 \Rightarrow b = 3.$

$\therefore a = 13 - 3\ (3) = 4.$

$\therefore$ (4, 3) is the solution.

c) $2s + t = 20 \Rightarrow t = 20 - 2s.$

$\therefore 2s - 5t = 30 \Rightarrow 2s - 5(20 - 2s) = 30 \Rightarrow s = \frac{65}{6}$

$\therefore t = 20 - \frac{65}{3} = -\frac{5}{3}.$

$\therefore \left(\frac{65}{6}, -\frac{5}{3}\right)$ is the solution.

d) $x + y = 2 \Rightarrow y = 2 - x$

$\therefore 2x + 2y = 4 \Rightarrow 2x + 2(2 - x) = 4 \Rightarrow 0 = 0.$

Note that the second equation is equivalent to the first one. Thus, any solution of the system is a solution of x + y = 2.
Thus, for any value of $x \in \mathbb{R}$, (x, 2 − x) is a solution. For example, (0, 2) is a solution.
This system has infinitely many solutions.

e) $3x = y + 5 \Rightarrow y = 3x - 5$
$\therefore 9 + y = 3x \Rightarrow 9 + 3x - 5 = 3x \Rightarrow 4 = 0$, a false statement.
Thus the system is inconsistent.

E9) a) $2x + y + z = 9$ ....(17)
$-x - y + z = 1$ ....(18)
$3x - y + z = 9$ ....(19)

To eliminate y we add (17) and (18). We get
$x + 2z = 10.$ ....(20)
To eliminate z we subtract (18) from (19). We get
$4x = 8$, that is, $x = 2$.
Substituting this value of x in (20), we get

$2z = 10 - 2 = 8 \Rightarrow z = 4.$
Then (17) gives us
$2(2) + y + 4 = 9 \Rightarrow y = 1.$
Thus, (2, 1, 4) is the solution. (Verify this !)

b) $\left(\frac{29}{41}, \frac{22}{41}, \frac{49}{41}\right)$

E10) $3x + 5y + 2z = 55$ ....(21)

$x + 7y + 3z = 45$ ....(22)

$2y = 10$ ....(23)

$x + 6y + 2z = 45$ ....(24)

$(23) \Rightarrow y = 5.$

Then $(21) \Rightarrow 3x + 2z = 30$ ....(25)

and $(22) \Rightarrow x + 3z = 10$ ....(26)

Eliminating x from (25) and (26), we get

$z = 0.$

Then $(24) \Rightarrow x + 6(5) + 2(0) = 45 \Rightarrow x = 15.$

Now we need to check if (15, 5, 0) satisfies all the equations in the system. It doesn't satisfy (21). But our calculations have been right.

Conclusion: the system is inconsistent !

The dietitian will have to alter his constraints !

E11) $\left(-\frac{32}{15}, \frac{34}{15}, \frac{10}{3}\right)$

E12) The system is

$7x + 2y = 150$

$x + y = 30.$

It has a unique solution, namely, (18, 12).

E13) (200, [illegible]0, 150).

E14) The solution set is $\{(x, 15 - 4x, 3x + 5) \mid x \in \mathbf{R}\}$.

E15) a) $\{(z - 3, 3 - 2z, z) \mid z \in \mathbf{R}\}$ is the solution set.

b) (1, – 5, 4) is the unique solution.

UNIT 5 CRAMER'S RULE

Structure

5.1 INTRODUCTION

In the previous unit we introduced you to linear systems and two methods for solving them. In this unit we shall discuss a method for solving a particular type of linear system. We shall first briefly introduce you to an efficient notation for dealing with systems of linear equations, namely, a matrix. You can find a detailed study in our course on linear algebra.

After that we shall explain a concept that is intimately linked with a certain type of matrix, and hence with the solution of certain systems of linear equations. The concept is that of a determinant, which seems to have been first used by ancient Chinese mathematicians for solving simultaneous linear equations. In 1683 the Japanese mathematician Seki Kowa started developing the theory of determinants for the same purpose. About the same time the German mathematician Leibniz also defined determinants and developed their use for solving simultaneous linear equations. So, you can see that mathematicians through the ages and around the globe have felt that determinants are very important. Nowadays scientists and social scientists also increasingly feel the need to understand and use this concept.

We shall end this unit with a discussion on a method which uses determinants for solving certain systems of linear equations. This method is due to the eighteenth century mathematician Cramer. It only applies to some of those linear systems in which the number of variables equals the number of equations.

Let us now list the objectives of this unit.

Objectives

After studying this unit, you should be able to

- define a matrix, and a square matrix, in particular;
- evaluate any determinant of order 1, 2 or 3;
- identify a non-singular matrix;
- identify the linear systems which can be solved by using Cramer's rule, and apply the rule to solve them.

And now let us look at a convenient way of representing a linear system.

5.2 WHAT IS A MATRIX?

Consider the set of linear equations

$$\left.\begin{aligned} 2x + y - z &= 5 \\ x + 5y - 3z &= -6 \\ -x + 2y + 2z &= 1 \end{aligned}\right] \qquad \text{.................................(1)}$$

While writing (1) we have had to write each variable three times. Wouldn't it be more satisfactory if there were a notation that would enable us to avoid this repetition? After all, it is the coefficients that really matter in obtaining their solutions. Let us do away with writing x, y, z each time, and only write their coefficients in a table in the following manner :

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 5 & -3 \\ -1 & 2 & 2 \end{bmatrix}$$

How did we prepare this table? The first row consists of the coefficients of x, y and z, respectively, in the first equation; the second row consists of the coefficients in the second equation; and the third row consists of the coefficients in the third equation.

In fact, we can symbolically rewrite (1) as

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 5 & -3 \\ -1 & 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ 1 \end{bmatrix} \qquad \text{..............................(2)}$$

Each group of numbers or variables enclosed in the square brackets is an example of a matrix.

Writing systems of equations in matrix notation leads to a saving of effort in dealing with them, especially as the number of equations grows. Nowadays computers are being used increasingly for solving large systems of linear equations. Their efficiency increases considerably if matrix methods are used. You may be interested to know that matrices (plural of "matrix") were first used in 250 B.C. for solving systems of linear equations in the Chinese mathematics text "Nine Chapters on the Mathematical Art". But the development of matrix theory is mainly due to the 19th century British mathematicians Arthur Cayley and J.J. Sylvester. We will discuss matrix theory in great detail in our course on linear algebra. In this section we shall only acquaint you with matrices. Let us start with the definition.

Definition : A **matrix** is a rectangular arrangement of numbers in the form of horizontal and vertical lines.

The numbers occurring in a matrix are called its **elements** or **entries**.

The set of entries in one horizontal line of a matrix is called a **row** of the matrix; and the set of entries in a vertical line is called a **column** of the matrix.

For example,

$\begin{bmatrix} 1 & 3 & -4 \\ 2 & 1/2 & 5 \end{bmatrix}$ is a matrix with two rows and three columns, and $[-1]$ is a matrix with one row and one column.

Note that each row of a matrix has the same number of elements. Similarly, each column of a matrix has the same number of elements. This is why we say that the arrangement of numbers in a matrix is 'rectangular'.

Now a few words about notation.

As you have seen in the examples of matrices given so far, we use **square brackets** to enclose the entries of matrix.

We usually denote matrices by capital letters. For instance, the general matrix with m rows and n columns is

$$A = \begin{bmatrix} a_{11} & a_{12}\ldots\ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22}\ldots\ldots\ldots\ldots\ldots a_{2n} \\ \vdots & \vdots \qquad\qquad\qquad \vdots \\ a_{m1} & a_{m2}\ldots\ldots\ldots\ldots\ldots a_{mn} \end{bmatrix} = \left[a_{ij} \right], \text{ in brief.}$$

Here a_{11} denotes the element lying in the 1st row and the 1st column, a_{12} denotes the element lying in the 1st row and the 2nd column, and, in general, a_{ij} denotes the element lying in the ith row and the jth column. We also say that a_{ij} **is the (i, j)th entry of A.**

Thus, the (1, 3)th entry of $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 3 & 0.3 & -1 \end{bmatrix}$ is 0, and the (3, 1)th entry is 3.

We can also denote a matrix A consisting of m rows and n columns by $A_{m \times n}$ or $A^{(m,n)}$; and we say that it **has order m × n** or **is an m × n matrix.**

We will often refer to the ith row of a matrix, meaning the ith row from the top. Similarly, the ith column of a matrix refers to the ith column, counting from left to right.

If the number of rows in a matrix equals the number of columns, the matrix is called a **square matrix.** Isn't the name appropriate?

Some examples of square matrices are [2], $\begin{bmatrix} 3 & 5 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Such matrices are very important in matrix theory.

You may like to try some exercises on matrices now.

E1) Write the order of each of the following matrices.

$$\begin{bmatrix} 1 & 0 & 2 \\ -4 & i & 9 \\ 3 & 0 & 8i \end{bmatrix}, \; [5], \; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}, \; [2 - 2i].$$

Also write their (2, 3)th and (1, 1)th entries and their 3rd columns, if they exist.

E2) Write down a 3 × 4 matrix in which the (i, j)th entry is 0 whenever $i < j$, and non-zero otherwise.

E3) a) Rewrite the linear systems that you obtained in E5 and E6 of Unit 4, in matrix notation.

b) What would happen to these matrices if the first and second equations in each of the linear systems were interchanged?

Now, go back to the system of equations (1). We rewrote them in matrix notation in (2). We can also write (2) in shorthand notation as AX=B, where A is the 3 × 3 coefficient matrix, X is the 3 × 1 matrix $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and B is the 3 × 1 matrix $\begin{bmatrix} 5 \\ -6 \\ 1 \end{bmatrix}$.

The coefficient matrix of a linear system is the matrix formed by taking the coefficients of the equations in the system.

If the number of equations in a linear system equals the number of variables, the coefficient matrix will be a square matrix. In this situation we can sometimes use the concept of a determinant to solve the system. Let us see what this concept is.

5.3 DETERMINANTS

You have just seen that we can represent a set of n linear equations in n variables by a matrix equation AX=B, where A is an n×n matrix. Associated with this square matrix A, we can define a unique number — its determinant. In this section we will discuss determinants of matrices whose elements are real numbers. We will also discuss some of their properties.

Let us start with a definition.

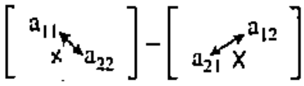
Definition: The determinant of a 1 × 1 matrix A = [a], denoted by **| A | or det (A),** is a.

For example, if A = [3], then | A | is 3. Similarly, if $A = \left[-\frac{1}{2}\right]$, then $|A| = -\frac{1}{2}$.

Now let us consider the determinant of a 2 × 2 matrix.

Definition: The determinant of the 2 × 2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is the number $a_{11}a_{22} - a_{12}a_{21}$, and is denoted by $|A|$.

This is simply the cross multiplication



For example, if $A = \begin{bmatrix} 0 & 2 \\ -1 & 5 \end{bmatrix}$, then $|A| = 0 \times 5 - 2 \times (-1) = 2$.

Another common way of writing $|A|$ is to write its elements within parallel vertical lines, instead of within square brackets. For example, we can write the determinant of $\begin{bmatrix} 0 & 2 \\ -1 & 5 \end{bmatrix}$ as $\begin{vmatrix} 0 & 2 \\ -1 & 5 \end{vmatrix}$.

We will often see this way of writing a determinant.

You may like to calculate some determinants now.

E4) Evaluate

(a) $\begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix}$, (b) $\begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix}$, (c) $\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix}$,

(d) $\begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}$, (e) $\begin{vmatrix} 2 & -2 \\ -1 & 0 \end{vmatrix}$, (f) $\begin{vmatrix} 1 & -1 \\ -4 & 0 \end{vmatrix}$.

Compare the determinants obtained in (a) and (e), and (a) and (f).

Now let us use determinants of 2 × 2 matrices to obtain the determinant of a 3 × 3 matrix.

Definition: The **determinant of the 3 × 3 matrix**

The determinant of A is denoted by |A| or det (A).

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ is}$$

$$|A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= (-1)^{1+1}a_{11}|A_{11}| + (-1)^{1+2}a_{12}|A_{12}| + (-1)^{1+3}a_{13}|A_{13}|,$$

where A_{1j} is the matrix obtained from A after deleting the first row and the jth column, for j = 1, 2, 3.

In obtaining $|A|$, we **expanded by** (or **along**) **the first row.** We could also have expanded by the second row, third row, or either of the columns. So, for example, expanding along the third column, we get

$$|A| = (-1)^{1+3} a_{13}|A_{13}| + (-1)^{2+3} a_{23}|A_{23}| + (-1)^{3+3} a_{33}|A_{33}|,$$

where A_{i3} denotes the matrix obtained from A after deleting the ith row and the third column.

All 6 ways of obtaining $|A|$ lead to the same value. We will not prove this here. However, let us consider an example.

Let $A = \begin{bmatrix} 1 & 2 & 6 \\ 5 & 4 & 1 \\ 7 & 3 & 2 \end{bmatrix}$. Then, expanding by the first row, we get

$$\begin{vmatrix} 1 & 2 & 6 \\ 5 & 4 & 1 \\ 7 & 3 & 2 \end{vmatrix} = (-1)^{1+1}.1.\begin{vmatrix} 4 & 1 \\ 3 & 2 \end{vmatrix} + (-1)^{1+2}.2.\begin{vmatrix} 5 & 1 \\ 7 & 2 \end{vmatrix} + (-1)^{1+3}.6.\begin{vmatrix} 5 & 4 \\ 7 & 3 \end{vmatrix}$$

$$= (4 \times 2 - 1 \times 3) - 2\,(5 \times 2 - 1 \times 7) + 6\,(5 \times 3 - 4 \times 7)$$

$$= -79.$$

Now, why don't you try this exercise?

E5) Obtain $|A|$, for A in the example above, by expanding along the 3rd row and by expanding along the 2nd column.

As you have seen, we define the determinant of a 3×3 matrix in terms of the determinants of 2×2 matrices. In the same way we can obtain the determinant of any $n \times n$ square matrix $(n \geq 2)$ in terms of the determinants of $(n-1) \times (n-1)$ square matrices.

Definition: The **determinant of an $n \times n$ matrix** $A = [a_{ij}]$, where $n > 1$, is given by

$$|A| = (-1)^{1+1} a_{11} |A_{11}| + (-1)^{1+2} a_{12} |A_{12}| + \ldots + (-1)^{1+n} a_{1n} |A_{1n}|,$$

where A_{1j} = matrix obtained from A on deleting the 1st row and the jth column, $\forall$ j = 1,...., n.

What we have stated for the 3×3 case is true for the $n \times n$ case $(n \geq 2)$, namely, that we can expand along any row or column to obtain the determinant of A. Thus,

$$|A| = (-1)^{i+1} \cdot a_{i1} |A_{i1}| + (-1)^{i+2} \cdot a_{i2} |A_{i2}| + \ldots + (-1)^{i+n} \cdot a_{in} |A_{in}| \; \forall\, i = 1, \ldots, n,$$

and

$$|A| = (-1)^{1+j} \cdot a_{1j} |A_{1j}| + (-1)^{2+j} \cdot a_{2j} |A_{2j}| + \ldots + (-1)^{n+j} \cdot a_{nj} |A_{nj}| \; \forall\, j = 1, \ldots, n.$$

We call the determinant of an $n \times n$ matrix **a determinant of order** n (or size n).

So far we have spent some time on evaluating determinants of orders 1, 2 and 3. In this course we shall not go to higher orders. (They are discussed in great detail in our course on linear algebra.) We will only introduce you to some elementary properties of determinants now. While doing E4 you may have realised some of them. These properties help us in evaluating a determinant in a shorter time. Let us see what they are.

Theorem 1 : Let A be a square matrix. Then $|A|$ satisfies the following properties.

P1: If all the elements of any row or column of A are zero, then $|A| = 0$.

P2: If B is the matrix obtained from A by interchanging any two rows (or any two columns), then $|B| = -|A|$.

P3: If B is the matrix obtained from A by multiplying all the elements of a row (or all the elements of a column) of A by a number c, then $|B| = c\,|A|$.

P4: If B is the matrix obtained from A by adding the multiple of a row to another row (or the multiple of a column to another column), then $|B| = |A|$.

A multiple of a row (or of a column) by a non-zero number k is the row (or column) obtained by multiplying each of its entries by k.

P5: If two rows (or two columns) of A are equal, then $|A| = 0$.

We shall not prove these properties here. If you are interested in the proofs, you can study Block 3 of our next level course on linear algebra. In this course we shall only see how to apply these properties. Let us first verify them in some cases.

If $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$, then $|A| = 1 \times 0 - 2 \times 0 = 0$ (an example of P1).

If $A = \begin{bmatrix} 3 & 0 \\ -2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 3 \\ 5 & -2 \end{bmatrix}$, that is, B is obtained by interchanging the columns of A, then $|A| = 15$ and $|B| = -15 = -|A|$ (an example of P2).

If $A = \begin{bmatrix} 1 & -1 \\ 2 & 6 \end{bmatrix}$ and B is obtained by multiplying the second row of A by 5, that is,

$B = \begin{bmatrix} 1 & -1 \\ 10 & 30 \end{bmatrix}$, then $|A| = 8$ and $|B| = 40 = 5|A|$ (an example of P3).

Now, let us take an example which satisfies the hypothesis of P4. Let $A = \begin{bmatrix} 1 & 2 \\ -3 & -1 \end{bmatrix}$

and $B = \begin{bmatrix} 4 & 3 \\ -3 & -1 \end{bmatrix}$. Can you make out if B is related to A in any manner suggested in P4? How have we got B? We multiplied the elements in the second row of A by (–1) and added them to the corresponding elements in the first row, that is, we subtracted the second row of A from the first row.

Thus, $B = \begin{bmatrix} 1+(-3)(-1) & 2+(-1)(-1) \\ -3 & -1 \end{bmatrix}$.

Now, $|A| = (1)(-1) - (2)(-3) = 5$ and $|B| = 5 = |A|$. So P4 seems to work in this case.

Now, suppose we add 3 times the first column of A to the second column. We get the matrix

$$C = \begin{bmatrix} 1 & 2+3 \\ -3 & (-1)+(-9) \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -3 & -10 \end{bmatrix}$$

Then, $|C| = |A|$, verifying P4 again.

Now let us see if P5 holds for a general 2 × 2 matrix satisfying the hypothesis. One such matrix is

$$A = \begin{bmatrix} a & a \\ b & b \end{bmatrix}, \text{ a, b} \in \mathbf{R}.$$

Then $|A| = ab - ab = 0$, verifying P5.

Now, let us try and obtain the determinant of the 3 × 3 matrix

$$A = \begin{bmatrix} 3 & -5 & 5 \\ 2 & 1 & -1 \\ 3 & 9 & -9 \end{bmatrix}$$

The third column is (–1) times the second column. So let us add the second column to the third column. By P4 we get

$$|A| = \begin{vmatrix} 3 & -5 & 0 \\ 2 & 1 & 0 \\ 3 & 9 & 0 \end{vmatrix} = 0, \text{ by P1.}$$

So, you have seen some ways in which P1 – P5 can be used for calculating determinants. Why don't you try this exercise now?

E6) Using P1 to P5, evaluate the determinants of

a) $A = \begin{bmatrix} a & b & c \\ 0 & 0 & 0 \\ d & e & f \end{bmatrix}$, where a, b, c, d, e, f ∈ R

b) $B = \begin{bmatrix} -1 & 2 & 1 \\ -3 & 6 & 3 \\ 1 & 5 & 1 \end{bmatrix}$,

c) $C = \begin{bmatrix} 1 & 5 & 1 \\ -3 & 6 & 3 \\ -1 & 2 & 1 \end{bmatrix}$,

d) $D = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 6 & 3 \\ 2 & 5 & 1 \end{bmatrix}$.

e) $E = \begin{bmatrix} a & 2a & d \\ b & 2b & e \\ c & 2c & f \end{bmatrix}$, where a, b, c, d, e, f $\in$ **R**.

As we said earlier, the importance of the properties P1 – P5 lies in their use for decreasing the effort involved in computing a determinant. You must have realised this while doing E6. Let us look at another example of their use. For our convenience we **will use R_i to denote the ith row and C_i to denote the ith column.**

Let $A = \begin{bmatrix} 2 & 3 & 6 \\ 6 & 5 & -5 \\ 1 & 1 & 3 \end{bmatrix}$. To evaluate |A| by expanding along any row or column will require us to evaluate 3 determinants of order 2. But, if we use P4, we can multiply R_3 by (–2) and add it to R_1, to get

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 6 & 5 & -5 \\ 1 & 1 & 3 \end{bmatrix}.$$

Now, R_1 has two zeros in it. So we can obtain | B |, which is the same as |A| by expanding along R_1. This means that we only need to evaluate | B_{12} |.
Thus, $|A| = |B| = (-1).1.|B_{12}| = -23$.

The following remark will be very useful to you for evaluating a determinant.

Remark 1: Whenever you have to compute the determinant of a matrix, it is best to expand along the row or column with the maximum number of zeros. Therefore, one should use the property P4 so as to get as many zeros as possible in some row or column.

You may like to use the remark above to obtain the following interesting results.

E7) Let $A = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$ and $B = \begin{bmatrix} p & 0 & 0 \\ q & s & 0 \\ r & t & u \end{bmatrix}$,

A and B are the general forms of 3×3 triangular matrices over **R**.

where a, b, ..., f, p, q, ..., u $\in$ **R**.

Show that

a) |A| = adf, that is, the product of the elements lying on the **principal diagonal**

b) |B| = psu.

E8) Let $C = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, where a, b, c $\in$ **R**.

C is the general 3×3 diagonal matrix over **R**.

Show that |C| = abc.

E9) Obtain $\begin{vmatrix} 3 & -2 & 4 \\ 6 & 8 & 1 \\ -9 & 6 & 12 \end{vmatrix}$

E10) Show that the analogues of E7 and E8 are true for general 2 × 2 triangular and diagonal matrices.

Now consider $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. What is | I | ? We use E8, and find | I | = 1.

And $1 \neq 0$. So $| I | \neq 0$. Because of this property of $| I |$, I belongs to the class of matrices that we will now define.

Definition: A square matrix A is called **non-singular** if $| A | \neq 0$. Otherwise A is called **singular.**

You can check that some more examples of non-singular matrices are [5], $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$ and $\begin{bmatrix} 2 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$; and $\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$ is an example of a singular matrix.

Why don't you try some exercises now?

E11) Which of the following matrices are non-singular? Give reasons for your choice.

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 5 \\ 0 & 0 \end{bmatrix}, [-3], \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ -4 & 9 \end{bmatrix}, \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & 6 \\ \sqrt{2} & 1 & -1 \\ 0 & 0 & 5 \end{bmatrix}.$$

E12) When are [a], $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\begin{bmatrix} a & b & c \\ d & e & d \\ f & g & f \end{bmatrix}$ singular? Here a, b, c, d, e, f, g $\in$ **R**.

Non-singular matrices form an important part of matrix theory. In the next section we shall introduce you to a rule for solving any system of linear equations whose coefficient matrix is non-singular.

5.4 CRAMER'S RULE

Fig. 1 : Cramer (1704-1752)

In 1750 the German mathematician Gabriel Cramer published a rule for solving a set of n linear equations in n unknowns simultaneously. Though this rule is named after Cramer, it seems to have been discovered by the British mathematician Colin Maclaurin twenty years earlier.

Let us see what this rule is.

Consider the general system of 2 equations in 2 unknowns:

$$ax + by + c = 0$$

$$dx + ey + f = 0,$$

where $ae - db \neq 0$.

Then, if you use the substitution method, what solution do you get? We get $x = \frac{bf - ce}{ae - db}$, $y = \frac{cd - af}{ae - db}$.

Notice that this is the same as

$$x = \frac{\begin{vmatrix} -c & b \\ -f & e \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a & -c \\ d & -f \end{vmatrix}}{\begin{vmatrix} a & b \\ d & e \end{vmatrix}}$$

But how did we get x and y in the determinant form? What we did was to first write the system of equations as

$$ax + by = -c$$

$$dx + ey = -f,$$

that is, AX = B, where A is the coefficient matrix $\begin{bmatrix} a & b \\ d & e \end{bmatrix}$, X is $\begin{bmatrix} x \\ y \end{bmatrix}$ and B is the matrix of constant terms, $\begin{bmatrix} -c \\ -f \end{bmatrix}$.

Then we calculated D = | A |, which is given to be non-zero. After that we calculated D_1, the determinant of the matrix obtained from A by replacing the first column by B; thus, $D_1 = \begin{vmatrix} -c & b \\ -f & e \end{vmatrix}$. Similarly, we calculated D_2, the determinant of the matrix obtained from A by replacing the second column by B; thus, $D_2 = \begin{vmatrix} a & -c \\ d & -f \end{vmatrix}$. Then $x = \frac{D_1}{D}$ and $y = \frac{D_2}{D}$.

Cramer extended this result to a system of n linear equations in n unknowns.

Let us consider his general rule.

Cramer's Rule: Consider the following linear system of n equations in n unknowns :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots\dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots\dots + a_{2n}x_n &= b_2 \\ \vdots \qquad & \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots\dots + a_{nn}x_n &= b_n \end{aligned}$$

that is, AX = B, where

$$A = [a_{ij}],\ X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix},\ B = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{bmatrix}.$$

For i=1,.....,n, define A_i to be the matrix obtained from A after substituting B for the ith column of A.

Define $D_i = | A_i |\ \forall\ i = 1, \dots\dots, n$, and

D = | A |.

Then, if **D ≠ 0**, the system has the unique solution

$$x_1 = \frac{D_1}{D},\ x_2 = \frac{D_2}{D}, \dots\dots\dots, x_n = \frac{D_n}{D}.$$

Does all this seem too much to take in? Don't worry. In this course we shall only be applying the rule for n = 2 or 3.

Just remember that

> Cramer's rule can only be applied if
>
> i) the number of equations in the linear system equals the number of variables; and
>
> ii) the determinant of the coefficient matrix is non-zero.

Let us apply Cramer's rule in an example. Consider the system

$$\begin{aligned} 2x - 3y + z &= 1 \\ x + y + z &= 2 \\ 3x - 4z - 17 &= 0 \end{aligned}$$

We first rewrite the system as

$$\begin{aligned} 2x - 3y + z &= 1 \\ x + y + z &= 2 \\ 3x - 0.y - 4z &= 17. \end{aligned}$$

This is of the form AX = B, where

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 1 & 1 \\ 3 & 0 & -4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 \\ 2 \\ 17 \end{bmatrix}.$$

Now let us first see whether $|A| = 0$ or not. Let us expand along the third row. We get

$$D = |A| = 3\begin{vmatrix} -3 & 1 \\ 1 & 1 \end{vmatrix} - 4\begin{vmatrix} 2 & -3 \\ 1 & 1 \end{vmatrix} = -32 \neq 0.$$

So we can go ahead and apply Cramer's rule. For this we evaluate

$$D_1 = \begin{vmatrix} 1 & -3 & 1 \\ 2 & 1 & 1 \\ 17 & 0 & -4 \end{vmatrix} = -96$$

$$D_2 = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 17 & -4 \end{vmatrix} = -32$$

$$D_3 = \begin{vmatrix} 2 & -3 & 1 \\ 1 & 1 & 2 \\ 3 & 0 & 17 \end{vmatrix} = -64.$$

Then $x = \frac{D_1}{D} = 3$, $y = \frac{D_2}{D} = 1$, $z = \frac{D_3}{D} = -2$.

On verifying, we find that (3, 1, −2) is indeed the solution of the given system.

You may like to try your hand at applying Cramer's rule now.

E13) Solve the following systems by Cramer's rule, if applicable. Otherwise use the Gaussian elimination method (see Sec. 4.4).

a) $x + y + 1 = 0$
$2x - y = 7.$

b) $x + y - z + 2 = 0$
$2x - y + z + 5 = 0$
$x - 2y + 3z - 4 = 0.$

c) $3x + 5y + 2z = 1$
$4x + y - 7 = 0$
$9x + 15y + 6z = 3.$

E14) Consider the $n \times n$ linear system

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0$$
$$\vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = 0$$

If A is the coefficient matrix and $|A| \neq 0$, can we apply Cramer's rule to solve the system? If so, obtain the solution set.

E15) I have Rs. 2480/- in five, ten and twenty rupee notes. The total number of notes is 290 and all the ten rupee notes add up to Rs. 60/- more than the sum of the twenty rupee notes. How many of each type of note do I have?

In this unit and the previous one we have discussed three methods for solving linear systems. The examples and exercises we did involved a maximum of 3 equations and a maximum of four unknowns. But, in practical applications in the sciences and social sciences, one needs to solve very large systems. These may consist of 15, 20 or more equations in as many or more variables. As you may have guessed, these systems require computers for solving them. Then the best method to apply is the Gaussian elimination method. We have discussed this method as well as Cramer's rule in greater detail in our course on linear algebra.

Now why don't you try solving the following exercise by any of the three methods we have covered in this unit and the previous one?

E16) Which of the following linear systems are consistent? Obtain the solution sets, wherever possible.

a) $2x - 5y + 7z = 6$
$x - 3y + 4z = 3$
$3x - 8y + 11z = 11.$

b) $2x - y + 3z - 5w = -7$
$-7y + 3z - 7w = -13$
$3x + 4y + 2z = 0.$

c) $x - y + z = 0$
$-3x + y - 4z = 0$
$7x - 3y - 9z = 0$
$4x - 2y - 5z = 0.$

d) $x - 2y + z = 6$
$3x + y - 4z = -7$
$5x - 3y + 2z = 5.$

Let us now summarise the contents of this unit.

5.5 SUMMARY

In this unit we

1) defined an $m \times n$ matrix, and a square matrix, in particular.

2) introduced you to the concept of a determinant of a square matrix.

3) discussed some properties of determinants.

4) used the definition and properties of determinants to evaluate determinants of orders 1,2 and 3.

5) defined a non-singular matrix.

6) applied Cramer's rule for solving a linear system of equations whose coefficient matrix is non-singular.

With this unit we finish our discussion on simultaneous linear equations. In the next unit we shall look at some commonly used inequalities. But before going to Unit 6, **please go back to the objectives given in Sec. 5.1** and check if you have achieved them. You may also like to go through the next section, in which we have given our solutions to the exercises in the unit. This may be useful to you for counterchecking your solutions.

5.6 SOLUTIONS/ANSWERS

E1) Their orders are 3×3, 1×1, 3×3, 3×1 and 1×3, respectively.
The (2, 3)th and (1, 1)th entries of the first matrix are 9 and 1.
The (1, 1)th entry of the second one is 5. It has no (2, 3)th entry.
The (2, 3)th and (1, 1)th entries of the third one are 0 and 1.
The (1, 1)th entry of the fourth one is 4; it has no (2, 3)th entry.
The (1, 1)th entry of the fifth one is 2; it has no (2,3)th entry.

Only the first and third matrices have third columns, which are $\begin{bmatrix} 2 \\ 9 \\ 8i \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, respectively.

E2) The required matrix will be of the form

$$\begin{bmatrix} a & 0 & 0 & 0 \\ b & c & 0 & 0 \\ d & e & f & 0 \end{bmatrix}, \text{ where } a, \ldots, f \in \mathbf{R} \setminus \{0\}.$$

E3) a) $\begin{bmatrix} 3 & 5 & 2 \\ 1 & 7 & 3 \\ 0 & 2 & 0 \\ 1 & 6 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 55 \\ 45 \\ 10 \\ 45 \end{bmatrix}$ gives us the system in E5.

$\begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 150 \\ 30 \end{bmatrix}$ gives us the system in E6.

b) The first two rows of both the coefficient matrices, as well as of the matrices on the right hand sides, would be interchanged.

E4) a) $1 \times 0 - (-1)(-1) = -1$

b) $1 \times 2 - 1 \times 2 = 0$

c) 0

d) $-i^2 = 1$

E5) By expanding along the 3rd row, we get

$|A| = (-1)^{3+1}.7.|A_{31}| + (-1)^{3+2}.3.|A_{32}| + (-1)^{3+3}.\ 2.\ |A_{33}|.$

$$= 7\begin{vmatrix} 2 & 6 \\ 4 & 1 \end{vmatrix} - 3\begin{vmatrix} 1 & 6 \\ 5 & 1 \end{vmatrix} + 2\begin{vmatrix} 1 & 2 \\ 5 & 4 \end{vmatrix}$$

$= -79.$

By expanding along the second column, we get

$|A| = (-1)^{2+1}.2.|A_{21}| + (-1)^{2+2}.4.|A_{22}| + (-1)^{2+3}.3.|A_{23}|$

$= -79.$

E6) a) 0 by P1.

b) $|B| = 3\begin{vmatrix} -1 & 2 & 1 \\ -1 & 2 & 1 \\ 1 & 5 & 1 \end{vmatrix}$, by P3

$= 3.0$, by P5

$= 0.$

c) C is obtained from B above, by interchanging the first and third rows.
$\therefore |C| = -|B| = 0.$

d) D = 0, as in (b).

e) $|E| = 2\begin{vmatrix} a & a & d \\ b & b & e \\ c & c & f \end{vmatrix} = 0$, by P3.

E7) a) We expand along a row or column which has the maximum number of zeros. So, expanding along C_1, we get

$|A| = (-1)^{1+1}.\ a.\ |A_{11}| = a = \begin{vmatrix} d & e \\ 0 & f \end{vmatrix} = adf.$

b) We can expand along R_1 to get

$|B| = psu.$

Note that we would get the same answer if we'd expanded along any other row or column.

E8) Expand along R_1 to get $|C| = abc$.

E9) Let $A = \begin{bmatrix} 3 & -2 & 4 \\ 6 & 8 & 1 \\ -9 & 6 & 12 \end{bmatrix}$

We want to make some entries zero. Looking at C_1, you may have noticed that replacing R_2 by $R_2 + (-2) R_1$ and R_3 by $R_3 + 3R_1$ will make the (2, 1)th and (3, 1)th entries zero. Then, by P4 we get

$$|A| = \begin{vmatrix} 3 & -2 & 4 \\ 6+(-2)3 & 8+(-2)(-2) & 1+(-2)(4) \\ -9+3(3) & 6+3(-2) & 12+3(4) \end{vmatrix}$$

$$= \begin{vmatrix} 3 & -2 & 4 \\ 0 & 12 & -7 \\ 0 & 0 & 24 \end{vmatrix}$$

Now, by E7, $|A| = 3 \times 12 \times 24 = 864$.

E10) The general 2×2 real triangular matrix $A = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$ or $B = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$, where a, b, c $\in$ **R**. In both cases their determinant is ac, the product of the diagonal elements. The general 2×2 real diagonal matrix is $C = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where a, b $\in$ **R**.

$|C| = ab$, the product of the diagonal elements.

The diagonal elements of an $n \times n$ matrix are its (i, i) th entries $\forall$ i = 1, . . . , n.

E11) $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$; hence, the first matrix is non-singular.

$\begin{bmatrix} 1 & 5 \\ 0 & 0 \end{bmatrix}$ is singular, since its determinant is zero.

$|[-3]| = -3 \neq 0$; thus, [−3] is non-singular.

$\begin{bmatrix} 1 & 3 \\ 2 & 1 \\ -4 & 9 \end{bmatrix}$ is not a square matrix, and hence can't be non-singular. Note that it is not singular either, since a singular matrix has to be square too.

The last matrix has zero determinant, and hence is singular.

E12) [a] is singular iff a = 0.

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is singular iff $ad - bc = 0$.

$\begin{bmatrix} a & b & c \\ d & e & d \\ f & g & f \end{bmatrix} = (a - c)(ef - dg)$. This is zero iff a = c or ef = dg. That is, the given matrix is singular iff a = c or $\begin{bmatrix} e & d \\ g & f \end{bmatrix}$ is singular.

E13) a) $x + y = -1$

$2x - y = 7$

is the same as $\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 7 \end{bmatrix}$.

Since $D = \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3 \neq 0$, Cramer's rule can be applied.

Now, $D_1 = \begin{vmatrix} -1 & 1 \\ 7 & -1 \end{vmatrix} = -6$ and

$D_2 = \begin{vmatrix} 1 & -1 \\ 2 & 7 \end{vmatrix} = 9$.

Then $x = \frac{D_1}{D} = 2$ and $y = \frac{D_2}{D} = -3$.

Thus, the solution is (2, −3).

b) The given system is equivalent to

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \\ 4 \end{bmatrix}$$

Now $D = \begin{vmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \\ 1 & -2 & 3 \end{vmatrix} = -3 \neq 0$. So we can apply Cramer's rule. We calculate

$$D_1 = \begin{vmatrix} -2 & 1 & -1 \\ -5 & -1 & 1 \\ 4 & -2 & 3 \end{vmatrix} = 7,$$

$$D_2 = \begin{vmatrix} 1 & -2 & -1 \\ 2 & -5 & 1 \\ 1 & 4 & 3 \end{vmatrix} = -22,$$

$$D_3 = \begin{vmatrix} 1 & 1 & -2 \\ 2 & -1 & -5 \\ 1 & -2 & 4 \end{vmatrix} = -21.$$

Therefore, $x = -\frac{7}{3}$, $y = \frac{22}{3}$, $z = 7$.

c) The system is equivalent to

$$\begin{bmatrix} 3 & 5 & 2 \\ 4 & 1 & 0 \\ 9 & 15 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix}$$

Note that the 3rd row of the coefficient matrix is 3 times the 1st row. Thus, its determinant is zero. So we can't apply Cramer's rule. Let us solve the system by successive elimination.

The system is

$3x + 5y + 2z = 1$ ....(3)
$4x + y = 7$ ....(4)
$9x + 15y + 6z = 3$ ....(5)

Note that (5) is equivalent to (3). So we can drop (5). Now let us eliminate y from (3) and (4). For this we calculate (3) −5 × (4), which is

$-17x + 2z = -34$ ....(6)

Now, we can't eliminate any further, between (3), (4) and (6). So let us use (4) and (6) to write y and z in terms of x.

$(4) \Rightarrow y = 7 - 4x$, and

$(6) \Rightarrow z = \frac{17}{2}x - 17$.

Thus, we get a 1-parameter set of infinitely many solutions,

$$\left\{ \left(x, 7 - 4x, \frac{17}{2}x - 17\right) \mid x \in R \right\}.$$

E14) Since $|A| \neq 0$, we can apply Cramer's rule. Now, to apply this rule, we calculate D_i

for i=1, ..., n by substituting the ith column with the constant column $\begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$.

Thus, each $D_i = 0$, by P1 of Theorem 1.

Thus, $x_i = 0 \ \forall \ i = 1, \ldots, n$.

The solution (0, 0, ..., 0) is called the **trivial solution.**

E15) Let x, y, z denote the number of five, ten and twenty notes, respectively. Then we know that

$5x + 10y + 20z = 2480$

$x + y + z = 290$

$10y - 20z = 60$

We can solve this by Cramer's rule to get

$x = 164, y = 86, z = 40$.

E16) a) This is a 3 × 3 system. First let us see if we can apply Cramer's rule. Since the determinant of the coefficient matrix is zero, we can't apply Cramer's rule.

So let us try to solve it by elimination.

Adding the first two equations, we get

$3x - 8y + 11z = 9$.

Subtracting this from the third equation in the system, we get 0 = 2, a false statement.

Thus the system is inconsistent.

b) The given system is

$2x - y + 3z - 5w = -7$ ....(7)

$- 7y + 3z - 7w = -13$ ....(8)

$3x + 4y + 2z = 0$ ....(9)

We shall try and solve by elimination.

To eliminate w from (7) and (8), we calculate 7 × (7) – 5 × (8)

We get

$14x + 28y + 6z = 16$, that is,

$7x + 14y + 3z = 8$ ....(10)

Eliminating z from (9) and (10), we get

$5x + 16y = 16$ ....(11)

Now, we can't eliminate any further. So we shall try and obtain all the variables in terms of the minimum number of variables possible.

$(11) \Rightarrow x = \frac{16}{5}(1-y)$.

Then $(9) \Rightarrow \frac{48}{5}(1-y) + 4y + 2z = 0 \Rightarrow z = \frac{14y - 24}{5}$.

Then (8) $\Rightarrow -7y + \frac{3}{5}(14y - 24) - 7w = -13 \Rightarrow w = \frac{1}{5}(y-1)$.

Thus, we get a 1-parameter set of solutions, namely,

$$\left\{\left(\frac{16(1-y)}{5}, y, \frac{14y-24}{5}, \frac{y-1}{5}\right) \middle| y \in \mathbf{R}\right\}.$$

To check that these are the solutions, we substitute the 4-tuple $\left(\frac{16}{5}(1-y), y, \frac{1}{5}(14y-24), \frac{1}{5}(y-1)\right)$ in each of the equations of the system and find that it satisfies them.

c) Since the system is a 4×3 system, we shall apply the Gaussian elimination method.

$$x - y + z = 0 \quad \text{........(12)}$$

$$-3x + y - 4z = 0 \quad \text{........(13)}$$

$$7x - 3y - 9z = 0 \quad \text{........(14)}$$

$$4x - 2y - 5z = 0 \quad \text{........(15)}$$

$$3 \times (12) + (13) \Rightarrow -2y - z = 0 \Rightarrow 2y + z = 0 \quad \text{........(16)}$$

$$7 \times (12) - (14) \Rightarrow -4y + 16z = 0 \Rightarrow y - 4z = 0 \quad \text{........(17)}$$

$4 \times (16) + (17) \Rightarrow y = 0$.

Then (17) $\Rightarrow z = 0$.

Then (12) $\Rightarrow x = 0$.

We check that (0, 0, 0) satisfies all the equations. Thus the system only has the trivial solution.

d) Since the determinant of the coefficient matrix is non-zero, we can apply Cramer's rule as well as the elimination method. Let us apply Cramer's rule. For this we calculate

$$D = \begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -4 \\ 5 & -3 & 2 \end{vmatrix} = 28,$$

$$D_1 = \begin{vmatrix} 6 & -2 & 1 \\ -7 & 1 & -4 \\ 5 & -3 & 2 \end{vmatrix} = -32,$$

$$D_2 = \begin{vmatrix} 1 & 6 & 1 \\ 3 & -7 & -4 \\ 5 & 5 & 2 \end{vmatrix} = -100,$$

$$D_3 = \begin{vmatrix} 1 & -2 & 6 \\ 3 & 1 & -7 \\ 5 & -3 & 5 \end{vmatrix} = 0$$

$\therefore\ x = \frac{-32}{28}$, $y = \frac{-100}{28}$, $z = 0$, that is, the unique solution is $\left(\frac{-8}{7}, \frac{-25}{7}, 0\right)$.

UNIT 6 INEQUALITIES

Structure

6.1 INTRODUCTION

So far we have discussed equations of various kinds. Now we shall consider some **inequalities**; not of the social kind, but between real numbers. A mathematical inequality is a mathematical expression of the condition that of two quantities one is greater than, greater than or equal to, less than or less than or equal to the other. An inequality that holds for every real number is called an **absolute inequality**. In this unit we shall restrict ourselves to such inequalities.

We will discuss six famous absolute inequalities. We have divided them into two sections —those that have been used for centuries and those that were discovered by some famous nineteenth century European mathematicians. These inequalities have several applications also. We will discuss a few of them. You may come across some applications in other courses too, at which time we hope that you will find that you didn't study this unit in vain!

Let us list our unit objectives now.

Objectives

After reading this unit you should be able to prove and apply

- the inequality of the means;
- the triangle inequality;
- the Cauchy-Schwarz (Bunyakovskii) inequality;
- Weierstrass' inequalities;
- Tchebychev's inequalities.

Let us discuss the inequalities one by one.

6.2 INEQUALITIES KNOWN TO THE ANCIENTS

In this section we shall discuss two inequalities handed down to us by ancient mathematicians. But first we will give a list of some properties of inequalities you must be familiar with. They are the following:

for a, b, c, d $\in$ **R**

i) $a \geq b, c \geq 0 \Rightarrow ac \geq bc$

ii) $a \geq b \Leftrightarrow -a \leq -b$

iii) $a \geq b \Leftrightarrow \frac{1}{a} \leq \frac{1}{b}$ provided $a \neq 0$, $b \neq 0$.

iv) $a \geq b$, $c \geq d \Rightarrow a+c \geq b+d$

v) $a^n \geq b^n$, $a \geq 0 \Rightarrow a \geq b$, where $n \in N$.

We will often use these properties implicitly while proving the inequalities mentioned in the unit objectives.

Now let us discuss the inequality that relates three averages.

6.2.1 Inequality of the Means

An important part of arithmetic that can be traced back to the Babylonians and Pythagoreans (approximately 6th century B.C.) is the theory of means or averages. The word "average" comes from the Latin word "havaria", which was the insurance paid to compensate for damage to goods in transit in the olden days. All of us are familiar with the term "average". In fact, all of us must have often calculated the average of a finite set of numbers by adding them up and dividing the sum by the total number of these numbers. But this is only one of many types of averages. We will discuss three of these types here. Let us start with the "usual" average.

Definition: The **arithmetic mean (AM)** of n real numbers $x_1, x_2, \ldots\ldots, x_n$ is

$$\frac{x_1 + x_2 + \ldots\ldots + x_n}{n}, \text{ that is, } \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right).$$

For example, the AM of $\frac{1}{2}$, $\frac{-1}{3}$ and 0 is $\dfrac{\frac{1}{2}-\frac{1}{3}+0}{3} = \frac{1}{18}$.

The AM is often used in statistics for studying data.

Another type of average is the geometric mean. This is the best mean to use if we want to find the mean of any finite set of positive numbers that follow geometric progression. Thus, this mean is very useful for studying population growth. Let us see how the geometric mean is defined.

Definition: The **geometric mean (GM)** of n positive real numbers

$x_1, x_2, \ldots\ldots, x_n$ is

$$(x_1 x_2 \ldots\ldots x_n)^{1/n}, \text{ that is, } \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}.$$

For example, the GM of 3 and 4 is $\sqrt{3 \times 4} = \sqrt{12}$, and the GM of 2, 4 and 8 is $(2 \times 4 \times 8)^{1/3} = 4$.

Yet another kind of average of numbers is their harmonic mean, which we now define.

Definition: The **harmonic mean (HM)** of n non-zero real numbers

$x_1, x_2, \ldots\ldots, x_n$ is

$$\frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}.$$

Thus, the HM of $x_1, x_2, \ldots\ldots, x_n$ is the inverse of the AM of $\frac{1}{x_1}, \frac{1}{x_2}, \ldots\ldots, \frac{1}{x_n}$.

For example, the HM of -2, $\frac{1}{3}$ and 7 is $\dfrac{3}{-\frac{1}{2}+3+\frac{1}{7}} = \frac{42}{37}$.

The HM is the most appropriate type of average to use when we want to find the average rate of a set of varying rates. Thus, it is the best average to use for obtaining the average velocity of a vehicle covering various distances at different speeds.

At this point we would like to make a remark.

Note: We can obtain the AM of any n real numbers. But, we only define the GM of n **positive** real numbers; and the HM of n **non-zero** real numbers.

Now let us look at the three different means together. To do so, we clearly need to restrict ourselves to positive real numbers. What is the AM of 2, 4 and 8? How is it related to their GM? And, how is their GM related to their HM? The following result answers these questions.

Theorem 1: Let $\{x_1, x_2, \ldots, x_n\}$ be any finite set of positive real numbers, and let A, G and H denote their arithmetic, geometric and harmonic means, respectively. Then

This proof is due to Cauchy, who you will meet again in Sec. 6.3.

$A \geq G \geq H$,

and $A=G=H$ iff $x_1 = x_2 = \ldots = x_n$.

We will only give a broad outline of the proof here. The inequality $A \geq G$ is first proved by induction (see Unit 2) for all those integers n that are powers of two. That is,

$$\frac{x_1 + x_2 + \ldots + x_{2^m}}{2^m} \geq (x_1 x_2 \ldots x_{2^m})^{2^{-m}}, m \in N \qquad \ldots\ldots\ldots(1)$$

and equality holds iff $x_1 = x_2 = \ldots = x_{2^m}$.

Now, given any $n \in N$, we can always choose $r \in N$ such that $2^r > n$.

We apply (1) to the 2^r numbers $x_1, x_2, \ldots, x_n, A, \ldots, A$, where the number of As is $2^r - n$. We get

$$\frac{x_1 + x_2 + \ldots + x_n + A + A + \ldots + A}{2^r} \geq (x_1 x_2 \ldots x_n A \ldots A)^{2^{-r}}$$

(with equality iff $x_1 = x_2 = \ldots = x_n = A$.)

$$\Rightarrow \frac{nA + (2^r - n)A}{2^r} \geq (G^n A^{(2^r - n)})^{2^{-r}}, \text{ since} \sum_{i=1}^{n} x_i = nA.$$

$$\Rightarrow A^{2^r} \geq G^n A^{2^r - n}$$

$$\Rightarrow A^n \geq G^n$$

$\Rightarrow A \geq G$, since A and G are positive real numbers.

Note that $A=G$ iff $x_1 = x_2 = \ldots = x_n$.

Thus, the result is true $\forall\, n \in N$.

Now let us consider the n positive numbers $\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}$.

Since their AM is greater than or equal to their GM, we get

$$\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}\right) \geq \left(\frac{1}{x_1} \cdot \frac{1}{x_2} \cdots \frac{1}{x_n}\right)^{\frac{1}{n}}$$

$$\Rightarrow \frac{1}{H} \geq \frac{1}{G}$$

$\Rightarrow G \geq H$.

Note that $H=G$ iff $\frac{1}{x_1} = \frac{1}{x_2} = \ldots = \frac{1}{x_n}$, that is, $x_1 = x_2 = \ldots = x_n$.

Thus, $A \geq G \geq H$, with equality iff $x_1 = x_2 = ... = x_n$.

In about 320 A.D. the geometer Pappus of Alexandria gave a geometric construction of the AM, GM and HM of two numbers. His construction is as follows:

Draw a semicircle with a+b as diameter (see Fig.1). Let its diameter be AC, with mid-point O.

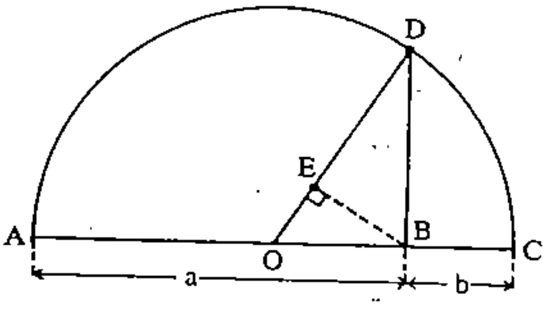


Fig. 1 : The AM, GM and HM of a and b are DO, DB and DE, respectively.

Then OA is the radius of the circle. Mark off the point B on AC such that AB=a. Then BC=b. Draw BD $\perp$ AC to meet the semicircle in D. Then draw BE $\perp$ DO, as in Fig.1. Then Pappus proved that

DO is the AM of a and b,

DB is the GM of a and b,

DE is the HM of a and b.

Since DO $\geq$ DB $\geq$ DE, this gives us a geometric proof for Theorem 1, when n=2.

Now let us apply Theorem 1 to prove some more inequalities.

Example 1: Show that $\left(\sum_{i=1}^{n} i^r\right)^n > n^n(n!)^r$, where n! denotes factorial n and $r > 0$.

Solution: Let r be a fixed positive real number. Consider the n positive numbers 1^r, 2^r,........n^r. By Theorem 1

$$\frac{1^r + 2^r + ... + n^r}{n} \geq (1^r \cdot 2^r \ldots . n^r)^{1/n} = (n\,!)^r)^{1/n}.$$

Since the numbers $1^r, 2^r, ..., n^r$ are not equal, their AM is strictly greater than their GM. Thus

$$\left(\frac{1^r + 2^r + ... + n^r}{n}\right)^n > (n!)^r$$

$$\Rightarrow \left(\sum_{i=1}^{n} i^r\right)^n > n^n(n!)^r$$

We can prove several inequalities, which are particularly useful in mathematics, by using Theorem 1. We ask you to prove some of them in the following exercises.

E 1) Show that $(ab+xy)(ax+by) \geq 4abxy$, where a,b,x,y are positive real numbers. Under what conditions on a,b,x and y would the equality hold?

E 2) For any $n \in N$ and positive real numbers x and y, show that

a) $(xy^n)^{1/(n+1)} \leq \frac{x+ny}{n+1}$,

b) $\left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n+1}\right)^{n+1}$,

c) $\left(1+\frac{1}{m}\right)^m < \left(1+\frac{1}{n+1}\right)^n$, where $m \in \mathbf{N}$ such that $m < n$.

E 3) Is Theorem 1 true if we remove the condition that the numbers are positive? Why?

Now, you know that the inequalities in Theorem 1 become equalities when $x_1 = x_2 = \ldots = x_n$. When this happens, then $x_i = A = G = H \; \forall \; i = 1, \ldots, n$. Thus, we see that

> if $x_1, x_2, \ldots, x_n$ are n positive real numbers such that $x_1 + x_2 + \ldots + x_n$ is a constant, then their arithmetic mean attains its lowest value and their geometric mean attains its maximum value when $x_1 = x_2 = \ldots = x_n = A = G$

Let us see how to use this fact for obtaining some maximum and minimum values. For convenience, we shall denote the set of positive real numbers by $\mathbf{R}^+$.

Example 2: Find the greatest value of xyz, where $x, y, z \in \mathbf{R}^+$ are subject to the condition $yz+zx+xy=12$.

Solution: xyz has greatest value when $(xyz)^2 = (yz)(zx)(xy)$ has greatest value. Since $yz+zx+xy$ is a constant, we know that the maximum value of $(yz)(zx)(xy)$ is attained when $yz = zx = xy$, that is, when $x = y = z$.

Then, $yz + zx + xy = 12 \Rightarrow x = y = z = 2$.

Hence, the maximum value of xyz is $2^3 = 8$.

Example 3: If the sum of the sides of a triangle is the constant 2s, prove that the area is greatest when the triangle is equilateral.

Solution: Let a, b, c be the sides of the triangle, where

$a+b+c = 2s$,

and let Δ denote the area of the triangle.

Then $\Delta = \sqrt{s(s-a)(s-b)(s-c)}$.

So, Δ will be greatest when $(s-a)(s-b)(s-c)$ is maximum.

Now $(s-a) + (s-b) + (s-c) = s$, a constant.

Thus, $(s-a)(s-b)(s-c)$ is maximum when

$s-a = s-b = s-c$, that is, $a = b = c$.

Thus, the area is maximum when the triangle is an equilateral triangle.

Why don't you try these exercises now?

E4) a) Prove that if the sum of two positive numbers is given, their product is greatest when they are equal.

b) Is (a) true if the words 'sum' and 'product' are interchanged? Why?

E5) Find the greatest value of $(5+x)^3 (5-x)^4$, for $-5 < x < 5$.
(Hint : The greatest value of $(5+x)^3 (5-x)^4$ occurs when the greatest value of $\left(\frac{5+x}{3}\right)^3 \left(\frac{5-x}{4}\right)^4$ occurs.)

E6) When does a cuboid, with dimensions x,y and z such that x+y+z is fixed, have maximum volume?

E7) Under what conditions on the dimensions, will a cuboid with fixed volume have minimal surface area?
(Hint: Use the inequality $G \geq H$.)

You can study other techniques for obtaining maximum values in our course on calculus.

Let us now consider another inequality, which follows from Theorem 1

Theorem 2: If $x_1, \ldots\ldots, x_n \in \mathbf{R}^+$ such that not all of them are equal, and $m \in \mathbf{Q}$, $m \neq 0$, $m \neq 1$, then

$$\frac{x_1^m + x_2^m + \ldots\ldots\ldots + x_n^m}{n} \leq \left(\frac{x_1 + x_2 + \ldots\ldots\ldots + x_n}{n}\right)^m, \text{ if } 0 < m < 1, \text{ and}$$

$$\frac{x_1^m + x_2^m + \ldots\ldots\ldots + x_n^m}{n} \geq \left(\frac{x_1 + x_2 + \ldots\ldots\ldots + x_n}{n}\right)^m, \text{ if } m < 0 \text{ or } m > 1.$$

The proof of this result uses Theorem 1. We shall not give it here.

A result that follows from Theorem 2 (and Theorem 1) is that

> if $x_1 + x_2 + \ldots\ldots + x_n = c$, a constant, then
>
> for $0 < m < 1$ the maximum value of $\sum_{i=1}^{n} x_i^m$ is $n^{1-m} c^m$, and
>
> for $m < 0$ or $m > 1$ the minimum value of $\sum_{i=1}^{n} x_i^m$ is $n^{1-m} c^m$.
>
> These values are attained when $x_1 = x_2 = \ldots\ldots = x_n$

Again, we shall not prove this result in this course. But let us consider an example of its use for finding some maximum and minimum values.

Example 4: Find the least value of $\frac{1}{x} + \frac{1}{y} + \frac{1}{z}$, where $x, y, z \in \mathbf{R}^+$ and $x+y+z = 27$.

Solution: $\frac{1}{x} + \frac{1}{y} + \frac{1}{z}$ is of the form $x^m+y^m+z^m$, where $m = -1 < 0$. Since $x+y+z = 27$, the least value is obtained when $x=y=z$.

And then $x+y+z=27$ gives us $x=y=z = 9$.

Thus, the minimum value of $\frac{1}{x} + \frac{1}{y} + \frac{1}{z}$ is $\frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$.

Note that we could have also obtained the answer by applying the inequality $G \geq H$ (of Theorem 1), exactly on the lines of the solution of E7.

Now for some exercises.

E8) Show that the sum of the mth powers of the first n even numbers is greater than $n(n+1)^m$, if $m > 1$.

E9) Show that $\sqrt{1} + \sqrt{2} + \ldots\ldots\ldots + \sqrt{n} < n\sqrt{\frac{n+1}{2}}$, where $n \in \mathbf{N}$.

E10) Let $a_1, a_2, \ldots\ldots\ldots, a_n \in \mathbf{R}^+$ and $p, q \in \mathbf{N}$ such that $p > q$. Show that
$a_1^q + a_2^q + \ldots\ldots + a_n^q < n^{p-q}(a_1^p + a_2^p + \ldots\ldots + a_n^p)$

(Hint: Put $m = \frac{-q}{p}$, $x_i = a_i^p$ in Theorem 2.)

So far we have discussed various inequalities related to the arithmetic, geometric and harmonic means. Now let us consider an inequality that had its origin in ancient Greek geometry.

6.2.2 Triangle Inequality

If you look up any translation of the ancient Greek mathematician Euclid's "Elements" you will find that Proposition 20 of Book 1 says:

"In any triangle, two sides taken together in any manner are greater than the remaining one."

This result is the basis of the triangle inequality, which is a statement about the absolute value of numbers.

Recall that the absolute value of $x \in \mathbf{R}$ is defined by

$|x| = x$, if $x \geq 0$

$= -x$, if $x < 0$.

Thus it satisfies the following properties

i) $|x| = |-x| \; \forall \; x \in \mathbf{R}$ and

ii) $x \leq |x| \; \forall \; x \in \mathbf{R}$.

You can study the absolute value of real numbers in more detail in our course on calculus.

Now let us state the triangle inequality.

Theorem 3: Let $x_1, x_2, \ldots\ldots, x_n \in \mathbf{R}$. Then

$|x_1 + x_2 + \ldots\ldots + x_n| \leq |x_1| + |x_2| + \ldots\ldots + |x_n|$.

Moreover, equality holds only when all the non-zero x_i s have the same sign.

Proof: Let us prove the result for n=2 first.

Now

$$\left(|x_1 + x_2|\right)^2 = (x_1 + x_2)^2, \text{ since } |x|^2 = x^2 \; \forall \; x \in \mathbf{R}$$
$$= x_1^2 + 2x_1 x_2 + x_2^2$$
$$\leq |x_1|^2 + 2|x_1||x_2| + |x_2|^2, \text{ since } x \leq |x| \; \forall \; x \in \mathbf{R}.$$
$$= (|x_1| + |x_2|)^2.$$

Now we take the square root on both sides, keeping in mind that $|x| \geq 0 \; \forall \; x \in \mathbf{R}$. We get

$|x_1 + x_2| \leq |x_1| + |x_2|$, which is what we wanted to prove.

Note that if $x_1 < 0$, say $x_1 = -a$, and $x_2 > 0$ say $x_2 = b$, where $a, b > 0$, then $|x_1 + x_2| = |b - a|$, while $|x_1| + |x_2| = a+b$. Thus, when x_1 and x_2 have opposite signs

$|x_1 + x_2| < |x_1| + |x_2|$.

So, Theorem 3 is true for n = 2.

Now, let us prove the result for the general case, by induction. So let $n > 2$ and assume that Theorem 3 is true for any $n-1$ numbers. Now consider

$$|x_1 + x_2 + \ldots\ldots + x_n| = |(x_1 + x_2 + \ldots\ldots + x_{n-1}) + x_n|$$
$$\leq |x_1 + x_2 + \ldots\ldots + x_{n-1}| + |x_n|$$
$$\leq |x_1| + |x_2| + \ldots\ldots + |x_{n-1}| + |x_n|, \text{ since the result is true for } n-1 \text{ numbers.}$$

Thus, $|x_1 + x_2 + \ldots\ldots + x_n| \leq |x_1| + |x_2| + \ldots\ldots + |x_n| \; \forall \; n \in \mathbf{N}$.

Further, just as we have shown for the case n = 2, strict inequality holds if all the non-zero x_i s don't have the same sign.

Theorem 3 is not only true for real numbers. In our course on linear algebra we have proved that if $z_1, z_2, \ldots, z_n \in \mathbf{C}$, then $|z_1+z_2+\ldots+z_n| \leq |z_1| + |z_2| + \ldots + |z_n|$,

where $|z|$ is the modulus of z.

Let us verify Theorem 3 for the numbers −2, 1, 5, 0.

Since $|-2+1+5+0| = |4| = 4$, and

$|-2| + |1| + |5| + |0| = 2+1+5+0 = 8$,

we find that strict inequality in Theorem 3 is true in this case. Note that −2 and 1 have opposite signs.

Why don't you try some exercises now?

E11) The absolute value of the AM of n numbers is less than or equal to the AM of their absolute values. True or false? Why?

E12) Prove or disprove that

$|x-y| \leq |x| - |y| \ \forall \ x, y \in \mathbf{R}$.

(To disprove a statement means to show that it is false. See the appendix of Block 1.)

E13) Prove or disprove that

$|x-y| \geq ||x| - |y|| \ \forall \ x, y \in \mathbf{R}$.

(**Hint:** Write $|x| = |(x-y)+y|$, and also use the fact that $|x| = |-x| \ \forall \ x \in \mathbf{R}$.)

Now let us discuss some "newer" inequalities.

6.3 LESS ANCIENT INEQUALITIES

In this section we shall discuss four important inequalities which are due to some mathematical giants of the nineteenth century. We start with an inequality due to three mathematicians.

6.3.1 Cauchy - Schwarz Inequality

Fig. 2 : Cauchy (1789-1857)

Augustin - Louis Cauchy, the famous French mathematician, was responsible for developments in infinite series, function theory, differential equations, determinants, probability and several other areas of mathematics. One of his contributions was a result, which was later generalised by the German mathematician H.A. Schwarz (1843 -1921). We now state this result, which was also proved independently by the Russian mathematician Bunyakovskii.

Theorem 4 : (Cauchy-Schwarz Inequality) : Let $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbf{R}$.

Then

$$(a_1b_1+a_2b_2+\ldots+a_nb_n)^2 \leq (a_1^2 + a_2^2 + \ldots + a_n^2)(b_1^2 + b_2^2 + \ldots + b_n^2),$$

with equality iff $a_i = cb_i \ \forall \ i = 1, \ldots, n$, where c is a fixed real number.

When $a_i = cb_i \ \forall \ i = 1, \ldots, n$, where c is a constant, we say that the n-tuples $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are **proportional**.

Proof: To help you understand the proof we shall prove it for n=3 first. Then you can try and generalise it (see E14).

Now

$(a_1^2+a_2^2+a_3^2)\ (b_1^2+b_2^2+b_3^2) - (a_1 b_1 + a_2 b_2 + a_3 b_3)^2$

$= (a_1^2 b_2^2 + a_2^2 b_1^2 - 2a_1a_2 b_1 b_2) + (a_2^2 b_3^2 + a_3^2 b_2^2 - 2a_2 a_3 b_2 b_3) + (a_3^2 b_1^2 + a_1^2 b_3^2$

$- 2a_3 a_1 b_3 b_1)$

$= (a_1 b_2 - a_2b_1)^2 + (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2$

$\geq 0.$

When will the equality sign hold? Equality holds iff $a_1b_2 - a_2b_1 = 0$, $a_2b_3 - a_3b_2 = 0$ and $a_3b_1 - a_1b_3 = 0$, that is, $a_1=cb_1$, $a_2=cb_2$, $a_3=cb_3$, for a fixed real number c.

Thus, we have proved the result for n=3.

Now, to complete the proof of Theorem 4, why don't you try this exercise?

E14) Prove Theorem 4 for any $n \in \mathbf{N}$.

Let us consider an application of Theorem 4 for locating the roots of a polynomial. Before going further, you may like to keep Unit 3 nearby for easy reference.

Theorem 5: If all the roots of the real polynomial equation

$x^n+a_1 x^{n-1}+a_2 x^{n-2}+.....+a_n = 0$ are real, then they lie between $\frac{-a_1}{n} - \frac{n-1}{n}\sqrt{a_1^2 - \left(\frac{2n}{n-1}\right)a_2}$ and $\frac{-a_1}{n} + \frac{n-1}{n}\sqrt{a_1^2 - \left(\frac{2n}{n-1}\right)a_2}$

Proof: From Theorem 1 of Unit 3, you know that the given equation has n roots. Let x be a root. If $x_1,......,x_{n-1}$ are the other roots, then by Theorem 4 of Unit 3,

$x+x_1+.....+x_{n-1} = -a_1$

$\Rightarrow \left(a_1 + x\right)^2 = \left(x_1 + ...+x_{n-1}\right)^2 \leq (\underbrace{1^2+1^2+........+1^2}_{(n-1)\text{ times}}) \left(x_1^2 + ...+x_{n-1}^2\right)$, by Theorem 4.

Also, by Theorem 4 of Unit 3,

$x^2 + x_1^2 + ...+x_{n-1}^2 = a_1^2 - 2a_2$

$\therefore (a_1 + x)^2 \leq (n-1)(a_1^2 - 2a_2 - x^2)$

$\Rightarrow nx^2+2a_1x-(n-2)a_1^2 + 2a_2(n-1) \leq 0$

$$\Rightarrow \left[x - \left(\frac{-a_1}{n} + \frac{n-1}{n}\sqrt{a_1^2 - \left(\frac{2n}{n-1}\right)a_2}\right)\right]\left[x - \left(\frac{-a_1}{n} - \frac{n-1}{n}\sqrt{a_1^2 - \left(\frac{2n}{n-1}\right)a_2}\right)\right] \leq 0,$$

by the quadratic formula.

This holds for all x such that

$$\frac{-a_1}{n} - \frac{n-1}{n}\sqrt{a_1^2 - \left(\frac{2n}{n-1}\right)a_2} \leq x \leq \frac{-a_1}{n} + \frac{n-1}{n}\sqrt{a_1^2 - \left(\frac{2n}{n-1}\right)a_2}.$$

Thus, any root of the given polynomial equation must lie between the bounds given in the statement of the theorem.

Before giving an example of the use of Theorem 5, we shall make some related observations.

Remark 1: Consider the polynomial equation $a_0+a_1x+a_2x^2+....+a_nx^n = 0$, where $a_i \in \mathbf{Z}$

$\forall$ i=0, ..., n, and $a_n \neq 0$.

Then any rational root of this equation is of the form $\frac{d}{a_n}$, where d is a factor of a_0.

Remark 2: For cubic cases, we know from the discriminant (Sec. 3.3.2) when the roots are all real. And then Theorem 5 can be very useful, especially if we know that the roots are rational.

Let us now consider an example of the use of Theorem 5.

Example 5: Solve $x^3-23x^2+167x-385 = 0$.

Solution: The discriminant of this equation (see Sec. 3.3.2) is positive. Hence the given equation has three distinct real roots. If the roots are rational they must be integral factors of – 385. Thus, they must belong to the set $\{\pm1, \pm5, \pm7, \pm11, \pm35, \pm55, \pm77, \pm385\}$.

But, by Theorem 5 the roots must lie between $\frac{23}{3}-\frac{4}{3}\sqrt{7}$ and $\frac{23}{3}+\frac{4}{3}\sqrt{7}$. Thus, if they are rational, they can only be 5,7, 11. On substituting these values in the equation, we find that they are indeed roots of the given equation. Also, you know that the equation can only have 3 roots. Hence, these values are the only roots.

You may now like to try to apply Theorem 5 yourself.

E15) Solve $x^3-2x^2-x+2 = 0$.

Now let us consider another example of the use of Theorem 4. In this example, we shall apply the Cauchy-Schwarz inequality twice to get an inequality that we want.

Example 6: Let x, y, z $\in$ $\mathbf{R}^+$ such that $x^2+y^2+z^2 = 27$. Show that

$x^3+y^3+z^3 \geq 81$.

Solution: Let us first apply Theorem 4 to the two triples of real numbers, $(x^{3/2}, y^{3/2}, z^{3/2})$ and $(x^{1/2}, y^{1/2}, z^{1/2})$. We get

$(x^{3/2}\,x^{1/2}+y^{3/2}\,y^{1/2}+z^{3/2}\,z^{1/2})^2 \leq (x^3+y^3+z^3)(x+y+z)$, that is

$(x^2+y^2+z^2)^2 \leq (x^3+y^3+z^3)(x+y+z)$ ....(2)

Now let us apply the Cauchy-Schwarz inequality to the triples

(x, y, z) and (1, 1, 1). We get

$(x.1+y.1+z.1)^2 \leq (x^2+y^2+z^2)(1^2+1^2+1^2)$, that is

$(x+y+z)^2 \leq 3\,(x^2+y^2+z^2)$

$=81$

$\Rightarrow x+y+z \leq 9$

Thus, by (2)

$(x^2+y^2+z^2)^2 \leq 9\,(x^3+y^3+z^3)^2$

But $x^2+y^2+z^2 = 27$. Thus,

$(x^3+y^3+z^3) \geq \frac{(27)^2}{9}$, that is,

$x^3+y^3+z^3 \geq 81$.

Why don't you try some exercises now?

E16) If a, b, x, y $\in$ $\mathbf{R}$ such that $a^2+b^2 = 1$ and $x^2+y^2 = 1$, then prove that $ax+by \leq 1$.

E17) Prove that if $a_1,......,a_n \in \mathbf{R}^+$, then

a) $(a_1+a_2+......+a_n)\left[\frac{1}{a_1}+\frac{1}{a_2}+......+\frac{1}{a_n}\right] \geq n^2$.

b) $(a_1+a_2+........+a_n)^2 \leq n(a_1^2+a_2^2+....+a_n^2)$

c) $(\sqrt{a_1}+\sqrt{a_2}+......+\sqrt{a_n})^2 \leq n(a_1+a_2+....+a_n)$

E18) (Another form of the triangle inequality) If a, b, x, y $\in$ **R**, then show that

$\sqrt{(a-b)^2+(x-y)^2} \leq \sqrt{a^2+x^2}+\sqrt{b^2+y^2}$

(**Hint:** Write $(a-b)^2+(x-y)^2 = (a^2+x^2)+(b^2+y^2)-2(ab+xy)$, and then apply Theorem 4 to (a, x) and (b, y).)

E19) If x, y, z $\in$ $\mathbf{R}^+$ such that $x^3+y^3+z^3 = 81$, then prove that $x+y+z \leq 9$.

E20) Prove or disprove the following generalisation of Theorem 4 :

Let $p \in$ **N**, $p \neq 1$, and $a_1, ..., a_n, b_1, b_2, ..., b_n \in$ **R**.

Then $(a_1b_1 + a_2b_2 + ... + a_nb_n)^p \leq (a_1^p+a_2^p+.........+a_n^p)(b_1^p+b_2^p+....+b_n^p)$

The Cauchy—Schwarz inequality has several applications in physics and mathematics, especially in the context of inner product spaces.

Now let us consider another useful set of inequalities.

6.3.2 Weierstrass' Inequalities

It is generally thought that a good mathematician must have started serious mathematical studies at an early age. But the German mathematician Weierstrass (1815–1897) is an exception to this rule. This outstanding mathematician started serious mathematics at the age of forty. He was responsible for making analysis more rigorous, and is considered to be the "father of modern analysis". He is responsible for the following result.

Fig 3 : Karl Theodor Weierstrass

Theorem 6 (Weierstrass' Inequalities) : Let $a_1, a_2, ..., a_n$ be positive real numbers less than 1 and $s_n = a_1 + a_2 + ... + a_n$. Then

(i) $1-s_n \leq (1-a_1)(1 \quad ....(1-a_n) < \frac{1}{1+s_n}$,

(ii) $1+s_n \leq (1+a_1)(1+a_2).....(1+a_n) < \frac{1}{1-s_n}$, where it is assumed that $s_n < 1$.

Proof : We prove (i) by induction on n, a principle that we introduced you to in Unit 2.

If n = 1, then $s_1 = a_1$ and hence, $(1-s_1) = (1-a_1)$.

Also, since $0 < a_1^2 < 1$, $(1-a_1)(1+a_1) < 1$, that is, $(1-a_1) < \frac{1}{1+s_1}$.

So, (i) is true when n = 1.

Let us assume that (i) is true for n = m, where m $\in$ **N**.

We will see if it is also true for n = m+1.

Now $s_{m+1} = a_1 + ... + a_{m+1} = (a_1 + ... + a_m) + a_{m+1} = s_m + a_{m+1}$

also $(1-a_1)(1-a_2)......(1-a_m) \geq 1-s_m$, by our assumption.

Thus, $(1-a_1)(1-a_2) .... (1-a_m)(1-a_{m+1}) \geq (1-s_m)(1-a_{m+1})$

$= 1-(s_m + a_{m+1}) + s_m a_{m+1}$

$= 1-s_{m+1} + s_m a_{m+1}$

$> 1-s_{m+1}$, since $s_m a_{m+1} > 0$.

So, $(1-a_1).......(1-a_{m+1}) > 1 - s_{m+1}$ .... (3)

Further, since $(1-a_1)(1-a_2).....(1-a_m) < \frac{1}{1+s_m}$, by our assumption, and $(1-a_{m+1}) < \frac{1}{1+a_{m+1}}$, we find that

$$(1-a_1)(1-a_2).....(1-a_{m+1}) < \frac{1}{(1+s_m)(1+a_{m+1})}$$

$$= \frac{1}{1+s_{m+1}+s_m a_{m+1}}$$

$$< \frac{1}{1+s_{m+1}} \quad .... (4)$$

(3) and (4), taken together, tell us that (i) is true for $n = m+1$. Hence, by induction, (i) is true $\forall\, n \in N$.

Now, to complete the proof you can try E21.

E21) Prove (ii) of Theorem 6.

E22) For $0 < a_1, a_2, ..., a_n < 1$, prove that

$$1 - \prod_{i=1}^{n} a_i < n - \sum_{i=1}^{n} a_i.$$

(**Hint** : $0 < 1-a_i < 1$.)

E23) Does Theorem 6 hold if $a_i < 0$ or $a_i > 1$ for any $i = 1, ..., n$? Give reasons for your answer.

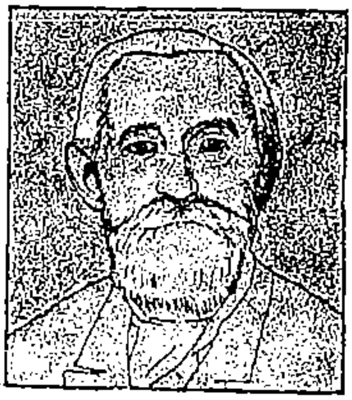


Fig. 4 : Tchebychev (1821-1894)

When you study mathematical analysis, you will find that Weierstrass' inequalities and their generalisations are very useful.

And finally, we shall discuss some inequalities due to a leading Russian mathematician.

6.3.3 Tchebychev's Inequalities

The mathematician Pafnutii L.Tchebychev (pronounced Che-bee-cheff) is most known for his tremendous work in analytic number theory and the theory of orthogonal polynomials. Over here we shall prove and apply some inequalities that are named after him.

Theorem 7 (Tchebychev's Inequalities) : If $a_1, ..., a_n, b_1, ..., b_n \in \mathbf{R}$ such that

i) $a_1 \le a_2 \le ... \le a_n$, $b_1 \le b_2 \le .... \le b_n$, then

$$n(a_1b_1 + a_2b_2 + ... + a_nb_n) \ge (a_1 + a_2 + ... + a_n)(b_1 + b_2 + ... + b_n).$$

ii) $a_1 \ge a_2 \ge ..... \ge a_n$, $b_1 \le b_2 \le ..... \le b_n$, then

$$n(a_1b_1 + a_2b_2 + ... + a_nb_n) \le (a_1 + a_2 + ... + a_n)(b_1 + b_2 + ... + b_n).$$

Proof : Let us prove (i) for the case $n = 3$, so that you can understand the proof more easily.

Since $a_1 \le a_2$ and $b_1 \le b_2$, we find that

$(a_1-a_2)(b_1-b_2) \ge 0$. Therefore,

$$a_1b_1 + a_2b_2 \ge a_1b_2 + a_2b_1 \quad .... (5)$$

Similarly, we get

$$a_2b_2 + a_3b_3 \ge a_2b_3 + a_3b_2 \quad .... (6)$$

$a_3b_3 + a_1b_1 \geq a_3b_1 + a_1b_3$ .... (7)

Adding (5), (6) and (7), we get

$2(a_1b_1 + a_2b_2 + a_3b_3) \geq a_1(b_2 + b_3) + a_2(b_3 + b_1) + a_3(b_1+b_2)$

Now we add $a_1b_1 + a_2b_2 + a_3b_3$ to both sides and simplify to get

$3(a_1b_1 + a_2b_2 + a_3b_3) \geq (a_1+a_2+a_3)(b_1 + b_2 + b_3)$.

The proof of (i) for any $n \in \mathbf{N}$ is on exactly the same lines. Let us prove it by induction on n.

The result is true for n = 3 (and, in fact, for n = 1 and 2).

Assume that it is true for n–1. Then

$(n-1)(a_1b_1 + ... + a_{n-1}b_{n-1}) \geq (a_1 + a_2 + ... + a_{n-1})(b_1 + b_2 + ... + b_{n-1})$

Also $a_1b_1 + a_nb_n \geq a_1b_n + a_nb_1$, since $(a_1-a_n)(b_1-b_n) \geq 0$.

Similarly $a_2b_2 + a_nb_n \geq a_2b_n + a_nb_2$

. . . .

. . . .

. . . .

$a_{n-1}b_{n-1} + a_nb_n \geq a_{n-1}b_n + a_nb_{n-1}$

Adding up the left hand sides and the right hand sides of these n inequalities, we get

$n(a_1b_1 + ... + a_nb_n) - a_nb_n \geq (a_1 + ... a_{n-1} + a_n)(b_1 + ... b_{n-1} + b_n) - a_nb_n$.

$\Rightarrow n(a_1b_1 + ... + a_nb_n) \geq (a_1 + ... + a_n)(b_1 + ... + b_n)$.

And now, to complete the proof of the theorem, try **E24**.

E24) Prove (ii) of Theorem 7.

(Hint : Put $x_i = -a_i \; \forall \; i = 1, ..., n$, and use (i).)

Now let us consider an application of Tchebychev's inequalities.

Example 7 : Show that

$$\sqrt{1} + \sqrt{2} + ... + \sqrt{n} \leq n\sqrt{\frac{n+1}{2}}.$$

Solution : Put $a_i = b_i = \sqrt{i} \; \forall \; i = 1,....,n$ in Theorem 7. Then we get

$n(\sqrt{1}\sqrt{1} + \sqrt{2}\sqrt{2} + ... + \sqrt{n}\sqrt{n}) \geq (\sqrt{1} + \sqrt{2} + ... + \sqrt{n})^2$, that is,

$n(1 + 2 + ... + n) \geq (\sqrt{1} + \sqrt{2} + ... + \sqrt{n})^2$, that is

$$n\left[\frac{n(n+1)}{2}\right] \geq (\sqrt{1} + \sqrt{2} + ... + \sqrt{n})^2, \text{ since } \sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Taking the square root on both sides, we get

$$\sqrt{1} + \sqrt{2} + ... + \sqrt{n} \leq n\sqrt{\frac{n+1}{2}}.$$

Now you can try some exercises.

E25) Show that

$$\frac{1}{\sqrt{n}}\left(\sqrt{1} + \sqrt{\frac{1}{2}} + ... + \sqrt{\frac{1}{n}}\right) \leq (2n-1)^{1/4}.$$

(Hint : First apply Tchebychev's inequality to the n-tuples

$\left(1, \frac{1}{2}, \ldots, \frac{1}{n}\right)$ and $\left(1, \frac{1}{2}, \ldots, \frac{1}{n}\right)$; and then apply it again to the n-tuples

$$\left(\sqrt{\frac{1}{1}}, \sqrt{\frac{1}{2}}, \ldots, \sqrt{\frac{1}{n}}\right) \text{ and } \left(\sqrt{\frac{1}{1}}, \sqrt{\frac{1}{2}}, \ldots, \sqrt{\frac{1}{n}}\right).$$

E26) If $a,b,c \in \mathbf{R}^+$, then show that

$$\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b} \geq \frac{3}{2}.$$

(**Hint** : See if it is possible to apply Theorem 7 to b+c, c+a, a+b and their inverses.)

With this inequality we come to the end of this unit. This doesn't mean that we've exhausted all the inequalities, or even all the important ones. We have just exposed you to a few elementary ones and some of their applications. As you study more mathematics you will come across these and several others.

Now let us quickly go through what we have covered in this unit.

6.4 SUMMARY

We have discussed several inequalities and their applications in this unit. Let us list them one by one.

1) The inequality of the means: The AM of any finite set of elements of $\mathbf{R}^+$ is greater than or equal to their GM, which is greater than or equal to their HM.

2) If $x_1, \ldots, x_n \in \mathbf{R}$ such that not all of them are equal, and $m \in \mathbf{Q}$, $m \neq 0, 1$, then

$$\frac{1}{n}\left(\sum_{i=1}^{n} x_i^m\right) \leq \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^m \text{ if } 0 < m < 1, \text{ and}$$

$$\frac{1}{n}\left(\sum_{i=1}^{n} x_i^m\right) \geq \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^m \text{ if } m < 0 \text{ or } m > 1.$$

3) The triangle inequality : For $x_1, x_2, \ldots, x_n \in \mathbf{R}$

$|x_1+x_2+\ldots+x_n| \leq |x_1|+|x_2|+\ldots+|x_n|.$

The inequality is strict in case all the non-zero x_is don't have the same sign.

4) Cauchy-Schwarz (or Bunyakovskii) inequality: If $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbf{R}$, then

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

with equality iff $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are proportional.

5) Weierstrass' inequalities: For $0 < a_1, \ldots, a_n < 1$ and $s_n = \sum_{i=1}^{n} a_i$,

$$1 - s_n \leq \prod_{i=1}^{n} (1 - a_i) < \frac{1}{1+s_n},$$

$$1 + s_n \leq \prod_{i=1}^{n} (1 + a_i) < \frac{1}{1-s_n} \text{ (here } s_n < 1 \text{ is assumed).}$$

6) Tchebychev's inequalities: If $(a_1,\ldots,a_n),(b_1,\ldots,b_n) \in \mathbf{R}$ such that

i) $a_1 \le a_2 \le \ldots \le a_n$, $b_1 \le b_2 \le \ldots \le b_n$, then

$n(a_1b_1+a_2b_2+\ldots+a_nb_n) \ge (a_1+a_2+\ldots+a_n)(b_1+b_2+\ldots+b_n)$;

ii) $a_1 \le a_2 \le \ldots \le a_n$, $b_1 \ge b_2 \ge \ldots \ge b_n$, then

$n(a_1b_1+a_2b_2+\ldots+a_nb_n) \le (a_1+a_2+\ldots+a_n)(b_1+b_2+\ldots+b_n)$.

As usual, we suggest that you go back to the beginning of the unit and see if you have achieved the objectives. We have given our solutions to the exercises in the unit in the last section. Please go through them too.

With this we have come to the end of this course. We hope you have enjoyed it, and will find it of use in the future.

6.5 SOLUTIONS/ANSWERS

E1) Applying the inequality $A \ge G$ to the positive real numbers ab and xy, we get

$ab+xy \ge 2\sqrt{abxy}$ .... (8)

Now we apply $A \ge G$ to ax and by, and we get

$ax+by \ge 2\sqrt{abxy}$ .... (9)

(8) and (9) $\Rightarrow (ab+xy)(ax+by) \ge 4\sqrt{abxy}\sqrt{abxy} = 4\,abxy$.

Note that equality holds iff ab=xy and ax=by $\Leftrightarrow$ a=y and b=x.

E2) a) Applying the inequality $A \ge G$ to the n+1 numbers

x,y,y,...,y (n times), we get

$$\frac{x+ny}{n+1} \ge \left(x\,.\,y^n\right)^{\frac{1}{n+1}}$$

Note that equality holds iff x=y.

b) Put x=1 and $y=1+\frac{1}{n}$ in (a). Then we get

$$\left[1.\left(1+\frac{1}{n}\right)^n\right]^{\frac{1}{n+1}} < \frac{1+n\left(1+\frac{1}{n}\right)}{n+1}$$

$$\Leftrightarrow \left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n}\right)^{n+1}$$

c) Let $m, n \in \mathbf{N}$ and $m < n$. Then, by (b),

$$\left(1+\frac{1}{m}\right)^m < \left(1+\frac{1}{m+1}\right)^{m+1} < \left(1+\frac{1}{m+2}\right)^{m+2} < \ldots\ldots < \left(1+\frac{1}{n}\right)^n$$

Hence the result.

E3) No, since the GM of negative numbers is not defined. Even $A \ge H$ does not remain true any longer.
For example, take the three numbers −2, 1 and 3.

Their AM is $\frac{2}{3}$ and their HM is $\frac{3}{\frac{-1}{2}+1+\frac{1}{3}} = \frac{18}{5}$, and $\frac{2}{3} \ngeq \frac{18}{5}$.

E4) a) Let $x,y \in \mathbf{R}^+$ such that $x+y = c$, a constant. Then, xy is maximum when $\sqrt{xy}$ is maximum, that is, when x=y.

b) No. For example, let $x,y \in \mathbf{R}^+$ such that $xy=1$.
If $x=y$, then $x=1=y$. And then $x+y=2$. But this is not the maximum value of $x+y$, since $x=5$ and $y=\frac{1}{5}$ for example, give a larger value of $x+y$.

E5) Since $-5 < x < 5$, $5+x > 0$ and $5-x > 0$.

Since $3\left(\frac{5+x}{3}\right)+4\left(\frac{5-x}{4}\right)=10$, a constant, the maximum value of $\left(\frac{5+x}{3}\right)^3\left(\frac{5-x}{4}\right)^4$ occurs when $\frac{5+x}{3}=\frac{5-x}{4}$, that is, when $x=\frac{-5}{7}$.

Thus, $(5+x)^3\ (5-x)^4$ is maximum when $x=\frac{-5}{7}$.

E6) Let $x+y+z=c$.

The volume of the cuboid is xyz.

This is maximum when $x=y=z$, since $x+y+z=c$.

Thus, the cube is a cuboid with maximum volume, under the given conditions.

E7) Let x,y and z be dimensions of the cuboid, where $xyz=c$, a constant.

Now, the surface area of the cuboid is $2(xy+yz+zx)$. This is minimum when $xy+yz+zx$ is minimum, that is, when

$xyz\left(\frac{1}{z}+\frac{1}{x}+\frac{1}{y}\right)$ is minimum.

Now, $xyz=c$, and the HM of x,y and z is $\dfrac{3}{\frac{1}{x}+\frac{1}{y}+\frac{1}{z}}$.

By Theorem 1 we know that the HM is maximum when $x=y=z$. Thus $\frac{1}{x}+\frac{1}{y}+\frac{1}{z}$ is minimum when $x=y=z$. Thus the surface area of a cuboid with fixed volume is minimum when the cuboid is a cube.

E8) We apply Theorem 2 to the n numbers $2,4,6,\ldots,2n$.

We get

$$\frac{2^m+4^m+6^m+\ldots+(2n)^m}{n} > \left(\frac{2+4+6+\ldots+2n}{n}\right)^m = 2^m\left(\frac{1+2+\ldots+n}{n}\right)^m$$

$$\Leftrightarrow 2^m+4^m+\ldots+(2n)^m > n\,.\,2^m.\left(\frac{n(n+1)}{2n}\right)^m, \text{ since } \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

$$= n(n+1)^m.$$

E9) Applying Theorem 2 to $1,2,\ldots,n$, we get

$$\frac{\sqrt{1}+\sqrt{2}+\ldots+\sqrt{n}}{n} < \sqrt{\frac{n(n+1)}{2n}}$$

Hence the result.

E10) Put $m=\frac{q}{p}$. Then $0 < m < 1$. Now we apply Theorem 2 to the n positive numbers $a_1^p, a_2^p, \ldots, a_n^p$. We get

$$\frac{\left(a_1^p\right)^m+\left(a_2^p\right)^m+\ldots+\left(a_n^p\right)^m}{n} \le \left(\frac{a_1^p+a_2^p+\ldots+a_n^p}{n}\right)^m$$

$$\Leftrightarrow \frac{a_1^q + a_2^q + \dots + a_n^q}{n} \le \left(\frac{a_1^p + a_2^p + \dots + a_n^p}{n}\right)^{q/p}$$

$$\Leftrightarrow a_1^q + a_2^q + \dots + a_n^q \le \left(n^{p-q}\right)^{1/q} \left(a_1^p + a_2^p + \dots + a_n^p\right)^{q/p}$$

$$< n^{p-q} \left(a_1^p + a_2^p + \dots + a_n^p\right) \text{ since } \frac{1}{p} < 1 \text{ and } \frac{q}{p} < 1.$$

E11) Let $x_1, \dots, x_n$ be n numbers.

Then their AM, $A = \frac{x_1 + x_2 + \dots + x_n}{n}$.

$\therefore A = \left|\frac{x_1 + x_2 + \dots + x_n}{n}\right| \le \frac{|x_1| + |x_2| + \dots + |x_n|}{n}$, by theorem 3.

$= \text{AM of } |x_1|, |x_2|, \dots, |x_n|$.

Hence the statement is true.

E12) False. For example, take $x=1$, $y=-3$.

Then $|x - y| = |4| = 4$,

and $|x| - |y| = 1 - 3 = -2$.

E13) $|x| = |(x-y) + y| \le |x - y| + |y|$. ...... (10)

$\therefore |x| - |y| \le |x - y|$

Similarly, $|y| = |x + (y - x)| \le |x| + |y - x| = |x| + |x - y|$,

since $|x| = |-x|$.

Therefore, $|y| - |x| \le |x - y|$ ....(11)

(10) and (11) $\Rightarrow ||x| - |y|| \le |x - y|$

E14) $\left(a_1^2 + \dots + a_n^2\right)\left(b_1^2 + \dots + b_n^2\right) - \left(a_1 b_1 + a_2 b_2 + \dots + a_n b_n\right)^2$

$$= \sum_{\substack{i, j = 1 \\ i \ne j}}^{n} \left(a_i b_j - a_j b_i\right)^2 \ge 0,$$

with equality iff $a_i b_j = a_j b_i \ \forall\ i, j = 1, \dots, n, i \ne j$,

that is, iff $\frac{a_i}{b_i} = \frac{a_j}{b_j} \ \forall\ i \ne j$,

that is, iff $a_i = cb_i \ \forall\ i = 1, \dots, n$, where c is a constant.

E15) ~~The discriminant of the equation is positive. Thus, the given equation has distinct~~ real roots. They must lie between $\frac{2}{3}(1 - \sqrt{7})$ and $\frac{2}{3}(1 + \sqrt{7})$. If they are rational, they have to be factors of 2. Hence, they can be 1, –1, 2 or –2. Of these, ± 1 and 2 lie within the given bounds. On substitution we find that they actually are the roots. Since the equation has only 3 roots, –1, 1 and 2 are its roots.

E16) Just applying Theorem 4 to the numbers a,b,x,y, we get

$(ax+by)^2 \le (a^2+b^2)(x^2+y^2) = 1$.

$\therefore ax+by \le 1$.

E17) (a) We apply the Cauchy–Schwarz inequality to the n-tuples

$(\sqrt{a_1}, \sqrt{a_2}, \ldots\ldots, \sqrt{a_n})$ and $\left(\frac{1}{\sqrt{a_1}}, \frac{1}{\sqrt{a_2}}, \ldots, \frac{1}{\sqrt{a_n}}\right)$ we get

$$\left(\sqrt{a_1} \cdot \frac{1}{\sqrt{a_1}} + \ldots\ldots + \sqrt{a_n} \cdot \frac{1}{\sqrt{a_n}}\right)^2 \leq \left(a_1 + a_2 + \ldots\ldots\ldots + a_n\right)\left(\frac{1}{a_1} + \frac{1}{a_2} + \ldots\ldots + \frac{1}{a_n}\right)$$

$$\Leftrightarrow n^2 \leq \left(a_1 + a_2 + \ldots\ldots + a_n\right)\left(\frac{1}{a_1} + \frac{1}{a_2} + \ldots\ldots\ldots\ldots + \frac{1}{a_n}\right).$$

b) Applying Theorem 4 to the n-tuples $(1,1,\ldots\ldots,1)$ and $(a_1, a_2, \ldots\ldots, a_n)$, we get the result.

c) Applying Theorem 4 to the n-tuples $(1,1,\ldots,1)$ and $(\sqrt{a_1}, \sqrt{a_2}, \ldots\ldots\ldots, \sqrt{a_n})$ we get the result.

E18) $(a-b)^2+(x-y)^2=(a^2+x^2)+(b^2+y^2)-2\,(ab+xy)$.

$$\leq (a^2+x^2) + (b^2+y^2) + 2\,|(ab+xy)| \qquad \ldots\ldots\ldots(12)$$

Also, by Theorem 4

$(ab+xy)^2 \leq (a^2+x^2)\,(b^2+y^2)$

$$\Rightarrow |ab+xy| \leq \sqrt{a^2+x^2}\,\sqrt{b^2+y^2} \qquad \ldots\ldots\ldots(13)$$

(12) and (13) $\Rightarrow (a-b)^2 + (x-y)^2 \leq \left[\sqrt{(a^2+x^2)} + \sqrt{(a^2+y^2)}\right]^2$,

which gives us the desired result.

E19) By the Cauchy–Schwarz inequality applied to (x,y,z) and (1,1,1), we have

$$(x+y+z)^2 \leq 3\,(x^2+y^2+z^2). \qquad \ldots\ldots\ldots(14)$$

Again, applying Theorem 4 to $(\sqrt{x}, \sqrt{y}, \sqrt{z})$ and $\left(x^{3/2}, y^{3/2}, z^{3/2}\right)$

we get $(x^2+y^2+z^2)^2 \leq (x+y+z)\,(x^3+y^3+z^3)$

$$= 81\,(x+y+z) \qquad \ldots\ldots\ldots(15)$$

(14) and (15) $\Rightarrow (x+y+z)^4 \leq 9 \quad 81\,(x+y+z)$.

$\Rightarrow (x+y+z)^3 \leq (9)^3$

$\Rightarrow (x+y+z) \leq 9$.

E20) False. For example, take p=3 and the pairs (1,0), (1,–1).

Then $|(1)\,(1)+(0)\,(-1)|^3 \not\leq (1^3+0^3)\,(1^3+(-1)^3.)$

E21) We use the principle of induction on n.

For n=1, $a_1=s_1$, and hence, $1+s_1 \leq 1+a_1$.

Also, $1-a_1^2 < 1$. Therefore, $1+a_1 < \frac{1}{1-a_1}$.

Assume that the result holds for n=m.

Then $(1+a_1)\,(1+a_2)\ldots\ldots\ldots(1+a_m) \geq 1+s_m$

$\therefore (1+a_1)\,(1+a_2)\ldots\ldots(1+a_m)\,(1+a_{m+1}) \geq 1+s_m+a_{m+1}+s_m a_{m+1}$

$> 1+s_{m+1}$

Also, $(1+a_1)\,(1+a_2)\ldots\ldots\ldots(1+a_m) < \frac{1}{\cdot}$.

$\therefore (1+a_1).....(1+a_m)(1+a_{m+1}) < \dfrac{1}{(1-s_m)(1-a_{m+1})}$

$$< \frac{1}{1-s_{m+1}}$$

Thus, (ii) is true for n=m+1, and hence $\forall\, n \in \mathbf{N}$.

E22) We apply Theorem 6 (i) to the n numbers $1-a_1, 1-a_2, \ldots, 1-a_n$.

Then

$1-(1-a_1+1-a_2+......+1-a_n) \le a_1 a_2.....a_n$.

$$\Rightarrow 1-n+\sum_{i=1}^{n} a_i \le \prod_{i=1}^{n} a_i.$$

$$\Rightarrow 1-\prod_{i=1}^{n} a_i \le n-\sum_{i=1}^{n} a_i$$

E23) No. For example, let us take $a_1 = -1$ and $a_2 = 2$.

Then $1-(a_1+a_2) \le (1-a_1)(1-a_2)$

$\Rightarrow 0 \le 2(-1) = -2$, which is false.

E24) If we put $x_i = -a_i$, then $x_1 \le x_2 \le ...... \le x_n$. Also $b_1 \le b_2 \le ...... \le b_n$.

So, by Theorem 7 (i),

$n(x_1 b_1 + x_2 b_2 + ... + x_n b_n) \ge (x_1+x_2+.....+x_n)(b_1+b_2+.....+b_n)$

$\Leftrightarrow -n(a_1 b_1 + a_2 b_2 + ... + a_n b_n) \ge -(a_1+a_2+....+a_n)(b_1+b_2+....+b_n)$

$\Leftrightarrow n(a_1 b_1 + ... + a_n b_n) \le (a_1+a_2+......+a_n)(b_1+.......+b_n)$.

E25) Applying Tchebychev's inequality to $\left(1, \frac{1}{2}, ..., \frac{1}{n}\right)$ and $\left(1, \frac{1}{2}, ..., \frac{1}{n}\right)$,

we get

$$\left(1+\frac{1}{2}+...........+\frac{1}{n}\right)^2 \le n\left(\frac{1}{1^2}+\frac{1}{2^2}+.......+\frac{1}{n^2}\right)$$

$$\le n\left(1+\frac{1}{1.2}+.........+\frac{1}{(n-1)n}\right), \text{since } \frac{1}{i^2} \le \frac{1}{(i-1)i} \;\forall\, i.$$

$$= n\left\{1+\left(1-\frac{1}{2}\right)+.....+\left(\frac{1}{n-1}-\frac{1}{n}\right)\right\}$$

$$= n\left(1+1-\frac{1}{n}\right).$$

$$\Rightarrow 1+\frac{1}{2}+...............+\frac{1}{n} \le \sqrt{2n-1} \qquad ....(16)$$

Again, applying Theorem 7 to $\left(\sqrt{\frac{1}{1}}, \sqrt{\frac{1}{2}}, ..., \sqrt{\frac{1}{n}}\right)$ and $\left(\sqrt{\frac{1}{1}}, \sqrt{\frac{1}{2}}, ..., \sqrt{\frac{1}{n}}\right)$

we get

$$\left(\sqrt{\frac{1}{1}}+\sqrt{\frac{1}{2}}+...........+\sqrt{\frac{1}{n}}\right)^2 \le n\left(1+\frac{1}{2}+..........+\frac{1}{n}\right). \qquad ....(17)$$

(16) and (17) $\Rightarrow \frac{1}{\sqrt{n}}\left(\sqrt{\frac{1}{1}}+\sqrt{\frac{1}{2}}+\ldots\ldots\ldots+\sqrt{\frac{1}{n}}\right) \leq (2n-1)^{\frac{1}{4}}$.

E26) Given a,b and c, they can always be ordered. Let us assume that $a \leq b \leq c$.

Then $a+b \leq a+c \leq b+c$.

Therefore, $\frac{1}{a+b} \geq \frac{1}{a+c} \geq \frac{1}{b+c}$.

We apply Theorem 7 (i) to $\frac{1}{a+b}, \frac{1}{a+c}, \frac{1}{b+c}$ and to c,b,a.

We get

$$3\left(\frac{c}{a+b}+\frac{b}{a+c}+\frac{a}{b+c}\right) \geq (a+b+c)\left(\frac{1}{b+c}+\frac{1}{a+c}+\frac{1}{a+b}\right) \quad \ldots\ldots\ldots\ldots \text{(18)}$$

Also, applying Theorem 7 (ii) to b+c, a+c, a+b and their inverses, we get

$$3(1+1+1) \leq (b+c+a+c+a+b), \left(\frac{1}{b+c}+\frac{1}{a+c}+\frac{1}{a+b}\right).$$

$$\Rightarrow 9 \leq 2(a+b+c)\left(\frac{1}{b+c}+\frac{1}{a+c}+\frac{1}{a+b}\right). \quad \ldots\ldots\ldots\ldots \text{(19)}$$

(18) and (19) $\Rightarrow \left(\frac{c}{a+b}+\frac{b}{a+c}+\frac{a}{b+c}\right) \geq \frac{3}{2}$.

MISCELLANEOUS EXERCISES

This section is optional.

As in the previous block, this section contains some extra problems related to the material covered in this block. Doing them may give you a better understanding of simultaneous linear equations and inequalities. As before, our solutions to the questions follow the list of problems.

1) Solve the following linear system, if possible:

$$3x + 6y + 15z + 6t = 42$$

$$3x + 8y + 21z - 2t = -8$$

$$2x + 9y + 25z + 7t = 41$$

2) a) If a+b+c = 0, solve the equation

$$\begin{vmatrix} a-x & c & b \\ c & b-x & a \\ b & a & c-x \end{vmatrix} = 0$$

b) Solve the equation

$$\begin{vmatrix} 15-2x & 11 & 10 \\ 11-3x & 17 & 16 \\ 7-x & 14 & 13 \end{vmatrix} = 0$$

3) Use Cramer's rule to solve the system

$$x - y + z = 0$$

$$2x + 3y - 5z = 7$$

$$3x - 4y - 2z = -1$$

4) A company produces three products, each of which must be processed through three different departments. In Table 1 we give the number of hours that each unit of each product must stay for in each department. We also give the weekly capacity of each department.

Table 1

Department	Product			No. of hours available per week
	P_1	P_2	P_3	
1	6	2	2	80
2	7	4	1	60
3	5	5	3	100

What combination of the three products will use up the weekly labour availability in all departments ?

5) Prove the following identities:

a) $$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (b-c)(c-a)(a-b)$$

b) $$\begin{vmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$$

6) A fish of species S_1 consumes 10 gm. of food F_1 and 5 gm. of food F_2 per day. A fish of species S_2 consumes 6 gm. of F_1 and 4 gm. of F_2 per day. If a given

environment has 2.2 kg of F_1 and 1.3 kg of F_2 available daily, what population sizes of the species S_1 and S_2 will consume exactly all of the available food ?

7) Does the system

$$x + y + z + w = 0$$
$$x + 3y - 2z + w = 0$$
$$2x \quad - 3z + 2w = 0$$

have a non-trivial solution?

8) Obtain a solution of the following system, if it exists.

$$x + 2y + 4z + t = 4$$
$$2x \quad - z - 3t = 4$$
$$x - 2y - z \quad = 0$$
$$3x - y - z - 5t = 5$$

9) Show that the following linear system has a two-parameter solution set.

$$2x_1 + x_3 - x_4 + x_5 = 2$$
$$x_1 + x_3 - x_4 + x_5 = 1$$
$$12x_1 + 2x_2 + 8x_3 + 2x_5 = 12.$$

10) Use Cramer's rule, if possible, for solving the following linear systems :

a) $3x + y = 3$

$5x + 2y = 1$

b) $2x - 3y + z = 1$

$x + y - z = 0$

$x - 2y + z = -1$

11) If the coordinate axes in a plane are rotated through an angle θ, then we can express the old coordinates (x,y) in terms of the new coordinates (x', y') as

$$x = x' \cos\theta - y' \sin\theta$$
$$y = x' \sin\theta + y' \cos\theta$$

Use Cramer's rule to write (x', y') in terms of (x,y).

Solutions

1) We apply the Gaussian elimination process. We get the solution set

$\{(4+z, -1-3z, z, 6) \mid z \in \mathbf{R}\}$.

2) a) Adding the 2nd and 3rd columns to the first column of the determinant doesn't change its value. On doing this and using the fact that a+b+c = 0, we get

$$\begin{vmatrix} -x & c & b \\ -x & b-x & a \\ -x & a & c-x \end{vmatrix} = 0$$

$$\Rightarrow (-x) \begin{vmatrix} 1 & c & b \\ 1 & b-x & a \\ 1 & a & c-x \end{vmatrix} = 0, \text{ by P3}$$

$$\Rightarrow (-x)\begin{vmatrix} 1 & c & b \\ 0 & b-x-c & a-b \\ 0 & a-c & c-x-b \end{vmatrix} = 0, \text{ applying P4 twice.}$$

$\Rightarrow (-x)\,[\,(b-c-x)(c-b-x) - (a-c)(a-b)\,] = 0$, expanding along the first column.

On solving this equation, we find that $x = 0$ or

$(b-c)^2 - x^2 + a^2 - a(b+c) + bc = 0$, that is,

$2x^2 = 3(a^2 + b^2 + c^2)$, using the condition that $a+b+c=0$.
Thus, the solution set is

$$\left(0, \pm\sqrt{\frac{3}{2}(a^2+b^2+c^2)} \right).$$

b) Using the properties P1 to P5 of determinants, we see that

$$\begin{vmatrix} 15-2x & 11 & 10 \\ 11-3x & 17 & 16 \\ 7-x & 14 & 13 \end{vmatrix} = 0$$

$$\Rightarrow \begin{vmatrix} 15-2x & 1 & 10 \\ 11-3x & 1 & 16 \\ 7-x & 1 & 13 \end{vmatrix} = 0$$

$\Rightarrow x = 4.$

3) In matrix notation the system is

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & -5 \\ 3 & -4 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ -1 \end{bmatrix}$$

$$\text{Here } D = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 3 & -5 \\ 3 & -4 & -2 \end{vmatrix} = -32 \neq 0.$$

Thus, we can apply Cramer's rule.

Now

$$D_1 = \begin{vmatrix} 0 & -1 & 1 \\ 7 & 3 & -5 \\ -1 & -4 & -2 \end{vmatrix} = -44,$$

$$D_2 = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 7 & -5 \\ 3 & -1 & -2 \end{vmatrix} = -42,$$

$$D_3 = \begin{vmatrix} 1 & -1 & 0 \\ 2 & 3 & 7 \\ 3 & -4 & -1 \end{vmatrix} = 2.$$

Thus, $x = \frac{D_1}{D} = \frac{11}{8}, y = \frac{D_2}{D} = \frac{21}{16}, z = \frac{D_3}{D} = -\frac{1}{16}.$

4) Let x, y and z be the required quantities of P_1, P_2 and P_3. Then we need to solve the system

$6x + 2y + 2z = 80$

$7x + 4y + z = 60$

$5x + 5y + 3z = 100.$

By Gaussian elimination (or Cramer's rule) we get

$x = 5, y = 0, z = 25.$

So, the ideal combination is 5 units of P_1, 25 units of P_3, and none of P_2, per week.

5) a) $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{vmatrix}$, subtracting the first row from the second and third rows.

$$= (b-a)\,(c-a) \begin{vmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 1 & c+a \end{vmatrix}$$

$$= (b-a)\,(c-a)\,(c+a-b-a)$$

$$= (b-c)\,(c-a)\,(a-b)$$

b) $\begin{vmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{vmatrix}$

$= \begin{vmatrix} -2a & c+a & a+b \\ -2p & r+p & p+q \\ -2x & z+x & x+y \end{vmatrix}$, subtracting the second and third columns from the first column.

$$= (-2) \begin{vmatrix} a & c+a & a+b \\ p & r+p & p+q \\ x & z+x & x+y \end{vmatrix}$$

$= (-2) \begin{vmatrix} a & c & b \\ p & r & q \\ x & z & y \end{vmatrix}$, subtracting the first column from the second and third columns.

$= 2 \begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$, interchanging the second and third columns.

6) Let x and y denote the number of fish of species S_1 and S_2, respectively. We have the information given in Table 2 below.

Table 2

Fish Species	Food consumed per week (in gms)	
	F_1	F_2
S_1	10	5
S2	6	4
Total food available per week	2200	1300

Thus, we need to solve the system

$10x + 6y = 2200$

$5x + 4y = 1300.$

Solving by any of the methods, we get

$x = 100, y = 200.$

Thus, the required sizes are 100 fish of species S_1 and 200 of species S_2.

7) By elimination we find that the system has infinitely many solutions (x, 0, 0, –x), where $x \in \mathbf{R}$. Thus, for any $x \neq 0$, we would get a non-trivial solution.

8) By Gaussian elimination, we reach a stage where we get 0 = 8/7. Hence, the given system is inconsistent.

9) We apply Gaussian elimination. After a few steps we get the following system

$x_1 = 1$

$x_2 + 4x_4 - 3x_5 = 0$

$x_3 - x_4 + x_5 = 0.$

Thus, $x_1 = 1$, $x_2 = -4x_4 + 3x_5$, $x_3 = x_4 - x_5$.

So, if $x_4 = s$ and $x_5 = t$, then our solution set is

$\{ (1, -4s + 3t, s-t, s, t) \mid s, t \in \mathbf{R}\}$.

Thus, we have expressed the solutions in terms of the two parameters, s and t.

10) a) In matrix notation, the system is

$$\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Since $D = \begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix} = 1 \neq 0$, we can apply Cramer's rule.

Here $D_1 = \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 5$, and

$D_2 = \begin{vmatrix} 3 & 3 \\ 5 & 1 \end{vmatrix} = -12.$

Thus, the solution is $x = 5$, $y = -12$.

b) Here the coefficient matrix is

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 1 \end{bmatrix} \text{ and } D = |A| = 1 \neq 0.$$

Thus, we can apply Cramer's rule.

$$D_1 = \begin{vmatrix} 1 & -3 & 1 \\ 0 & 1 & -1 \\ -1 & -2 & 1 \end{vmatrix} = -3,$$

$$D_2 = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 1 \end{vmatrix} = -5,$$

$$D_3 = \begin{vmatrix} 2 & -3 & 1 \\ 1 & 1 & 0 \\ 1 & -2 & -1 \end{vmatrix} = -8.$$

$\therefore x = -3, y = -5, z = -8.$

11) Here we can write the equations as

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Since $D = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = 1 \neq 0,$

we can apply Cramer's rule.

Now, $D_1 = \begin{vmatrix} x & -\sin\theta \\ y & \cos\theta \end{vmatrix} = x\cos\theta + y\sin\theta$, and

$D_2 = \begin{vmatrix} \cos\theta & x \\ \sin\theta & y \end{vmatrix} = y\cos\theta - x\sin\theta$.

Thus $x' = x\cos\theta + y\sin\theta$, $y' = y\cos\theta - x\sin\theta$.

NOTE