MINISTRY OF SCIENCE AND HIGHER EDUCATION

Mathematics for Social Sciences

Prepared by:

1. Dr. Berhanu Bekele
2. Ato Mulugeta Naizghi
3. Dr. Simon Derkee
4. Ato Wondwosen Zemene

MOSHE

September, 2019

Content **Page**

Chapter One
Propositional Logic and Set Theory

In this chapter, we study the basic concepts of propositional logic and some part of set theory. In the first part, we deal about propositional logic, logical connectives, quantifiers and arguments. In the second part, we turn our attention to set theory and discus about description of sets and operations of sets.

Main Objectives of this Chapter

At the end of this chapter, students will be able to:-

- ❖ Know the basic concepts of mathematical logic.
- ❖ Know methods and procedures in combining the validity of statements.
- ❖ Understand the concept of quantifiers.
- ❖ Know basic facts about argument and validity.
- ❖ Understand the concept of set.
- ❖ Apply rules of operations on sets to find the result.
- ❖ Show set operations using Venn diagrams.

1.1. Propositional Logic

Mathematical or symbolic logic is an analytical theory of the art of reasoning whose goal is to systematize and codify principles of valid reasoning. It has emerged from a study of the use of language in argument and persuasion and is based on the identification and examination of those parts of language which are essential for these purposes. It is formal in the sense that it lacks reference to meaning. Thereby it achieves versatility: it may be used to judge the correctness of a chain of reasoning (in particular, a "mathematical proof") solely on the basis of the form (and not the content) of the sequence of statements which make up the chain. There is a variety of symbolic logics. We shall be concerned only with that one which encompasses most of the deductions of the sort encountered in mathematics. Within the context of logic itself, this is "classical" symbolic logic.

Section objectives:

After completing this section, students will be able to:-

- ✓ Identify the difference between proposition and sentence.
- ✓ Describe the five logical connectives.
- ✓ Determine the truth values of propositions using the rules of logical connectives.

- ✓ Construct compound propositions using the five logical connectives.
- ✓ Determine the truth values of compound propositions.
- ✓ Distinguish a given compound proposition is whether tautology or contradiction.

1.1.1. Definition and examples of propositions

Consider the following sentences.

a. 2 is an even number.
b. A triangle has four sides.
c. Emperor Menelik ate chicken soup the night after the battle of Adwa.
d. May God bless you!
e. Give me that book.
f. What is your name?

The first three sentences are declarative sentences. The first one is true and the second one is false. The truth value of the third sentence cannot be ascertained because of lack of historical records but it is, by its very form, either true or false but not both. On the other hand, the last three sentences have not truth value. So they are not declaratives.

Now we begin by examining proposition, the building blocks of every argument. A proposition is a sentence that may be asserted or denied. Proposition in this way are different from questions, commands, and exclamations. Neither questions, which can be asked, nor exclamations, which can be uttered, can possibly be asserted or denied. Only propositions assert that something is (or is not) the case, and therefore only they can be true or false.

> **Definition 1.1:** A proposition (or statement) is a sentence which has a truth value (either True or False but not both).

The above definition does not mean that we must always know what the truth value is. For example, the sentence "The 1000^{th} digit in the decimal expansion of π is 7" is a proposition, but it may be necessary to find this information in a Web site on the Internet to determine whether this statement is true. Indeed, for a sentence to be a proposition (or a statement), it is not a requirement that we be able to determine its truth value.

Remark: Every proposition has a truth value, namely **true** (denoted by ***T***) or **false** (denoted by ***F***).

1.1.2. Logical connectives

In mathematical discourse and elsewhere one constantly encounters declarative sentences which have been formed by modifying a sentence with the word "not" or by connecting sentences with the words "and", "or", "if . . . then (or implies)", and "if and only if". These five words or combinations of words are called propositional connectives.

Note: Letters such as p, q, r, s etc. are usually used to denote actual propositions.

Conjunction

When two propositions are joined with the connective "**and**," the proposition formed is a logical ***conjunction***. "and" is denoted by "$\wedge$". So, the logical conjunction of two propositions, p and q, is written:

$p \wedge q$, read as "p and q," or "p conjunction q".

p and q are called ***the components of the conjunction***. $p \wedge q$ is true if and only if p is true and q is true.

The truth table for conjunction is given as follows:

p	q	$p \wedge q$
T	***T***	***T***
T	***F***	***F***
F	***T***	***F***
F	***F***	***F***

Example 1.1: Consider the following propositions:

p: 3 is an odd number. (True)

q: 27 is a prime number. (False)

r: Addis Ababa is the capital city of Ethiopia. (True)

a. $p \wedge q$: 3 is an odd number and 27 is a prime number. (False)
b. $p \wedge r$: 3 is an odd number and Addis Ababa is the capital city of Ethiopia. (True)

Disjunction

When two propositions are joined with the connective "**or**," the proposition formed is called a logical ***disjunction***. "or" is denoted by "$\vee$". So, the logical disjunction of two propositions, p and q, is written:

$p \vee q$ read as "p or q" or "p disjunction q."

$p \vee q$ is false if and only if both p and q are false.

The truth table for disjunction is given as follows:

p	q	$p \vee q$
T	***T***	***T***
T	***F***	***T***
F	***T***	***T***
F	***F***	***F***

Example 1.2: Consider the following propositions:

p: 3 is an odd number. (True)

q: 27 is a prime number. (False)

s: Nairobi is the capital city of Ethiopia. (False)

a. $p \vee q$: 3 is an odd number or 27 is a prime number. (True)

b. p ∨ r: 27 is a prime number or Nairobi is the capital city of Ethiopia. (False)

Note: The use of "**or**" in propositional logic is rather different from its normal use in the English language. For example, if Solomon says, "I will go to the football match in the afternoon or I will go to the cinema in the afternoon," he means he will do one thing or the other, but not both. Here "or" is used in the exclusive sense. But in propositional logic, "or" is used in the inclusive sense; that is, we allow Solomon the possibility of doing both things without him being inconsistent.

Implication

When two propositions are joined with the connective "**implies**," the proposition formed is called a *logical* ***implication***. "implies" is denoted by "$\Longrightarrow$." So, the logical implication of two propositions, p and q, is written:

$$p \Longrightarrow q \quad \text{read as "}p \text{ implies } q\text{."}$$

The function of the connective "implies" between two propositions is the same as the use of "If … then …" Thus $p \Longrightarrow q$ can be read as "if p, then q."

$p \Longrightarrow q$ is false if and only if p is true and q is false.

This form of a proposition is common in mathematics. The proposition p is called the hypothesis or the antecedent of the conditional proposition $p \Longrightarrow q$ while q is called its conclusion or the consequent.

The following is the truth table for implication.

p	q	$p \Longrightarrow q$
T	*T*	*T*
T	*F*	*F*
F	*T*	*T*
F	*F*	*T*

Examples 1.3: Consider the following propositions:

p: 3 is an odd number. (True)

q: 27 is a prime number. (False)

r: Addis Ababa is the capital city of Ethiopia. (True)

$p \Longrightarrow q$: If 3 is an odd number, then 27 is prime. (False)

$p \Longrightarrow r$: If 3 is an odd number, then Addis Ababa is the capital city of Ethiopia. (True)

We have already mentioned that the implication $p \Longrightarrow q$ can be expressed as both "If p, then q" and "p implies q." There are various ways of expressing the proposition $p \Longrightarrow q$, namely:

If p, then q.
q if p.
p implies q.
p only if q.
p is sufficient for q.
q is necessary for p

Bi-implication

When two propositions are joined with the connective "**bi-implication**," the proposition formed is called a *logical* ***bi-implication*** or a *logical* ***equivalence***. A bi-implication is denoted by "$\Leftrightarrow$". So the logical bi-implication of two propositions, p and q, is written:

$$p \Leftrightarrow q.$$

$p \Leftrightarrow q$ is false if and only if p and q have different truth values.

The truth table for bi-implication is given by:

p	q	$p \Leftrightarrow q$
T	***T***	***T***
T	***F***	***F***
F	***T***	***F***
F	***F***	***T***

Examples 1.4:

a. Let p: 2 is greater than 3. (False)
   q: 5 is greater than 4. (True)

Then

$p \Leftrightarrow q$: 2 is greater than 3 if and only if 5 is greater than 4. (False)

b. Consider the following propositions:
   p: 3 is an odd number. (True)
   q: 2 is a prime number. (True)
   $p \Leftrightarrow q$: 3 is an odd number if and only if 2 is a prime number. (True)

There are various ways of stating the proposition $p \Leftrightarrow q$.

p if and only if q (also written as p iff q),
p implies q and q implies p,
p is necessary and sufficient for q
q is necessary and sufficient for p
p is equivalent to q

Negation

Given any proposition p, we can form the proposition $\neg p$ called the **negation** of p. The truth value of $\neg p$ is F if p is T and T if p is F.

We can describe the relation between p and $\neg p$ as follows.

p	$\neg p$
T	**F**
F	**T**

Example 1.5: Let p: Addis Ababa is the capital city of Ethiopia. (True)
$\neg p$: Addis Ababa is not the capital city of Ethiopia. (False)

Exercises

1. Which of the following sentences are propositions? For those that are, indicate the truth value.
   a. 123 is a prime number.
   b. 0 is an even number.
   c. $x^2 - 4 = 0$.
   d. Multiply $5x + 2$ by 3.
   e. What an impossible question!
2. State the negation of each of the following statements.
   a. $\sqrt{2}$ is a rational number.
   b. 0 is not a negative integer.
   c. 111 is a prime number.
3. Let p: 15 is an odd number.
   q: 21 is a prime number.
   State each of the following in words, and determine the truth value of each.
   a. $p \vee q$.
   b. $p \wedge q$.
   c. $\neg p \vee q$.
   d. $p \wedge \neg q$.
   e. $p \Longrightarrow q$.
   f. $q \Longrightarrow p$.
   a. $\neg p \Longrightarrow \neg q$.
   g. $\neg q \Longrightarrow \neg p$.

4. Complete the following truth table.

p	q	$\neg q$	$p \wedge \neg q$
T	T		
T	F		
F	T		
F	F		

1.1.3. Compound (or complex) propositions

So far, what we have done is simply to define the logical connectives, and express them through algebraic symbols. Now we shall learn how to form propositions involving more than one connective, and how to determine the truth values of such propositions.

Definition 1.2: The proposition formed by joining two or more proposition by connective(s) is called a compound statement.

Note: We must be careful to insert the brackets in proper places, just as we do in arithmetic. For example, the expression $p \Rightarrow q \wedge r$ will be meaningless unless we know which connective should apply first. It could mean $(p \Rightarrow q) \wedge r$ or $p \Rightarrow (q \wedge r)$, which are very different propositions. The truth value of such complicated propositions is determined by systematic applications of the rules for the connectives.

The possible truth values of a proposition are often listed in a table, called a **truth table**. If p and q are propositions, then there are four possible combinations of truth values for p and q. That is, TT, TF, FT and FF. If a third proposition r is involved, then there are eight possible combinations of truth values for p,q and r. In general, a truth table involving "n" propositions $p_1, p_2, \ldots, p_n$ contains 2^n possible combinations of truth values for these propositions and a truth table showing these combinations would have n columns and 2^n rows. So, we use truth tables to determine the truth value of a compound proposition based on the truth value of its constituent component propositions.

Examples 1.6:

a. Suppose p and r are true and q and s are false.
   What is the truth value of $(p \wedge q) \Rightarrow (r \vee s)$?
   i. Since p is true and q is false, $p \wedge q$ is false.
   ii. Since r is true and s is false, $r \vee s$ is true.
   iii. Thus by applying the rule of implication, we get that $(p \wedge q) \Rightarrow (r \vee s)$ is true.
b. Suppose that a compound proposition is symbolized by

$$(p \vee q) \Rightarrow (r \Leftrightarrow \neg s)$$

and that the truth values of $p, q, r,$ and s are $T, F, F,$ and T, respectively. Then the truth value of $p \vee q$ is T, that of $\neg s$ is F, that of $r \Leftrightarrow \neg s$ is T. So the truth value of $(p \vee q) \Longrightarrow (r \Leftrightarrow \neg s)$ is T.

Remark: When dealing with compound propositions, we shall adopt the following convention on the use of parenthesis. Whenever "∨" or "∧" occur with "⟹" or "⟺", we shall assume that "∨" or "∧" is applied first, and then "⟹" or "⟺" is then applied. For example,

$$p \wedge q \Longrightarrow r \text{ means } (p \wedge q) \Longrightarrow r$$

$$p \vee q \Leftrightarrow r \text{ means } (p \vee q) \Leftrightarrow r$$

$$\neg q \Longrightarrow \neg p \text{ means } (\neg q) \Longrightarrow (\neg p)$$

$$\neg q \Longrightarrow p \Leftrightarrow r \text{ means } ((\neg q) \Longrightarrow p) \Leftrightarrow r$$

However, it is always advisable to use brackets to indicate the order of the desired operations. .

Definition 1.3: Two compound propositions P and Q are said to be ***equivalent*** if they have the same truth value for all possible combinations of truth values for the component propositions occurring in both P and Q. In this case we write $P \equiv Q$.

Example 1.7: Let $P: p \Longrightarrow q$.
Q: ¬(p ∧ ¬q)

p	q	$\neg p$	$\neg q$	$p \Longrightarrow q$	$\neg q \Longrightarrow \neg p$
T	***T***	***F***	***F***	***T***	***T***
T	***F***	***F***	***T***	***F***	***F***
F	***T***	***T***	***F***	***T***	***T***
F	***F***	***T***	***T***	***T***	***T***

Then, P is equivalent to Q, since columns 5 and 6 of the above table are identical.

Example 1.8: Let $P: p \Longrightarrow q$.
$Q: \neg p \Longrightarrow \neg q$.

Then

p	q	$\neg p$	$\neg q$	$p \Longrightarrow q$	$\neg p \Longrightarrow \neg q$
T	***T***	***F***	***F***	***T***	***T***
T	***F***	***F***	***T***	***F***	***T***
F	***T***	***T***	***F***	***T***	***F***
F	***F***	***T***	***T***	***T***	***T***

Looking at columns 5 and 6 of the table we see that they are not identical. Thus $P \not\equiv Q$.

It is useful at this point to mention the non-equivalence of certain conditional propositions. Given the conditional $p \Longrightarrow q$, we give the related conditional propositions:-

$q \Longrightarrow p$: Converse of $p \Longrightarrow q$

$\neg p \Longrightarrow \neg q$: Inverse of $p \Longrightarrow q$

$\neg q \Longrightarrow \neg p$: Contrapositive of $p \Longrightarrow q$

As we observed from example 1.7, the conditional $p \Longrightarrow q$ and its contrapositve $\neg q \Longrightarrow \neg p$ are equivalent. On the other hand, $p \Longrightarrow q \not\equiv q \Longrightarrow p$ and $p \Longrightarrow q \not\equiv \neg q \Longrightarrow \neg p$.

Do not confuse the contrapositive and the converse of the conditional proposition. Here is the difference:

Converse: The hypothesis of a converse statement is the conclusion of the conditional statement and the conclusion of the converse statement is the hypothesis of the conditional statement.

Contrapositive: The hypothesis of a contrapositive statement is the negation of conclusion of the conditional statement and the conclusion of the contrapositive statement is the negation of hypothesis of the conditional statement.

Example 1.9:

a. If Kidist lives in Addis Ababa, then she lives in Ethiopia.
   Converse: If Kidist lives in Ethiopia, then she lives in Addis Ababa.
   Contrapositive: If Kidist does not live in Ethiopia, then she does not live in Addis Ababa.
   Inverse: If Kidist does not live in Addis Ababa, then she does not live in Ethiopia.
b. If it is morning, then the sun is in the east.
   Converse: If the sun is in the east, then it is morning.
   Contrapositive: If the sun is not in the east, then it is not morning.
   Inverse: If it is not morning, then the sun is not the east.

Propositions, under the relation of logical equivalence, satisfy various laws or identities, which are listed below.

1. Idempotent Laws
   a. $p \equiv p \vee p$.
   b. $p \equiv p \wedge p$.
2. Commutative Laws
   a. $p \wedge q \equiv q \wedge p$.
   b. $p \vee q \equiv q \vee p$.
3. Associative Laws
   a. $p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$.
   b. $p \vee (q \vee r) \equiv (p \vee q) \vee r$.
4. Distributive Laws
   a. $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$.
   b. $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$.
5. De Morgan's Laws
   a. $\neg(p \wedge q) \equiv \neg p \vee \neg q$.

b. $\neg(p \vee q) \equiv \neg p \wedge \neg q$

6. Law of Contrapositive

$p \Longrightarrow q \equiv \neg q \Longrightarrow \neg p$

7. Complement Law

$\neg(\neg p) \equiv p.$

1.1.4. Tautology and contradiction

Definition: A compound proposition is a **tautology** if it is always true regardless of the truth values of its component propositions. If, on the other hand, a compound proposition is always false regardless of its component propositions, we say that such a proposition is a **contradiction**.

Examples 1.10:

a. Suppose p is any proposition. Consider the compound propositions $p \vee \neg p$ and $p \wedge \neg p$.

p	$\neg p$	$p \vee \neg p$	$p \wedge \neg p$
T	F	T	F
F	T	T	F

Observe that $p \vee \neg p$ is a tautology while $p \wedge \neg p$ is a contradiction.

b. For any propositions p and q. Consider the compound proposition $p \Longrightarrow (q \Longrightarrow p)$. Let us make a truth table and study the situation.

p	q	$q \Longrightarrow p$	$p \Longrightarrow (q \Longrightarrow p)$
T	T	T	T
T	F	T	T
F	T	F	T
F	F	T	T

We have exhibited all the possibilities and we see that for all truth values of the constituent propositions, the proposition $p \Longrightarrow (q \Longrightarrow p)$ is always true. Thus, $p \Longrightarrow (q \Longrightarrow p)$ is a tautology.

c. The truth table for the compound proposition $(p \Longrightarrow q) \Longleftrightarrow (p \wedge \neg q)$.

p	q	$\neg q$	$p \wedge \neg q$	$p \Longrightarrow q$	$(p \Longrightarrow q) \Longleftrightarrow (p \wedge \neg q)$
T	T	F	F	T	F
T	F	T	T	F	F
F	T	F	F	T	F
F	F	T	F	T	F

In example 1.10(c), the given compound proposition has a truth value F for every possible combination of assignments of truth values for the component propositions p and q. Thus $(p \Rightarrow q) \Leftrightarrow (p \wedge \neg q)$ is a contradiction.

Remark:

1. In a truth table, if a proposition is a tautology, then every line in its column has T as its entry; if a proposition is a contradiction, every line in its column has F as its entry.
2. Two compound propositions P and Q are equivalent if and only if "$P \Leftrightarrow Q$" is a tautology.

Exercises

1. For statements p, q and r, use a truth table to show that each of the following pairs of statements is logically equivalent.
   a. $(p \wedge q) \Leftrightarrow p$ and $p \Rightarrow q$.
   b. $p \Rightarrow (q \vee r)$ and $\neg q \Rightarrow (\neg p \vee r)$.
   c. $(p \vee q) \Rightarrow r$ and $(p \Rightarrow q) \wedge (q \Rightarrow r)$.
   d. $p \Rightarrow (q \vee r)$ and $(\neg r) \Rightarrow (p \Rightarrow q)$.
   e. $p \Rightarrow (q \vee r)$ and $((\neg r) \wedge p) \Rightarrow q$.
2. For statements p, q, and r, show that the following compound statements are tautology.
   a. $p \Rightarrow (p \vee q)$.
   b. $(p \wedge (p \Rightarrow q)) \Rightarrow q$.
   c. $\big((p \Rightarrow q) \wedge (q \Rightarrow r)\big) \Rightarrow (p \Rightarrow r)$.
3. For statements p and q, show that $(p \wedge \neg q) \wedge (p \wedge q)$ is a contradiction.
4. Write the contrapositive and the converse of the following conditional statements.
   a. If it is cold, then the lake is frozen.
   b. If Solomon is healthy, then he is happy.
   c. If it rains, Tigist does not take a walk.
5. Let p and q be statements. Which of the following implies that $p \vee q$ is false?
   a. $\neg p \vee \neg q$ is false.
   b. $\neg p \vee q$ is true.
   c. $\neg p \wedge \neg q$ is true.
   d. $p \Rightarrow q$ is true.
   e. $p \wedge q$ is false.
6. Suppose that the statements p, q, r, and s are assigned the truth values T, F, F, and T, respectively. Find the truth value of each of the following statements.
   a. $(p \vee q) \vee r$.
   b. $p \vee (q \vee r)$.
   c. $r \Rightarrow (s \wedge p)$.
   d. $p \Rightarrow (r \Rightarrow s)$.
   e. $p \Rightarrow (r \vee s)$.
   f. $(p \vee r) \Leftrightarrow (r \wedge \neg s)$.
   g. $(s \Leftrightarrow p) \Rightarrow (\neg p \vee s)$.
   h. $(q \wedge \neg s) \Rightarrow (p \Leftrightarrow s)$.
   i. $(r \wedge s) \Rightarrow (p \Rightarrow (\neg q \vee s))$.
   j. $(p \vee \neg q) \vee r \Rightarrow (s \wedge \neg s)$.
7. Suppose the value of $p \Rightarrow q$ is T; what can be said about the value of $\neg p \wedge q \Leftrightarrow p \vee q$?

8. a. Suppose the value of $p \Leftrightarrow q$ is T; what can be said about the values of $p \Leftrightarrow \neg q$ and $\neg p \Leftrightarrow q$?
   b. Suppose the value of $p \Leftrightarrow q$ is F; what can be said about the values of $p \Leftrightarrow \neg q$ and $\neg p \Leftrightarrow q$?
9. Construct the truth table for each of the following statements.
   a. $p \Rightarrow (p \Rightarrow q)$.
   b. $(p \vee q) \Leftrightarrow (q \vee p)$.
   c. $p \Rightarrow \neg(q \wedge r)$.
   d. $(p \Rightarrow q) \Leftrightarrow (\neg p \vee q)$.
   e. $\big(p \Rightarrow (q \wedge r)\big) \vee (\neg p \wedge q)$.
   f. $(p \wedge q) \Rightarrow ((q \wedge \neg q) \Rightarrow (r \wedge q))$.
10. For each of the following determine whether the information given is sufficient to decide the truth value of the statement. If the information is enough, state the truth value. If it is insufficient, show that both truth values are possible.
   a. $(p \Rightarrow q) \Rightarrow r$, where $r = T$.
   b. $p \wedge (q \Rightarrow r)$, where $q \Rightarrow r = T$.
   c. $p \vee (q \Rightarrow r)$, where $q \Rightarrow r = T$.
   d. $\neg(p \vee q) \Leftrightarrow (\neg p \wedge \neg q)$, where $p \vee q = T$.
   e. $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$, where $q = T$.
   f. $(p \wedge q) \Rightarrow (p \vee s)$, where $p = T$ and $s = F$.

1.2. Open propositions and quantifiers

In mathematics, one frequently comes across sentences that involve a variable. For example, $x^2 + 2x - 3 = 0$ is one such. The truth value of this statement depends on the value we assign for the variable x. For example, if $x = 1$, then this sentence is true, whereas if $x = -1$, then the sentence is false.

Section objectives:

After completing this section, students will be able to:-

- ✓ Define open proposition.
- ✓ Analyze the difference between proposition and open proposition.
- ✓ Differentiate the two types of quantifiers.
- ✓ Convert open propositions into propositions using quantifiers.
- ✓ Determine the truth value of a quantified proposition.
- ✓ Convert a quantified proposition into words and vise versa.
- ✓ Explain the relationship between existential and universal quantifiers.
- ✓ Analyze quantifiers occurring in combinations.

Definition 1.4: An open statement (also called a predicate) is a sentence that contains one or more variables and whose truth value depends on the values assigned for the variables. We represent an open statement by a capital letter followed by the variable(s) in parenthesis, e.g., $P(x), Q(x)$ etc.

Example 1.11: Here are some open propositions:

a. x is the day before Sunday.
b. y is a city in Africa.
c. x is greater than y.
d. $x + 4 = -9$.

It is clear that each one of these examples involves variables, but is not a proposition as we cannot assign a truth value to it. However, if individuals are substituted for the variables, then each one of them is a proposition or statement. For example, we may have the following.

a. Monday is the day before Sunday.
b. London is a city in Africa.
c. 5 is greater than 9.
d. –13 + 4= –9

Remark

The collection of all allowable values for the variable in an open sentence is called the **universal set** (the universe of discourse) and denoted by $\boldsymbol{U}$.

Definition 1.5: Two open proposition $P(x)$ and $Q(x)$ are said to be equivalent if and only if $P(a) = Q(a)$ for all individual a. Note that if the universe U is specified, then $P(x)$ and $Q(x)$ are equivalent if and only if $P(a) = Q(a)$ for all $a \in U$.

Example 1.12: Let $P(x): x^2 - 1 = 0$.

$Q(x): |x| \geq 1$.

Let $U = \{-1, -\frac{1}{2}, 0, 1\}$.

Then for all $a \in U$; $P(a)$ and $Q(a)$ have the same truth value.

$P(-1): (-1)^2 - 1 = 0 \quad (T) \qquad Q(-1): |-1| \geq 1 \quad (T)$

$P\left(-\frac{1}{2}\right): (-\frac{1}{2})^2 - 1 = 0 \quad (F) \qquad Q\left(-\frac{1}{2}\right): \left|-\frac{1}{2}\right| \geq 1 \quad (F)$

$P(0): 0 - 1 = 0 \quad (F) \qquad Q(0): |0| \geq 1 \quad (F)$

$P(1): 1 - 1 = 0 \quad (T) \qquad Q(1): |1| \geq 1 \quad (T)$

Therefore $P(a) = Q(a)$ for all $a \in U$.

Definition 1.6: Let U be the universal set. An open proposition $P(x)$ is a tautology if and only if $P(a)$ is always true for all values of $a \in U$.

Example 1.13: The open proposition $P(x): x^2 \geq 0$ is a tautology.

As we have observed in example 1.11, an open proposition can be converted into a proposition by substituting the individuals for the variables. However, there are other ways that an open proposition can be converted into a proposition, namely by a method called quantification. Let $P(x)$ be an open proposition over the domain S. Adding the phrase "For every $x \in S$" to $P(x)$ or "For some $x \in S$" to $P(x)$ produces a statement called a quantified statement.

Consider the following open propositions with universe $\mathbb{R}$.

a. $R(x): x^2 \geq 0$.
b. $P(x): (x+2)(x-3) = 0$.
c. $Q(x): x^2 < 0$.

Then $R(x)$ is always true for each $x \in \mathbb{R}$,
$P(x)$ is true only for $x = -2$ and $x = 3$,
$Q(x)$ is always false for all values of $x \in \mathbb{R}$.

Hence, given an open proposition $P(x)$, with universe U, we observe that there are three possibilities.

a. $P(x)$ is true for all $x \in U$.
b. $P(x)$ is true for some $x \in U$.
c. $P(x)$ is false for all $x \in U$.

Now we proceed to study open propositions which are satisfied by "**all**" and "**some**" members of the given universe.

a. The phrase "for every x " is called a **universal quantifier**. We regard "for every x," "for all x," and "for each x " as having the same meaning and symbolize each by "$(\forall x)$." Think of the symbol $\forall$ as an inverted A(representing all). If $P(x)$ is an open proposition with universe U, then $(\forall x)P(x)$ is a quantified proposition and is read as "every $x \in U$ has the property P."
b. The phrase "there exists an x " is called an **existential quantifier**. We regard "there exists an x," "for some x," and "for at least one x " as having the same meaning, and symbolize each by "$(\exists x)$." Think of the symbol $\exists$ as the backwards capital E(representing exists). If $P(x)$ is an open proposition with universe U, then $(\exists x)P(x)$ is a quantified proposition and is read as "there exists $x \in U$ with the property P."

Remarks:

i. To show that $(\forall x)P(x)$ is F, it is sufficient to find at least one $a \in U$ such that $P(a)$ is F. Such an element $a \in U$ is called a ***counter example***.
ii. $(\exists x)P(x)$ is F if we cannot find any $a \in U$ having the property P.

Example 1.14:

a. Write the following statements using quantifiers.
   i. For each real number $x > 0, x^2 + x - 6 = 0$.

Solution: $(\forall x > 0)(x^2 + x - 6 = 0)$.

ii. There is a real number $x > 0$ such that $x^2 + x - 6 = 0$.

Solution: $(\exists x > 0)(x^2 + x - 6 = 0)$.

iii. The square of any real number is nonnegative.

Solution: $(\forall x \in \mathbb{R})(x^2 \geq 0)$.

b.

i. Let $P(x): x^2 + 1 \geq 0$. The truth value for $(\forall x)P(x)$ [i.e $(\forall x)(x^2 + 1 \geq 0)$] is T.

ii. Let $P(x): x < x^2$. The truth value for $(\forall x)(x < x^2)$ is F. $x = \frac{1}{2}$ is a counterexample since $\frac{1}{2} \in \mathbb{R}$ but $\frac{1}{2} < \frac{1}{4}$. On the other hand, $(\exists x)P(x)$ is true, since $-1 \in \mathbb{R}$ such that $-1 < 1$.

iii. Let $P(x): |x| = -1$. The truth value for $(\exists x)P(x)$ is F since there is no real number whose absolute value is -1.

Relationship between the existential and universal quantifiers

If $P(x)$ is a formula in x, consider the following four statements.

a. $(\forall x)P(x)$.
b. $(\exists x)P(x)$.
c. $(\forall x)\neg P(x)$.
d. $(\exists x)\neg P(x)$.

We might translate these into words as follows.

a. Everything has property P.
b. Something has property P.
c. Nothing has property P.
d. Something does not have property P.

Now (d) is the denial of (a), and (c) is the denial of (b), on the basis of everyday meaning. Thus, for example, the existential quantifier may be defined in terms of the universal quantifier.

Now we proceed to discuss the negation of quantifiers. Let $P(x)$ be an open proposition. Then $(\forall x)P(x)$ is false only if we can find an individual "a" in the universe such that $P(a)$ is false. If we succeed in getting such an individual, then $(\exists x)\neg P(x)$ is true. Hence $(\forall x)P(x)$ will be false if $(\exists x)\neg P(x)$ is true. Therefore the negation of $(\forall x)P(x)$ is $(\exists x)\neg P(x)$. Hence we conclude that

$$\neg(\forall x)P(x) \equiv (\exists x)\neg P(x).$$

Similarly, we can easily verified that

$$\neg(\exists x)P(x) \equiv (\forall x)\neg P(x).$$

Remark: To negate a statement that involves the quantifiers ∀ and ∃, change each ∀ to ∃, change each ∃ to ∀, and negate the open statement.

Example 1.15:

Let $U = \mathbb{R}$.

a. $\neg(\exists x)(x < x^2) \equiv (\forall x)\neg(x < x^2)$
$$\equiv (\forall x)(x \geq x^2).$$

b. $\neg(\forall x)(4x + 1 = 0) \equiv (\exists x)\neg(4x + 1 = 0)$
$$\equiv (\exists x)(4x + 1 \neq 0).$$

Given propositions containing quantifiers we can form a compound proposition by joining them with connectives in the same way we form a compound proposition without quantifiers. For example, if we have $(\forall x)P(x)$ and $(\exists x)Q(x)$ we can form $(\forall x)P(x) \Leftrightarrow (\exists x)Q(x)$.

Consider the following statements involving quantifiers. Illustrations of these along with translations appear below.

a. All rationals are reals. $(\forall x)(\mathbb{Q}(x) \Rightarrow \mathbb{R}(x))$.
b. No rationals are reals. $(\forall x)(\mathbb{Q}(x) \Rightarrow \neg\mathbb{R}(x))$.
c. Some rationals are reals. $(\exists x)(\mathbb{Q}(x) \wedge \mathbb{R}(x))$.
d. Some rationals are not reals. $(\exists x)(\mathbb{Q}(x) \wedge \neg\mathbb{R}(x))$.

Example 1.16:

Let $U =$ The set of integers.

Let $P(x)$: x is a prime number.

$Q(x)$: x is an even number.

$R(x)$: x is an odd number.

Then

a. $(\exists x)[P(x) \Rightarrow Q(x)]$ is T; since there is an x, say 2, such that $P(2) \Rightarrow Q(2)$ is T.
b. $(\forall x)[P(x) \Rightarrow Q(x)]$ is F. As a counterexample take 7. Then $P(7)$ is T and $Q(7)$ is F. Hence $P(7) \Rightarrow Q(7)$.
c. $(\forall x)[R(x) \wedge P(x)]$ is F.
d. $(\forall x)[(R(x) \wedge P(x)) \Rightarrow Q(x)]$ is F.

Quantifiers Occurring in Combinations

So far, we have only considered cases in which universal and existential quantifiers appear simply. However, if we consider cases in which universal and existential quantifiers occur in combination, we are lead to essentially new logical structures. The following are the simplest forms of combinations:

1. $(\forall x)(\forall y)P(x, y)$
   "for all x and for all y the relation $P(x, y)$ holds";
2. $(\exists x)(\exists y)P(x, y)$
   "there is an x and there is a y for which $P(x, y)$ holds";
3. $(\forall x)(\exists y)P(x, y)$
   "for every x there is a y such that $P(x, y)$ holds";

4. $(\exists x)(\forall y)P(x,y)$
   "there is an x which stands to every y in the relation $P(x,y)$."

Example 1.17:

Let U = The set of integers.

Let $P(x,y): x + y = 5$.

a. $(\exists x)\,(\exists y)\,P(x,y)$ means that there is an integer x and there is an integer y such that $x + y = 5$. This statement is true when $x = 4$ and $y = 1$, since $4 + 1 = 5$. Therefore, the statement $(\exists x)\,(\exists y)\,P(x,y)$ is always true for this universe. There are other choices of x and y for which it would be true, but the symbolic statement merely says that there is at least one choice for x and y which will make the statement true, and we have demonstrated one such choice.
b. $(\exists x)\,(\forall y)\,P(x,y)$ means that there is an integer x_0 such that for every y, $x_0 + y = 5$. This is false since no fixed value of x_0 will make this true for all y in the universe; e.g. if $x_0 = 1$, then $1 + y = 5$ is false for some y.
c. $(\forall x)\,(\exists y)\,P(x,y)$ means that for every integer x, there is an integer y such that $x + y = 5$. Let $x = a$, then $y = 5 - a$ will always be an integer, so this is a true statement.
d. $(\forall x)\,(\forall y)\,P(x,y)$ means that for every integer x and for every integer y, $x + y = 5$. This is false, for if $x = 2$ and $y = 7$, we get $2 + 7 = 9 \neq 5$.

Example 1.18:

a. Consider the statement

For every two real numbers x and y,$x^2 + y^2 \geq 0$.

If we let

$$P(x,y): x^2 + y^2 \geq 0$$

where the domain of both x and y is $\mathbb{R}$, the statement can be expressed as

$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})P(x,y)$ or as $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(x^2 + y^2 \geq 0)$.

Since $x^2 \geq 0$ and $y^2 \geq 0$ for all real numbers x and y, it follows that $x^2 + y^2 \geq 0$ and so $P(x,y)$ is true for all real numbers x and y. Thus the quantified statement is true.

b. Consider the open statement

$$P(x,y): |x - 1| + |y - 2| \leq 2$$

where the domain of the variable x is the set E of even integers and the domain of the variable y is the set O of odd integers. Then the quantified statement

$$(\exists x \in E)(\exists y \in O)P(x,y)$$

can be expressed in words as

There exist an even integer x and an odd integer y such that $|x-1|+|y-2| \leq 2$.

Since $P(2,3)\colon 1+1 \leq 2$ is true, the quantified statement is true.

c. Consider the open statement

$$P(x,y)\colon xy = 1$$

where the domain of both x and y is the set $\mathbb{Q}^+$ of positive rational numbers. Then the quantified statement

$$(\forall x \in \mathbb{Q}^+)(\exists y \in \mathbb{Q}^+)P(x,y)$$

can be expressed in words as

For every positive rational number x, there exists a positive rational number y such that $xy = 1$. It turns out that the quantified statement is true. If we replace $\mathbb{Q}^+$ by $\mathbb{R}$, then we have

$$(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})P(x,y)\ .$$

Since $x = 0$ and for every real number $y, xy = 0 \neq 1$, $(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})P(x,y)$ is false.

d. Consider the open statement

$$P(x,y)\colon xy \text{ is odd}$$

where the domain of both x and y is the set $\mathbb{N}$ of natural numbers. Then the quantified statement

$$(\exists x \in \mathbb{N})(\forall y \in \mathbb{N})P(x,y),$$

expressed in words, is

There exists a natural number x such that for every natural numbers y, xy is odd. The statement is false.

In general, from the meaning of the universal quantifier it follows that in an expression $(\forall x)(\forall y)P(x,y)$ the two universal quantifiers may be interchanged without altering the sense of the sentence. This also holds for the existential quantifies in an expression such as $(\exists x)(\exists y)P(x,y)$.

In the statement $(\forall x)(\exists y)P(x,y)$, the choice of y is allowed to depend on x - the y that works for one x need not work for another x. On the other hand, in the statement $(\exists y)(\forall x)P(x,y)$, the y must work for all x, i.e., y is independent of x. For example, the expression $(\forall x)(\exists y)(x<y)$, where x and y are variables referring to the domain of real numbers, constitutes a true proposition, namely, "For every number x, there is a number y, such that x is less that y," i.e., "given any number, there is a greater number." However, if the order of the symbol $(\forall x)$ and $(\exists y)$ is changed, in this case, we obtain: $(\exists y)(\forall x)(x<y)$, which is a false proposition, namely, "There is a number which is greater than every number." By transposing $(\forall x)$ and $(\exists y)$, therefore, we get a different statement.

The logical situation here is:

$$(\exists y)(\forall x)P(x,y) \Rightarrow (\forall x)(\exists y)P(x,y).$$

Finally, we conclude this section with the remark that there are no mechanical rules for translating sentences from English into the logical notation which has been introduced. In every case one must first decide on the meaning of the English sentence and then attempt to convey that same meaning in terms of predicates, quantifiers, and, possibly, individual constants.

Exercises

1. In each of the following, two open statements $P(x,y)$ and $Q(x,y)$ are given, where the domain of both x and y is $\mathbb{Z}$. Determine the truth value of $P(x,y) \Longrightarrow Q(x,y)$ for the given values of x and y.
   a. $P(x,y): x^2 - y^2 = 0.$ and $Q(x,y): x = y.$ $(x,y) \in \{(1,-1),(3,4),(5,5)\}.$
   b. $P(x,y): |x| = |y|.$ and $Q(x,y): x = y.$ $(x,y) \in \{(1,2),(2,-2),(6,6)\}.$
   c. $P(x,y): x^2 + y^2 = 1.$ and $Q(x,y): x + y = 1.$
   $(x,y) \in \{(1,-1),(-3,4),(0,-1),(1,0)\}.$
2. Let O denote the set of odd integers and let $P(x): x^2 + 1$ is even, and $Q(x): x^2$ is even. be open statements over the domain O. State $(\forall x \in O)P(x)$ and $(\exists y \in O)Q(x)$ in words.
3. State the negation of the following quantified statements.
   a. For every rational number r, the number $\frac{1}{r}$ is rational.
   b. There exists a rational number r such that $r^2 = 2$.
4. Let $P(n): \frac{5n-6}{3}$ is an integer. be an open sentence over the domain $\mathbb{Z}$. Determine, with explanations, whether the following statements are true or false:
   a. $(\forall n \in \mathbb{Z})P(n).$
   b. $(\exists n \in \mathbb{Z})P(n).$
5. Determine the truth value of the following statements.
   a. $(\exists x \in \mathbb{R})(x^2 - x = 0).$
   b. $(\forall x \in \mathbb{N})(x + 1 \geq 2).$
   c. $(\forall x \in \mathbb{R})(\sqrt{x^2} = x).$
   d. $(\exists x \in \mathbb{Q})(3x^2 - 27 = 0).$
   e. $(\exists x \in \mathbb{R})(\exists y \in \mathbb{R})(x + y + 3 = 8).$
   f. $(\exists x \in \mathbb{R})(\exists y \in \mathbb{R})(x^2 + y^2 = 9).$
   g. $(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})(x + y = 5).$
   h. $(\exists x \in \mathbb{R})(\forall y \in \mathbb{R})(x + y = 5)$
6. Consider the quantified statement

   For every $x \in A$ and $y \in A$, $xy - 2$ is prime.

   where the domain of the variables x and y is $A = \{3,5,11\}$.
   a. Express this quantified statement in symbols.
   b. Is the quantified statement in (a) true or false? Explain.
   c. Express the negation of the quantified statement in (a) in symbols.
   d. Is the negation of the quantified in (a) true or false? Explain.
7. Consider the open statement $P(x,y): \frac{x}{y} < 1.$ where the domain of x is $A = \{2,3,5\}$ and the

domain of y is $B = \{2,4,6\}$.

a. State the quantified statement $(\forall x \in A)(\exists y \in B)P(x, y)$ in words.

b. Show quantified statement in (a) is true.

8. Consider the open statement $P(x, y): x - y < 0$. where the domain of x is $A = \{3,5,8\}$ and the domain of y is $B = \{3,6,10\}$.

a. State the quantified statement $(\exists y \in B)(\forall x \in A)P(x, y)$ in words.

b. Show quantified statement in (a) is true.

1. 3. Argument and Validity

Section objectives:

After completing this section, students will be able to:-

- ✓ Define argument (or logical deduction).
- ✓ Identify hypothesis and conclusion of a given argument.
- ✓ Determine the validity of an argument using a truth table.
- ✓ Determine the validity of an argument using rules of inferences.

Definition 1.7: An argument (logical deduction) is an assertion that a given set of statements $p_1, p_2, p_3, \ldots, p_n$, called **hypotheses** or ***premises***, yield another statement Q, called the ***conclusion***. Such a logical deduction is denoted by:

$$p_1, p_2, p_3, \ldots, p_n \vdash Q \text{ or}$$

$$\begin{array}{l} p_1 \\ p_2 \\ \vdots \\ \underline{p_n} \\ Q \end{array}$$

Example 1.19: Consider the following argument:

If you study hard, then you will pass the exam.
You did not pass the exam.
Therefore, you did not study hard.
Let p: You study hard.
q: You will pass the exam.
The argument form can be written as:

$\neg q$

When is an argument form accepted to be correct? In normal usage, we use an argument in order to demonstrate that a certain conclusion follows from known premises. Therefore, we shall require that under any assignment of truth values to the statements appearing, if the premises became all true, then the conclusion must also become true. Hence, we state the following definition.

Definition 1.8: An argument form $p_1, p_2, p_3, \dots, p_n \vdash Q$ is said to be ***valid*** if Q is true whenever all the premises $p_1, p_2, p_3, \dots, p_n$ are true; otherwise it is ***invalid***.

Example 1.20: Investigate the validity of the following argument:

a. $p \Rightarrow q, \neg q \vdash \neg p$

b. $p \Rightarrow q, \neg q \Rightarrow r \vdash p$

c. If it rains, crops will be good. It did not rain. Therefore, crops were not good.

Solution: First we construct a truth table for the statements appearing in the argument forms.

a.

p	q	$\neg p$	$\neg q$	$p \Rightarrow q$
T	***T***	***F***	***F***	***T***
T	***F***	***F***	***T***	***F***
F	***T***	***T***	***F***	***T***
F	***F***	***T***	***T***	***T***

The premises $p \Longrightarrow q$ and $\neg q$ are true simultaneously in row 4 only. Since in this case p is also true, the argument is valid.

b.

p	q	r	$\neg q$	$p \Rightarrow q$	$\neg q \Rightarrow r$
T	***T***	***T***	***F***	***T***	***T***
T	***T***	***F***	***F***	***T***	***T***
T	***F***	***T***	***T***	***F***	***T***
T	***F***	***F***	***T***	***F***	***F***
F	***T***	***T***	***F***	***T***	***T***
F	***T***	***F***	***F***	***T***	***T***
F	***F***	***T***	***T***	***T***	***T***
F	***F***	***F***	***T***	***T***	***F***

The 1st, 2nd, 5th, 6th and 7th rows are those in which all the premises take value T. In the 5th, 6th and 7th rows however the conclusion takes value F. Hence, the argument form is invalid.

c. Let p: It rains.

q: Crops are good.

$\neg p$: It did not rain.

$\neg q$: Crops were not good.

The argument form is $p \Longrightarrow q, \neg p \vdash \neg q$

Now we can use truth table to test validity as follows:

p	q	$\neg p$	$\neg q$	$p \Longrightarrow q$
T	T	F	F	T
T	F	F	T	F
F	T	T	F	T
F	F	T	T	T

The premises $p \Longrightarrow q$ and $\neg p$ are true simultaneously in row 4 only. Since in this case $\neg q$ is also true, the argument is valid.

Remark:

1. What is important in validity is the form of the argument rather than the meaning or content of the statements involved.
2. The argument form $p_1, p_2, p_3, \dots, p_n \vdash Q$ is valid iff the statement $(p_1 \wedge p_2 \wedge p_3 \wedge \dots \wedge p_n) \Longrightarrow Q$ is a tautology.

Rules of inferences

Below we list certain valid deductions called rules of inferences.

1. Modes Ponens

$$\begin{array}{l} p \\ \underline{p \Longrightarrow q} \\ q \end{array}$$

2. Modes Tollens

$$\begin{array}{l} \neg q \\ \underline{p \Longrightarrow q} \\ \neg p \end{array}$$

3. Principle of Syllogism

$$\begin{array}{l} p \Longrightarrow q \\ \underline{q \Longrightarrow r} \\ p \Longrightarrow r \end{array}$$

4. Principle of Adjunction

a.
$$\frac{\begin{array}{c} p \\ q \end{array}}{p \wedge q}$$

b.
$$\frac{q}{p \vee q}$$

5. Principle of Detachment
$$\frac{p \wedge q}{p, q}$$

6. Modes Tollendo Ponens
$$\frac{\begin{array}{c} \neg p \\ p \vee q \end{array}}{q}$$

7. Modes Ponendo Tollens
$$\frac{\begin{array}{c} \neg(p \wedge q) \\ p \end{array}}{\neg q}$$

8. Constructive Dilemma
$$\frac{\begin{array}{c} (p \Rightarrow q) \wedge (r \Rightarrow s) \\ p \vee r \end{array}}{q \vee s}$$

9. Principle of Equivalence
$$\frac{\begin{array}{c} p \Leftrightarrow q \\ p \end{array}}{q}$$

10. Principle of Conditionalization
$$\frac{p}{q \Rightarrow p}$$

Formal proof of validity of an argument

Definition 1.9: A formal proof of a conclusion Q given hypotheses $p_1, p_2, p_3, \dots, p_n$ is a sequence of stapes, each of which applies some inference rule to hypotheses or previously proven statements (antecedent) to yield a new true statement (the consequent).

A formal proof of validity is given by writing on the premises and the statements which follows from them in a single column, and setting off in another column, to the right of each statement, its justification. It is convenient to list all the premises first.

Example 1.21: Show that $p \Rightarrow \neg q, q \vdash \neg p$ is valid.

Solution:

1. q is true — premise
2. $p \Rightarrow \neg q$ — premise
3. $q \Rightarrow \neg p$ — contrapositive of (2)
4. ך p — Modes Ponens using (1) and (3)

Example 1.22: Show that the hypotheses

It is not sunny this afternoon and it is colder than yesterday.
If we go swimming, then it is sunny.
If we do not go swimming, then we will take a canoe trip.
If we take a canoe trip, then we will be home by sunset.

Lead to the conclusion:

We will be home by sunset.

Let p: It is sunny this afternoon.
q: It is colder than yesterday.
r: We go swimming.
s: We take a canoe trip.
t: We will be home by sunset.

Then

1. $\neg p \wedge q$ — hypothesis
2. $\neg p$ — simplification using (1)
3. $r \Rightarrow p$ — hypothesis
4. $\neg r$ — Modus Tollens using (2) and (3)
5. $\neg r \Rightarrow s$ — hypothesis
6. s — Modus Ponens using (4) and (5)
7. $s \Rightarrow t$ — hypothesis
8. t — Modus Ponens using (6) and (7)

Exercises

1. Use the truth table method to show that the following argument forms are valid.
   i. $\neg p \Rightarrow \neg q, q \vdash p.$
   ii. $p \Rightarrow \neg p, p, r \Rightarrow q \vdash \neg r.$
   iii. $p \Rightarrow q, \neg r \Rightarrow \neg q \vdash \neg r \Rightarrow \neg p.$
   iv. $\neg r \vee \neg s, (\neg s \Rightarrow p) \Rightarrow r \vdash \neg p.$
   v. $p \Rightarrow q, \neg p \Rightarrow r, r \Rightarrow s \vdash \neg q \Rightarrow s.$
2. For the following argument given a, b and c below:
   i. Identify the premises.
   ii. Write argument forms.

iii. Check the validity.

a. If he studies medicine, he will get a good job. If he gets a good job, he will get a good wage. He did not get a good wage. Therefore, he did not study medicine.

b. If the team is late, then it cannot play the game. If the referee is here, then the team is can play the game. The team is late. Therefore, the referee is not here.

c. If the professor offers chocolate for an answer, you answer the professor's question. The professor offers chocolate for an answer. Therefore, you answer the professor's question

3. Give formal proof to show that the following argument forms are valid.

a. $\neg p \Rightarrow \neg q, q \vdash p.$

b. $p \Rightarrow \neg q, p, r \Rightarrow q \vdash \neg r.$

c. $p \Rightarrow q, \neg r \Rightarrow \neg q \vdash \neg r \Rightarrow \neg p.$

d. $\neg r \wedge \neg s, (\neg s \Rightarrow p) \Rightarrow r \vdash \neg p.$

e. $p \Rightarrow, \neg p \Rightarrow r, r \Rightarrow s \vdash \neg q \Rightarrow s.$

f. $\neg p \vee q, r \Rightarrow p, r \vdash q.$

g. $\neg p \wedge \neg q, (q \vee r) \Rightarrow p \vdash \neg r.$

h. $p \Rightarrow (q \vee r), \neg r, p \vdash q.$

i. $\neg q \Rightarrow \neg p, r \Rightarrow p, \neg q \vdash r.$

4. Prove the following are valid arguments by giving formal proof.

a. If the rain does not come, the crops are ruined and the people will starve. The crops are not ruined or the people will not starve. Therefore, the rain comes.

b. If the team is late, then it cannot play the game. If the referee is here then the team can play the game. The team is late. Therefore, the referee is not here.

1.4. Set theory

In this section, we study some part of set theory especially description of sets, Venn diagrams and operations of sets.

Section objectives:

After completing this section, students will be able to:-

- ✓ Explain the concept of set.
- ✓ Describe sets in different ways.
- ✓ Identify operations of sets.
- ✓ Illustrate sets using Venn diagrams.

1.4.1. The concept of a set

The term set is an undefined term, just as a point and a line are undefined terms in geometry. However, the concept of a set permeates every aspect of mathematics. Set theory underlies the language and concepts of modern mathematics. The term set refers to a well-defined collection of objects that share a certain property or certain properties. The term "**well-defined**" here means that the set is described in such a way that one can decide whether or not a given object belongs in the set. If A is a set, then the objects of the collection A are called the elements or members of the set A. If x is an element of the set A, we write $x \in A$. If x is not an element of the set A, we write $x \notin A$.

As a convention, we use capital letters to denote the names of sets and lowercase letters for elements of a set.

Note that for each objects x and each set A, exactly one of $x \in A$ or $x \notin A$ but not both must be true.

1.4.2. Description of sets

Sets are described or characterized by one of the following four different ways.

1. **Verbal Method**

In this method, an ordinary English statement with minimum mathematical symbolization of the property of the elements is used to describe a set. Actually, the statement could be in any language.

Example 1.23:

a. The set of counting numbers less than ten.
b. The set of letters in the word "Addis Ababa."
c. The set of all countries in Africa.

2. **Roster/Complete Listing Method**

If the elements of a set can all be listed, we list them all between a pair of braces without repetition separating by commas, and without concern about the order of their appearance. Such a method of describing a set is called ***the roster/complete listing*** method.

Examples 1.24:

a. The set of vowels in English alphabet may also be described as $\{a, e, i, o, u\}$.
b. The set of positive factors of 24 is also described as $\{1, 2, 3, 4, 6, 8, 12, 24\}$.

Remark:

i. We agree on the convention that the order of writing the elements in the list is immaterial. As a result the sets $\{a, b, c\}, \{b, c, a\}$ and $\{c, a, b\}$ contain the same elements, namely a, b and c.
ii. The set $\{a, a, b, b, b\}$ contains just two distinct elements; namely a and b, hence it is the same set as $\{a, b\}$. We list distinct elements without repetition.

Example 1.25:

a. Let $A = \{a, b, \{c\}\}$. Elements of A are a, b and $\{c\}$.

Notice that c and $\{c\}$ are different objects. Here $\{c\} \in A$ but $c \notin A$.

b. Let $B = \{\{a\}\}$. The only element of B is $\{a\}$. But $a \notin B$.

c. Let $C = \{a, b, \{a, b\}, \{a, \{a\}\}\}$. Then C has four elements.

The readers are invited to write down all the elements of C.

3. **Partial Listing Method**

In many occasions, the number of elements of a set may be too large to list them all; and in other occasions there may not be an end to the list. In such cases we look for a common property of the elements and describe the set by partially listing the elements. More precisely, if the common property is simple that it can easily be identified from a list of the first few elements, then with in a pair of braces, we list these few elements followed (or preceded) by exactly three dotes and possibly by one last element. The following are such instances of describing sets by partial listing method.

Example 1.26:

a. The set of all counting numbers is $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$.
b. The set of non-positive integers is $\{\ldots, -4, -3, -2, -1, 0\}$.
c. The set of multiples of 5 is $\{\ldots, -15, -10, -5, 0\ 5, 10, 15, \ldots\}$.
d. The set of odd integers less than 100 is $\{\ldots, -3, -1, 1, 3, 5, \ldots\ 99\}$.

4. **Set-builder Method**

When all the elements satisfy a common property P, we express the situation as an open proposition $P(x)$ and describe the set using a method called the ***Set-builder Method*** as follows:

$$A = \{x \mid P(x)\} \text{ or } A = \{x : P(x)\}$$

We read it as "A is equal to the set of all x's such that $P(x)$ is true." Here the bar "| " and the colon ":" mean "such that." Notice that the letter x is only a place holder and can be replaced throughout by other letters. So, for a property P, the set $\{x \mid P(x)\}, \{t \mid P(t)\}$ and $\{y \mid P(y)\}$ are all the same set.

Example 1.27: The following sets are described using the set-builder method.

a. $A = \{x \mid x \text{ is a vowel in the English alphabet}\}$.
b. $B = \{t \mid t \text{ is an even integer}\}$.
c. $C = \{n \mid n \text{ is a natural number and } 2n - 15 \text{ is negative}\}$.
d. $D = \{y \mid y^2 - y - 6 = 0\}$.
e. $E = \{x \mid x \text{ is an integer and } x - 1 < 0 \Rightarrow x^2 - 4 > 0\}$.

Exercise: Express each of the above by using either the complete or the partial listing method.

Definition 1.10: The set which has no element is called the empty (or null) set and is denoted by ϕ or {}.

Example 1.28: The set of $x \in \mathbb{R}$ such that $x^2 + 1 = 0$ is an empty set.

Relationships between two sets

Definition 1.11: Set B is said to be a ***subset*** of set A (or is contained in A), denoted by $B \subseteq A$, if every element of B is an element of A, i.e.,

$$(\forall x)(x \in B \Longrightarrow x \in A).$$

It follows from the definition that set B is not a subset of set A if at least one element of B is not an element of A. i.e., $B \nsubseteq A \Leftrightarrow (\exists x)(x \in B \Longrightarrow x \notin A)$. In such cases we write $B \nsubseteq A$ or $A \nsupseteq B$.

Remarks: For any set $A, \phi \subseteq A$ and $A \subseteq A$.

Example 1.29:

a. If $A = \{a, b\}$, $B = \{a, b, c\}$ and $C = \{a, b, d\}$, then $A \subseteq B$ and $A \subseteq C$. On the other hand, it is clear that: $B \nsubseteq A$, $B \nsubseteq C$ and $C \nsubseteq B$.

b. If $S = \{x \mid x \text{ is a multiple of } 6\}$ and $T = \{x \mid x \text{ is even integer}\}$, then $S \subseteq T$ since every multiple of 6 is even. However, $2 \in T$ while $2 \notin S$. Thus $T \nsubseteq S$.

c. If $A = \{a, \{b\}\}$, then $\{a\} \subseteq A$and $\{\{b\}\} \subseteq A$. On the other hand, since $b \notin A$, $\{b\} \nsubseteq A$, and $\{a, b\} \nsubseteq A$.

Definition 1.12: Sets A and B are said to be ***equal*** if they contain exactly the same elements. In this case, we write $A = B$. That is,

$$(\forall x)(x \in B \Leftrightarrow x \in A).$$

Example 1.30:

a. The sets $\{1, 2, 3\}, \{2, 1, 3\}, \{1, 3, 2\}$ are all equal.

b. $\{x|xisacountingnumber\} = \{x|xisapositiveinteger\}$

Definition 1.13: Set A is said to be a **proper subset** of set B if every element of A is also an element of B, but B has at least one element that is not in A. In this case, we write $A \subset B$. We also say B is a proper super set of A, and write $B \supset A$. It is clear that

$$A \subset B \Leftrightarrow [(\forall x)(x \in A \Longrightarrow x \in B) \wedge (A \neq B)].$$

Remark: Some authors do not use the symbol $\subseteq$. Instead they use the symbol $\subset$ for both subset and proper subset. In this material, we prefer to use the notations commonly used in high school mathematics, and we continue using $\subseteq$ and $\subset$ differently, namely for subset and proper subset, respectively.

Definition 1.14: Let A be a set. The power set of A, dented by $P(A)$, is the set whose elements are all subsets of A. That is,

$$P(A) = \{B: B \subseteq A\}.$$

Example 1.31: Let $A = \{x, y, z\}$. As noted before, ϕ and A are subset of A. Moreover, $\{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}$ and $\{y, z\}$ are also subsets of A. Therefore,

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

Frequently it is necessary to limit the topic of discussion to elements of a certain fixed set and regard all sets under consideration as a subset of this fixed set. We call this set the **universal set** or the **universe** and denoted by $\boldsymbol{U}$.

Exercises

1. Which of the following are sets?
   a. 1,2,3
   b. {1,2},3
   c. {{1},2},3
   d. {1,{2},3}
   e. {1,2,a,b}.
2. Which of the following sets can be described in complete listing, partial listing and/or set-builder methods? Describe each set by at least one of the three methods.
   a. The set of the first 10 letters in the English alphabet.
   b. The set of all countries in the world.
   c. The set of students of Addis Ababa University in the 2018/2019 academic year.
   d. The set of positive multiples of 5.
   e. The set of all horses with six legs.
3. Write each of the following sets by listing its elements within braces.
   a. $A = \{x \in \mathbb{Z}: -4 < x \leq 4\}$
   b. $B = \{x \in \mathbb{Z}: x^2 < 5\}$
   c. $C = \{x \in \mathbb{N}: x^3 < 5\}$
   d. $D = \{x \in \mathbb{R}: x^2 - x = 0\}$
   e. $E = \{x \in \mathbb{R}: x^2 + 1 = 0\}$.
4. Let A be the set of positive even integers less than 15. Find the truth value of each of the following.
   a. $15 \in A$
   b. $-16 \in A$
   c. $\phi \in A$
   d. $12 \subset A$
   e. $\{2, 8, 14\} \in A$
   f. $\{2,3,4\} \subseteq A$
   g. $\{2,4\} \in A$
   h. $\phi \subset A$

i. $\{246\} \subseteq A$

5. Find the truth value of each of the following and justify your conclusion.
   a. $\phi \subseteq \phi$
   b. $\{1,2\} \subseteq \{1,2\}$
   c. $\phi \in A$ for any set A
   d. $\{\phi\} \subseteq A$, for any set A
   e. $5, 7 \subseteq \{5, 6, 7, 8\}$
   f. $\phi \in \{\{\phi\}\}$
   g. For any set $A, A \subset A$
   h. $\{\phi\} = \phi$
6. For each of the following set, find its power set.
   a. $\{ab\}$
   b. $\{1, 1.5\}$
   c. $\{a, b\}$
   d. $\{a, 0.5, x\}$
7. How many subsets and proper subsets do the sets that contain exactly $1, 2, 3, 4, 8, 10$ and 20 elements have?
8. If n is a whole number, use your observation in Problems 6and 7 to discover a formula for the number of subsets of a set with n elements. How many of these are proper subsets of the set?
9. Is there a set A with exactly the following indicated property?
   a. Only one subset
   b. Only one proper subset
   c. Exactly 3 proper subsets
   d. Exactly 4 subsets
   e. Exactly 6 proper subsets
   f. Exactly 30 subsets
   g. Exactly 14 proper subsets
   h. Exactly 15 proper subsets
10. How many elements does A contain if it has:
    a. 64 subsets?
    b. 31 proper subsets?
    c. No proper subset?
    d. 255 proper subsets?
11. Find the truth value of each of the following.
    a. $\phi \in P(\phi)$
    b. $For\ any\ set\ A, \phi \subseteq P(A)$

c. $For\ any\ set\ A, A \in P(A)$
d. $For\ any\ set\ A, A \subset P(A)$.

12. For any three sets A, B and C, prove that:
    a. If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.
    b. If $A \subset B$ and $B \subset C$, then $A \subset C$.

1.4.3. Set Operations and Venn diagrams

Given two subsets A and B of a universal set U, new sets can be formed using A and B in many ways, such as taking common elements or non-common elements, and putting everything together. Such processes of forming new sets are called ***set operations***. In this section, three most important operations, namely union, intersection and complement are discussed.

Definition 1.15: The union of two sets A and B, denoted by $A \cup B$, is the set of all elements that are either in A or in B (or in both sets). That is,

$$A \cup B = \{x: (x \in A) \vee (x \in B)\}.$$

As easily seen the union operator "∪" in the theory of set is the counterpart of the logical operator "∨".

Definition 1.16: The intersection of two sets A and B, denoted by $A \cap B$, is the set of all elements that are in A and B. That is,

$$A \cap B = \{x: (x \in A) \wedge (x \in B)\}.$$

As suggested by definition 1.15, the intersection operator "∩" in the theory of sets is the counterpart of the logical operator "∧".

Note: - Two sets A and B are said to be disjoint sets if $A \cap B = \phi$.

Example 1.32:

a. Let $A = \{0, 1, 3, 5, 6\}$ and $B = \{1, 2, 3, 4, 6, 7\}$. Then,
   $A \cup B = \{0, 1, 2, 3, 4, 5, 6, 7\}$ and $A \cap B = \{1, 3, 6\}$.

b. Let A = The set of positive even integers, and
   B = The set of positive multiples of 3. Then,

$A \cup B = \{x: x$ is a positive intger that is either even or a multiple of 3$\}$
$= \{2, 3, 4, 6, 8, 9, 10, 12, 14, 15, 16, \dots\}$

$A \cap B = \{x \mid x$ is a positive integer that is both even and multiple of 3$\}$
$= \{6, 12, 18, 24, \dots\}$
$= \{x \mid x$ is a positive multiple of 6.$\}$

Definition 1.17: The difference between two sets A and B, denoted by $A - B$, is the of all elements in A and not in B; this set is also called the **relative complement** of B with respect to A. Symbolically,

$$A - B = \{x: x \in A \wedge x \notin B\}.$$

Example 1.33: If $= \{1,3,5\}$, $B = \{1,2\}$, then $A - B = \{3,5\}$ and $B - A = \{2\}$.
Note: The above example shows that, in general, $A - B$ are $B - A$ disjoint.

Definition 1.18: Let A be a subset of a universal set U. The ***absolute complement*** (or simply ***complement***) of A, denoted by A' (or A^cor $\overline{A}$), is defined to be the set of all elements of U that are not in A. That is,

$$A' = \{x: x \in U \wedge x \notin A\} \text{ or } x \in A' \Leftrightarrow x \notin A \Leftrightarrow \neg(x \in A).$$

Notice that taking the absolute complement of A is the same as finding the relative complement of A with respect to the universal set U. That is,

$$A' = U - A.$$

Example 1.34:

a. If $U = \{0,1,2,3,4\}$, and if $A = \{3,4\}$, then $A' = \{3,4\}$.
b. Let $U = \{1, 2, 3, \ldots, 12\}$
   $A = \{x \mid x \text{ is a positive factor of } 12\}$
   and $B = \{x \mid x \text{ is an odd integer in } U\}$.
   Then, $A' = \{5, 7, 8, 9, 10, 11\}$, $B' = \{2, 4, 6, 8, 10, 12\}$,
   $(A \cup B)' = (8, 10\}$, $A' \cup B' = \{2, 4, 5, 6, \ldots, 12\}$,
   $A' \cap B' = \{8, 10\}$, and $(A \backslash B)' = \{1, 3, 5, 7, 8, 9, 10, 11\}$.
c. Let $U = \{a, b, c, d, e, f, g, h\}$, $A = \{a, e, g, h\}$ and
   $B = \{b, c, e, f, h\}$. Then
   $A' = \{b, c, d, f\}$, $B' = \{a, d, g\}$, $B - A = \{b, c, f\}$,
   $A - B = \{a, g\}$, and $(A \cup B)' = \{d\}$.

Find $(A \cap B)'$, $A' \cap B'$, $A' \cup B'$. Which of these are equal?

Theorem 1.1: For any two sets A and B, each of the following holds.

1. $(A')' = A$.
2. $A' = U - A$.
3. A– B = A ∩ B′A– B = A∩B' $A - B = A \cap B'$.
4. **$(A \cup B)' = A' \cap B'$.**
5. $(A \cap B)' = A' \cup B'$.
6. $A \subseteq B \Leftrightarrow B' \subseteq A'$.

Now we define the symmetric difference of two sets.

Definition 1.17: The symmetric difference of two sets A and B, denoted by $A\Delta B$, is the set
$$A\Delta B = (A - B) \cup (B - A).$$

Example 1.35: Let $U = \{1,2,3, \dots ,10\}$ be the universal set, $A = \{2,4,6,8,9,10\}$ and $B = \{3,5,7,9\}$. Then $B - A = \{3,5,7\}$ and $A - B = \{2,4,6,8,10\}$. Thus $A\Delta B = \{2,3,4,5,6,7,8,10\}$.

Theorem 1.2: For any three sets A, B and C, each of the following holds.

a. $A \cup B = B \cup A$. (∪ is commutative)
b. $A \cap B = B \cap A$. (∩ is commutative)
c. $(A \cup B) \cup C = A \cup (B \cup C)$. (∪ is associative)
d. $(A \cap B) \cap C = A \cap (B \cap C)$. (∩ is associative)
e. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. (∪ is distributive over ∩)
f. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. (∩ is distributive over ∪)

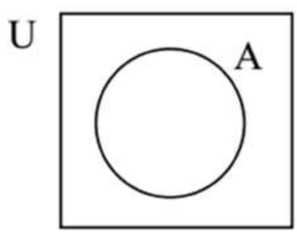
Let us prove property "e" formally.

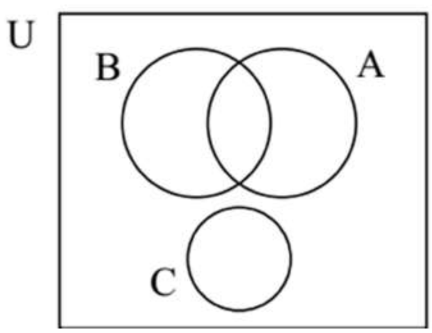
$x \in A \cup (B \cap C) \Leftrightarrow (x \in A) \vee (x \in B \cap C)$ (definition of ∪)
$\Leftrightarrow x \in A \vee (x \in B \wedge x \in C)$ (definition of ∩)
$\Leftrightarrow (x \in A \vee x \in B) \wedge (x \in A \vee x \in C)$ (∨ is distributive over ∧
$\Leftrightarrow (x \in A \cup B) \wedge (x \in A \cup C)$ (definition of ∪)
$\Leftrightarrow x \in (A \cup B) \cap (A \cup C)$ (definition of ∩)

Therefore, we have $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
The readers are invited to prove the rest part of theorem (1.2).

Venn diagrams

While working with sets, it is helpful to use diagrams, called **Venn diagrams,** to illustrate the relationships involved. A Venn diagram is a schematic or pictorial representative of the sets involved in the discussion. Usually sets are represented as interlocking circles, each of which is enclosed in a rectangle, which represents the universal set U.

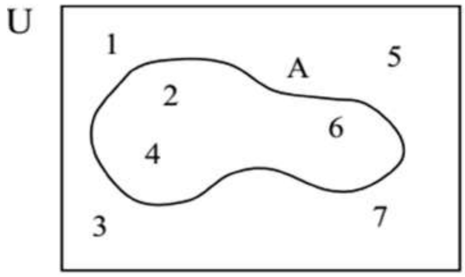




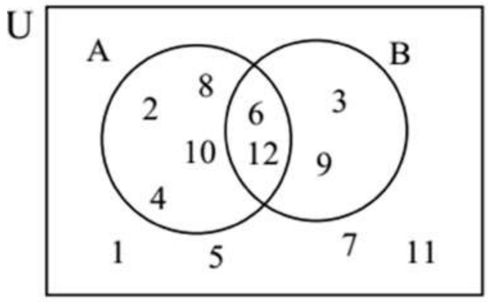
In some occasions, we list the elements of set A inside the closed curve representing A.

Example 1.36:

a. If $U = \{1, 2, 3, 4, 5, 6, 7\}$ and $A = \{2, 4, 6\}$, then a Venn diagram representation of these two sets looks like the following.

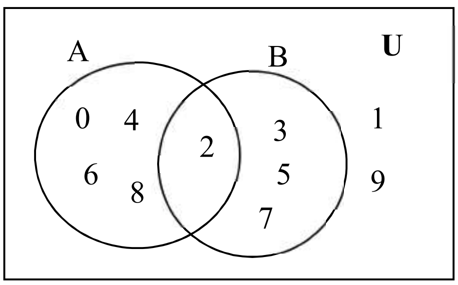


b. Let $U = \{x \mid x \text{ is a positive integer less than } 13\}$
$A = \{x \mid x \in U \text{ and } x \text{ is even}\}$
$B = \{x \mid x \in U \text{ and } x \text{ is a multiple of } 3\}$.
A Venn diagram representation of these sets is given below.

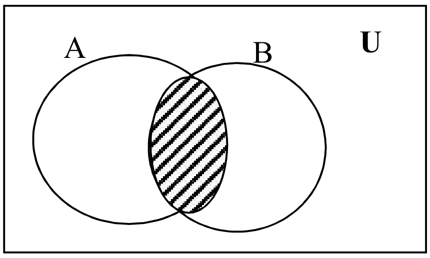


Example 1.37: Let U = The set of one digits numbers
A = The set of one digits even numbers
B = The set of positive prime numbers less than 10

We illustrate the sets using a Venn diagram as follows.

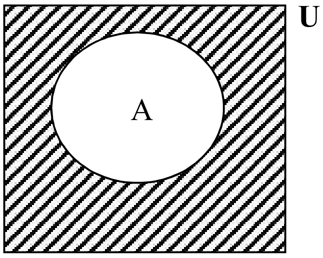


a. Illustrate $A \cap B$ by a Venn diagram



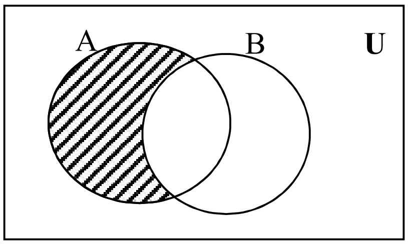
$A \cap B$: The shaded portion

b. Illustrate A' by a Venn diagram



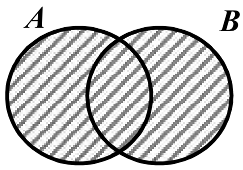
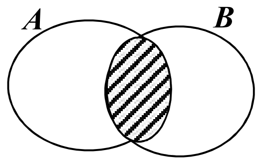
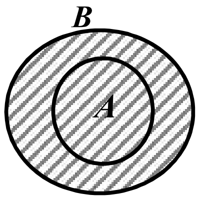
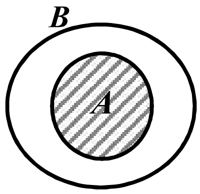
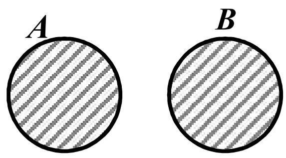
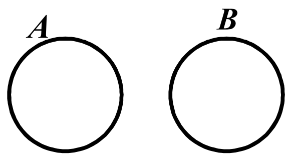
A' : The shaded portion

c. Illustrate $A \backslash B$ by using a Venn diagram



$A \setminus B$: The shaded portion

Now we illustrate intersections and unions of sets by Venn diagram.

Cases	Shaded is $A \cup B$	Shaded $A \cap B$
Only some common elements		
$A \subseteq B$		
No common element		 $A \cap B = \varnothing$

Exercises

1. If $B \subseteq A$, $A \cap B' = \{1,4,5\}$ and $A \cup B = \{1,2,3,4,5,6\}$, find B.
2. Let $A = \{2,4,6,7,8,9\}$,
   $B = \{1,3,5,6,10\}$ and
   $C = \{\, x: 3x + 6 = 0$ or $2x + 6 = 0\}$. Find
   a. $A \cup B$.
   b. Is $(A \cup B) \cup C = A \cup (B \cup C)$?
3. Suppose $U =$ The set of one digit numbers and
   $A = \{x: x$ is an even natural number less than or equal to 9$\}$
   Describe each of the sets by complete listing method:
   a. A'.
   b. $A \cap A'$.
   c. $A \cup A'$.
   d. $(A')'$
   e. $\phi - U$.
   f. ϕ'
   g. U'

4. Suppose $U =$ The set of one digit numbers and
   $A =\{x: x$ is an even natural number less than or equal to 9$\}$
   Describe each of the sets by complete listing method:
   h. A'.
   i. $A \cap A'$.
   j. $A \cup A'$.
   k. $(A')'$
   l. $\phi - U$.
   m. ϕ'
   n. U'
5. Use Venn diagram to illustrate the following statements:
   a. $(A \cup B)' = A' \cap B'$.
   b. $(A \cap B)' = A' \cup B'$.
   c. If $A \nsubseteq B$, then $A \backslash B \neq \phi$.
   d. $A \cup A' = U$.
6. Let $A = \{5,7,8,9\}$ and $C = \{6,7,8\}$. Then show that $(A\backslash B)\backslash c = A(B\backslash C)$.
7. Perform each of the following operations.
   a. $\phi \cap \{\phi\}$
   b. $\{\phi, \{\phi\}\} - \{\{\phi\}\}$
   c. $\{\phi, \{\phi\}\} - \{\phi\}$
   d. $\{\{\{\phi\}\}\} - \phi$
8. Let $U = \{2, 3, 6, 8, 9, 11, 13, 15\}$,
   $A = \{x | x$ is a positive prime factor of 66$\}$
   $B =\{ x \in U | x$ is composite number $\}$ and $C = \{x \in U | x - 5 \in U\}$. Then find each of the following.
   $A \cap B, (A \cup B) \cap C, (A - B) \cup C, (A - B) - C, A - (B - C), (A - C) - (B - A), A' \cap B' \cap C'$
9. Let $A \cup B = \{a, b, c, d, e, x, y, z\}$ and $A \cap B = \{b, e, y\}$.
   a. $If\ B - A = \{x, z\}$, then $A =$ ___________________
   b. $If\ A - B = \phi$, then $B =$ ____________________
   c. $If\ B = \{b, e, y, z\}$, then $A - B =$ ________________
10. Let $U = \{1, 2, \ldots, 10\}, A = \{3, 5, 6, 8, 10\}, B = \{1, 2, 4, 5, 8, 9\}$,
    $C = \{1, 2, 3, 4, 5, 6, 8\}$ and $D = \{2, 3, 5, 7, 8, 9\}$. Verify each of the following.
    a. $(A \cup B) \cup C = A \cup (B \cup C)$.
    b. $A \cap (B \cup C \cup D) = (A \cap B) \cup (A \cap C) \cup (A \cap D)$.
    c. $(A \cap B \cap C \cap D)' = A' \cup B' \cup C' \cup D'$.
    d. $C - D = C \cap D'$.
    e. $A \cap (B \cap C)' = (A - B) \cup (A - C)$.

11. Depending on question No. 10 find.

a. $A \Delta B$.

b. $C \Delta D$.

c. $(A \Delta C) \Delta D$.

d. $(A \cup B) \setminus (A \Delta B)$.

12. For any two subsets A and B of a universal set U, prove that:

a. $A \Delta B = B \Delta A$.

b. $A \Delta B = (A \cup B) - (A \cap B)$.

c. $A \Delta \phi = A$.

d. $A \Delta A = \phi$.

13. Draw an appropriate Venn diagram to depict each of the following sets.

a. U = The set of high school students in Addis Ababa.

A = The set of female high school students in Addis Ababa.

B = The set of high school anti-AIDS club member students in Addis Ababa.

C = The set of high school Nature Club member students in Addis Ababa.

b. U = The set of integers.

A = The set of even integers.

B = The set of odd integers.

C = The set of multiples of 3.

D = The set of prime numbers.

Chapter Two
Functions

Our everyday lives are filled with situations in which we encounter relationships between two sets. For example,

- To each automobile, there corresponds a license plate number
- To each circle, there corresponds a circumference
- To each number, there corresponds its square

In order to apply mathematics to a variety of disciplines, we must make the idea of a "relationship" between two sets mathematically precise.

On completion of this chapter students will be able to:

- understand the concept of real numbers
- use properties of real numbers to solve problems
- determine whether a given real number is rational number or not
- solve linear equations and inequalities
- solve quadratic equations and inequalities
- understand the notion of relation and function
- determine the domain and range of relations and functions
- find the inverse of a relation
- define polynomial and rational functions
- perform the fundamental operations on polynomials
- find the inverse of an invertible function
- apply the theorems on polynomials to find the zeros of polynomial functions
- apply theorems on polynomials to solve related problems
- sketch and analyze the graphs of rational functions
- define exponential, logarithmic, and trigonometric functions
- sketch the graph of exponential, logarithmic, and trigonometric functions
- use basic properties of logarithmic, exponential and trigonometric functions to solve problems

In this chapter, before discussing the idea of relations and functions we first review the system of real numbers, linear and quadratic equations and inequalities.

1.1 **The real number systems**

At the end of this section, students will be able to:

- understand the concept of real numbers
- use properties of real numbers to solve problems
- determine whether a given real number is rational number or not

In this section we will define what the real numbers are and what are their properties? To answer, we start with some simpler number systems.

- **The integers and the rational numbers**

The simplest numbers of all are the natural numbers,

$$1, 2, 3, 4, 5, 6, \cdots$$

With them we can count: our books, our friends, and our money. If we adjoin their negatives and zero, we obtain the integers;

$$\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots$$

When we try to measure length, weight, or voltage, the integers are inadequate. They are spaced too far apart to give sufficient precision. Thus, we are led to consider quotients (ratios) of integers, numbers such as:

$$\frac{3}{4}, \frac{-7}{8}, \frac{21}{5}, \frac{19}{-2}, \frac{16}{2} \text{ and } \frac{-17}{1}$$

Note that we included $\frac{16}{2}$ and $\frac{-17}{1}$, though we would normally write them as 8 and – 17, since they are equal to the latter by the ordinary meaning of division. We did not include $\frac{5}{0}$ or $\frac{-9}{0}$, since it is impossible to make sense out of these symbols. In fact, let us agree once and for all to banish division by zero from this section. Numbers which can be written in the form $\frac{m}{n}$, where m and n are integers with $n \neq 0$, are called rational numbers.

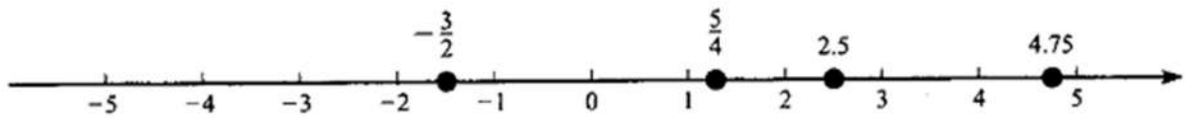
Do the rational numbers serve to measure all lengths? No. This surprising fact was discovered by the ancient Greeks long ago. They showed that while $\sqrt{2}$ measures the hypotenuse of a right triangle with sides of length 1, it cannot be written as a quotient of two integers(see exercise…). Thus, Thus, $\sqrt{2}$ is an irrational (not rational) number. So are $\sqrt{3}, \sqrt{5}, \sqrt[3]{7}, \pi$ and a host of other numbers.

- **The real numbers**

Consider the set of all numbers (rational and irrational) that can measure lengths, together with their negatives and zero. We call these numbers the real numbers.

The set of real numbers denoted by $\Re$ can be described as the union of the set of rational and irrational numbers. i.e $\Re = \{x : x \text{ is a rational number or an irrational number}\}$.

The real numbers may be viewed as labels for points along a horizontal line. There they measure the distance to the right or left (the directed distance) from a fixed point called the origin and labeled 0. Each point on the number line corresponds a unique real number and vice-versa.



Most students will remember that the number system can be enlarged still more to the so-called complex numbers. These are numbers of the form $a + b\sqrt{-1}$, where a and b are real numbers.

- **The four arithmetic operations**

Give two real numbers x and y, we may add or multiply them to obtain two new real numbers $x + y$ and $x \cdot y$ (also written simply as xy). The real numbers along with the operations of addition (+) and multiplication $(\cdot)$, obey the 11 properties listed below. Most of these properties are straightforward and may seem trivial. Nevertheless, we shall see that these 11 basic properties are quite powerful in that they are the basis for simplifying algebraic expressions.

The commutative Properties

1. *For addition*: $a + b = b + a$
2. *For multiplication*: $ab = ba$

The associative properties

3. *For addition*: $a + (b + c) = (a + b) + c$
4. *For multiplication*: $a(bc) = (ab)c$

The distributive property

5. $a(b + c) = ab + ac$ or $(b + c)a = ba + ca$

Identities

6. *For addition*: There is a unique number called the additive identity, represented by 0, which has the property that $a+0=a=0+a$ for all real numbers a.
7. *For multiplication*: There is a unique real number called the multiplicative identity, represented by 1, which has the property that $a\cdot 1=a=1\cdot a$ for all real numbers a.

Inverses

8. *For addition*: Each real number a has a unique additive inverse, represented by $-a$, which has the property that $a+(-a)=0=(-a)+a$
9. *For multiplication*: Each real number a, except 0, has a unique multiplicative inverse, represented by $\frac{1}{a}$, which has the property that $a\cdot(\frac{1}{a})=1=(\frac{1}{a})a$.

Closure properties

10. *For addition*: The sum of two real numbers is a real number.
11. *For multiplication*: The product of two real numbers is a real number.

Subtraction and division are defined by:

$x-y=x+(-y)$ and $x\div y=x\cdot\frac{1}{y}$, where $y\neq 0$.

In the product ab, a and b are called factors, in the sum $a+b$, a and b are called terms.

Example 2.1: The set of irrational numbers is not closed under addition and multiplication, because $\sqrt{2}+(-\sqrt{2})=0$ and $\sqrt{2}\sqrt{8}=\sqrt{16}=4$, which are rational numbers.

- **The order relation on the set of real numbers**

The nonzero real numbers separate nicely into two disjoint sets – the positive real numbers and the negative real numbers. This fact allows us to introduce the order relation < (read "is less than") by

$$x<y \Leftrightarrow y-x \text{ is positive}$$

We agree that $x<y$ and $y>x$ will mean the same thing. The order relation ≤ (read"is less than or equal to") is a first cousin of <. It is defined by

$$x\leq y \Leftrightarrow y-x \text{ is positive or zero}$$

The order relation $<$ has the following properties:

The order property

1. **Trichotomy**: If x and y are numbers, exactly one of the following holds:

$$x < y \text{ or } x = y \text{ or } x > y$$

2. **Transitivity**: $x < y$ and $y < z \Rightarrow x < z$
3. **Addition**: $< y \Leftrightarrow x + z < y + z$
4. **Multiplication**: When z is positive, $x < y \Leftrightarrow xz < yz$,

   When z is negative, $x < y \Leftrightarrow xz > yz$

- **Intervals**

Let a and b be two real numbers such that $a<b$, then the intervals which are subsets of **R** with end points a and b are denoted and defined as below:

i) $(a,b)=\{x : a < x < b\}$ open interval from a to b.

ii) $[a,b]=\{x : a \leq x \leq b\}$ closed interval from a to b.

iii) $(a,b]=\{x : a < x \leq b\}$ open-closed interval from a to b.

iv) $[a,b)=\{x : a \leq x < b\}$ closed-open interval from a to b.

Exercise 2.1

1. Simplify as much as possible:

   a) $4-3(8-12)-6$ c) $\frac{5}{6}-(\frac{1}{4}+\frac{2}{3})$

   b) $2[3-2(4-8)]$ d) $\dfrac{\frac{1}{2}-\frac{3}{4}+\frac{7}{8}}{\frac{1}{2}+\frac{3}{4}-\frac{7}{8}}$

2. Which of the following statements are true and which of them are false?

   a) The sum of any two rational numbers is rational.

   b) The sum of any two irrational numbers is irrational.

   c) The product of any two rational numbers is rational.

   d) The product of any two irrational numbers is irrational.

3. Find the value of each of the following, if undefined, say so.

   a) $0\cdot 0$ c) $\frac{0}{0}$ e) $\frac{0}{8}$

   b) $\frac{8}{0}$ d) 8^0 f) 0^8

4. Show that division by 0 is meaningless as follows: Suppose $a \neq 0$. If $\frac{a}{0}=b$, then $a = 0\cdot b = 0$, which is a contradiction. Now find a reason why $\frac{0}{0}$ is also meaningless.

5. Prove each if $a > 0$, $b > 0$

   a) $a < b \Leftrightarrow a^2 < b^2$ b) $a < b \Leftrightarrow \frac{1}{a} > \frac{1}{b}$

6. Which of the following are always correct if $a \leq b$?

   a) $a-4 \leq b-4$ b) $-a \leq -b$ c) $a^2 \leq ab$ d) $a^2 \leq a^2 b$

2.2 **Equations and Inequalities: Linear and Quadratic**

At the end of this section, students will be able to:

- solve linear equations and inequalities
- solve quadratic equations and inequalities identify the notions of the common sets of numbers

- **Linear Equations and inequalities**

An equation is a symbolic statement of equality. That is, rather than writing "twice a number is four less than the number," we write $2x = x - 4$. Our goal is to find the solution to a given equation. By solution we mean the value or values of the variable that make the algebraic statement true.

Definition 2.1: (Linear Equation) A linear equation in one variable is an equation that can be put in the form $ax + b = 0$, where a and b are constants, and $a \neq 0$.

Equations that have the same solutions are called equivalent equations. For example, $3x - 1 = 5$ and $3x = 6$ are equivalent equations because the solution set of both equations is $\{2\}$. Our goal here is to take an equation and with the help of a few properties, gradually, change the given equation into an equivalent equation of the form $x = a$, where x is the variable for which we are solving. These properties are:

1. **The addition property**
   If $a = b$, then $a + c = b + c$. That is, adding the same quantity to both sides of an equation will produce an equivalent equation.

2. **The multiplication property**
   If $a = b$, then $ac = bc$. That is, multiplying both sides of an equation by the same nonzero quantity will produce an equivalent equation.

Example 2.2:

1. Solve for x
   a) $820x = 10x + 30(50 - x)$ b) $3(2x + 1) = 2(1 - 5x) + 6x + 11$

Solution:

a) $820x = 10x + 30(50 - x)$ Simplify the right hand side
   $820x = 10x + 1500 - 30x$

$820x = 1500 - 20x$ Applying the addition property (add $20x$ to both sides)

$840x = 1500$

Thus, $x = \frac{1500}{840} = \frac{25}{14}$.

Remember to check by substituting $\frac{25}{14}$ for x in the original equation.

b) $3(2x+1) = 2(1-5x) + 6x + 11$ (The given equation)

$6x + 3 = 2 - 10x + 6x + 11$ (Removing parentheses by distribution)

$6x + 10x - 6x = 2 + 11 - 3$ (Collecting like terms: 'variables to the left and numbers to the right')

$10x = 10$

$x = 1$ (Dividing both sides by 10)

Therefore, the solution set (S.S) is $\{1\}$.

2. Find the solution set of $\frac{8x+3}{2} - 5(x+2) = -3(x + \frac{5}{6})$

Solution: $\frac{8x+3}{2} - 5(x+2) = -3(x + \frac{5}{6})$ (The given equation)

This gives us:

$$4x + \frac{3}{2} - 5x - 10 = -3x - \frac{5}{2}$$

$$4x - 5x + 3x = -\frac{5}{2} - \frac{3}{2} + 10$$ Using addition property

$$2x = 6$$

Hence, $x = 3$. That is, the solution set is $\{3\}$.

3. A computer discount store held an end of summer sale on two types of computers. They collected Birr 41,800 on the sale of 58 computers. If one type sold for Birr 600 and the other type sold for Birr 850, how many of each type were sold?

Solution: If we let x to be the number of Birr 600 computers sold, then $58 - x$ = the number of computers that are sold for Birr 850 (since 58 were sold all together).

Our equation involves the amount of money collected on the sale of each type of computer that is, the value of computers sold). Thus we have:

$$600x + 850(58 - x) = 41{,}800, \text{ which yields}$$

$$x = 30$$

Hence, there were 30 computers sold at Birr 600 and 28 computers sold at 850.

Remark: The solution set of some equation can be the set of all rational numbers. This is the case when the equation is satisfied by every rational number.

Example 2.3: Find the solution set of $5x-2(x-1)+4=3(x+2)$

Solution: $5x-2(x-1)+4=3(x+2)$ (The given equation)

$5x-2x+2+4=3x+6$ (Removing parentheses by distribution)

$3x+6=3x+6$ (Combining like terms)

This is always true whatever the value of x is. In fact, subtracting $3x$ from both sides of the last equation we get 6=6 which is always true. This means the given equation is satisfied if you take any number for x as you wish. Thus, S.S = $\Re$.

Remark: There are also some equations which cannot be satisfied by any number. For example, the equation $x+10=x$ says 'If you increase a number x by 10, the result is x itself (unchanged)'. Obviously, there is no such a number. The solution set of such equation is **empty set.** If you try to solve such equation, you end up with a false statement (false equality). For example, an attempt to solve $x+10=x$ leads to the following:

$10+x-x=x-x$ (Subtracting x from both sides of the equation)

$10=0$, which is false.

Hence, the solution set of $x+10=x$ is $\varnothing$ (empty set).

Example 2.4: Find the solution set of $6+3(1-x)=2(1-5x)+7x$

Solution: $6+3(1-x)=2(1-5x)+7x$ (The given equation)

$6+3-3x=2-10x+7x$ (Removing parentheses by distribution)

$9-3x=2-3x$ (Combining like terms)

$9-3x+3x=2-3x+3x$ (Adding $3x$ to both sides)

$9=2$, which is false.

This means the solution set of the given equation is empty, $\varnothing$.

Example 2.5: A man has a daughter and a son. The man is five times older than his daughter. Moreover, his age is twice of the sum of the ages of his daughter and son. His daughter is 3 years younger than his son. How old is the man and his children?

Solution: The unknowns in the problem are age of the man, age of his daughter, and age of his son. So, let m = *Age of the man;* d = *Age of the daughter;* and s = *Age of the son.* Then, 'The man is 5 times older than his daughter' means $m=5d$. Moreover, 'Age of the man is twice the sum of the ages of his daughter and son' means $m=2(d+s)$. 'His daughter is 3 years younger than his son' means $d=s-3$.

Now, from the last (3rd) equation you can get $s = d + 3$. Substitute this in the 2nd equation to get m=2(d +d+3) = 2(2d+3). Thais is, m=4d+6. Next substitute this in the 1st equation to get 4d+6 = 5d or 6 = 5d–4d=d. Hence, d= 6. From this, $s = d$ +3 = 6+3 = 9, and m=5d =5×6= 30. Therefore, the age of the man is 30, age of his daughter is 6 and age of his son is 9.

Definition 2.2: (Linear Inequalities)

A linear inequality is an inequality that can be put in the form $ax + b < 0$, where a and b are constants with $a \neq 0$. (The $<$ symbol can be replaced with $>, \leq$ or $\geq$)

To solve inequalities, we will need the following properties of inequalities.

For $a, b, c \in \Re$, if $a < b$, then
1) $a + c < b + c$ 2) $ac < bc$, when $c > 0$ 3) $ac > bc$, when $c < 0$

Thus, to produce an equivalent inequality, we may add (subtract) the same quantity to (from) both sides of an inequality, or multiply (divide) both sides by the same positive quantity. On the other hand, we must reverse the inequality symbol to produce an equivalent inequality if we multiply (divide) both sides by the same negative quantity.

Example 2.6:
1. Solve the linear inequality $5x + 8(20 - x) \geq 2(x - 5)$.

Solution:

$5x + 8(20 - x) \geq 2(x - 5)$	Simplify each side
$5x + 160 - 8x \geq 2x - 10$	
$160 - 3x \geq 2x - 10$	Now apply the inequality property
$-5x \geq -170$	Divide both sides by – 5
$x \leq 34$	Note that the inequality symbol is reversed

Thus, the solution set is $\{x \in \Re : x \leq 34\} = (-\infty, 34]$.

Example 2.7: Find the solution set of the inequality 3x –5(x+2) ≥ 0.

Solution:

3x –5(x –2) > 0	(The given inequality)
3x –5x + 10 > 0	(Removing the parentheses by distribution)
–2x + 10 > 0	(Combining like terms)
–2x > –10	(Subtracting 10 from both sides)
$x < \frac{-10}{-2}$	(Dividing both sides by –2 reverse the inequality)

That is, $x < 5$. Therefore, S.S = {x: $x < 5$}, the set of all real numbers less 5.
The solution of an inequality is sometimes required to be only in a given **domain** (set). If so, a solution set should contain only those solutions that belong to the specified domain.

Example 2.8: Find the solution set of $x-4(x+1)\geq -13-(x-2)$ in the set of natural numbers, ℕ.

Solution: $x-4(x+1)\geq -13-(x-2)$ (The given inequality)

$x-4x-4\geq -13-x+2$ (Removing parentheses by distribution)

$-3x-4\geq -11-x$ (Combining like terms; i.e., $x-4x=-3x$ and $-13+2=-11$)

$-3x+x\geq -11+4$ (Collecting like terms)

$-2x\geq -7$ (Next, division of both sides of this by –2 reverses the inequality)

$x\leq \frac{7}{2}$; i.e., $x\leq 3.5$

Thus, the solution of the given inequality in ℕ is {1, 2, 3}. (Recall: ℕ = {1, 2, 3, … })

Some inequalities may have no solution in the specified domain as in the following example.

Example 2.9: Find the solution set of $7x+6\leq 3x+2$ in the set of **whole numbers**, **W**.

Solution: $7x-2\leq 3x-6$ (The given inequality)

$7x-3x\leq -6+2$ (Collecting like terms)

$4x\leq -4$

$\frac{\not{4}x}{\not{4}}\leq \frac{-4}{4}$ or $x\leq -1$

However, there is no negative whole number. Therefore, the solution set of the given inequality in **W** is ∅, empty set. (Recall: **W** = {0, 1, 2, 3, … } **)**

Example 2.10: Find the solution set of the inequality $\frac{1}{6}(x+3)+\frac{1}{2}x-\frac{3}{2}\ \leq\ \frac{3}{2}(x+1)$ in ℚ.

Solution: The inequality involves fractional numbers. Thus, like for the case of linear equations, clear the denominators by multiplying both sides of the inequality by the LCM of the denominators. The denominators in this equation are 6 and 2; and their LCM is 6. Thus, multiply every term in both sides of the given inequality by 6. That is,

$6\left[\frac{1}{6}(x+3)\right]+6\left(\frac{1}{2}x\right)-6\left(\frac{3}{2}\right)\ \leq\ 6\left[\frac{3}{2}(x+1)\right]$ (The inequality is not reversed because 6>0)

$x+3+3x-9\ \leq\ 9(x+1)$ (Simplifying/clear denominators)

$4x-6\ \leq\ 9x+9$

$4x-9x\ \leq\ 9+6$ (Collecting like terms)

$-5x\leq 15$ (Next, division of both sides by -5)

$x\geq \frac{15}{-5}$ or $x\geq -3$.

Therefore, S.S = $\{\, x\in\mathbb{Q} \mid x\geq -3\,\}$.

- **Quadratic Equations and Inequalities**

A quadratic equation is a polynomial equation in which the highest degree of the variable is 2. We define the standard form of a quadratic a quadratic equation as $Ax^2 + Bx + c = 0$, where $A \neq 0$.
As with linear equations, the solutions of quadratic equations are values of the variable that make the equation a true statement. The solutions of $Ax^2 + Bx + C = 0$ are also called the roots of the polynomial equation $Ax^2 + Bx + C = 0$.

In solving the equation $Ax^2 + Bx + C = 0$, if the polynomial $Ax^2 + Bx + C$ can be factored, the we can use the zero product rule (which is stated below) to reduce the problem to that of solving two linear equations. For example, to solve the equation $x^2 + x - 6 = 0$, we van factor the left hand side to get $(x-2)(x+3) = 0$. Hence, we can conclude that $x - 2 = 0$ or $x + 3 = 0$, which yields $x = 2$ or $x = -3$.

The Zero-Product Rule: If $a \cdot b = 0$, then $a = 0$ or $b = 0$

Another method is to apply the Square Root Theorem.

The Square Root Theorem: If $x^2 = d$, then $x = \pm\sqrt{d}$.

Example 2.11: Solve the following

a) $4x^2 + 10x = 6$ b) $5x^2 - 6 = 8$ c) $(x-2)^2 = 6$

Solution: a)

$4x^2 + 10x = 6$	Put into standard form
$4x^2 + 10x - 6 = 0$	Factor the left hand side
$2(2x-1)(x+3) = 0$	Hence we have
$2x - 1 = 0$ or $x + 3 = 0$	Solving each linear equation, we get
$x = \frac{1}{2}$ or $x = -3$	

b) We note that there is no first-degree term, so our approach will be to apply the Square Root Theorem.

$5x^2 - 6 = 8$	Isolate x^2 on the left-hand side before applying the square root theorem
$5x^2 = 14$	
$x^2 = \frac{14}{5}$	Applying the square root theorem we get
$x = \pm\sqrt{\frac{14}{5}}$	

c) Since it is in the form of a squared quantity equal to a number, we will apply the Square Root Theorem to get $x = 2 \pm \sqrt{6}$.

Part (c) of the above example illustrates that if we can construct a perfect square binomial from a quadratic equation (i.e., get the equation in the form $(x+p)^2 = d$), then we can apply the Square Root Theorem and solve for x to get $x = -p \pm \sqrt{d}$.

The method of constructing a perfect square is called completing the square. It is based on the fact that in multiplying out the perfect square $(x+p)^2$, with p a constant, we get

$$(x+p)^2 = x^2 + 2px + p^2$$

Notice the relationship between the constant term, p^2, and the coefficient of the middle term, $2p$: The constant term is the square of half the coefficient of the middle term.

Example 2.12: Solve by completing the square: $2x^2 - 8x + 4 = 6$.

Solution:

$2x^2 - 8x + 4 = 6$	Divide both sides by 2, the coefficient of x^2
$x^2 - 4x + 2 = 3$	Isolate the constant term on the right-hand side
$x^2 - 4x = 1$	Take half the middle term coefficient, square it $\left(\frac{1}{2}(-4)\right)^2 = 4$, we add 4 to both sides of the equation
$x^2 - 4x + 4 = 1 + 4$	Factor the left hand side
$(x-2)^2 = 5$	Solve for x using the Square Root Theorem
$x = 2 \pm \sqrt{5}$.	

Unlike the factoring method, all quadratic equations can be solved by completing the square. If we were to complete the square for the general quadratic equation $Ax^2 + Bx + C = 0, A \neq 0$, we would arrive at the formula given below.

The Quadratic Formula: If $Ax^2 + Bx + C = 0$ and $A \neq 0$, then $x = \dfrac{-B \pm \sqrt{B^2 - 4AC}}{2A}$

Example 2.13: Solve the following using the quadratic formula: $x^2 - 8 = -6x$.

Solution: Writing the equation in standard form we get, $x^2 + 6x - 8 = 0$. By the quadratic formula we have:

$$x = \frac{-6 \pm \sqrt{6^2 - 4(1)(-8)}}{2(1)} = \frac{-6 \pm \sqrt{68}}{2} = \frac{-6 \pm 2\sqrt{17}}{2} = -3 \pm \sqrt{17}$$

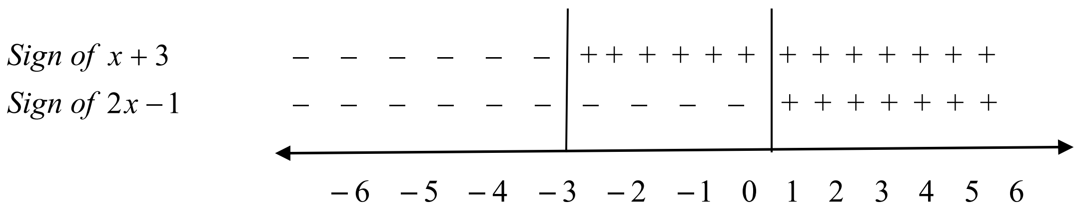
Thus, the solution set is $\{-3 - \sqrt{17}, -3 + \sqrt{17}\}$.

A **quadratic inequality** is in standard form if it is in the form $Ax^2 + Bx + C < 0$. (We can replace $<$ with $>, \le,$ or $\ge$.)

If we keep in mind that $u > 0$ means u is positive, then solving an inequality such as $2x^2 + 5x - 3 > 0$ means we are interested in finding the values of x that will make $2x^2 + 5x - 3$ positive. Or, since $2x^2 + 5x - 3 = (2x-1)(x+3)$, we are looking for values of x that make $(2x-1)(x+3)$ positive. For $(2x-1)(x+3)$ to be positive, the factors must be either both positive or both negative. To determine when this happens, we first find the values of x for which $(2x-1)(x+3)$ is equal to 0; we call these the cut points of $(2x-1)(x+3)$. The cut points are $\frac{1}{2}$ and -3.

Thus, our approach in solving quadratic inequalities will be primarily algebraic. After putting the inequality in standard form, we will determine the sign of each factor of the expression for various values of x. Then, we determine the solution by examining the sign of the product. This process is called a **sign analysis**.

Returning to the problem $2x^2 + 5x - 3 > 0$, we draw a number line and examine the sign of each factor as x takes on various values on the number line, especially around the cut points.



The above figure illustrates that the factor $x+3$ is negative when $x < -3$ and positive when $x > -3$. It is also shown that $2x-1$ is negative when $x < \frac{1}{2}$ and positive when $x > \frac{1}{2}$. Thus the product of the two factors is positive when $x < -3$ and $x > \frac{1}{2}$. Therefore, the solution set is $(-\infty, -3) \cup (\frac{1}{2}, \infty)$.

Remark: 1. The cut points of the inequalities will break up the number line into intervals.

2. The sign of the product does not change within an interval, i.e., if the expression is positive (or negative) for one value within the interval, it is positive (or negative) for all values within the interval.

Example 2.14: Solve the quadratic inequality $x^2 - 2x - 2 < 0$.

Solution: Since we cannot factor x^2-2x-2, we use the quadratic formula to find that its roots are $1\pm\sqrt{3}$. This gives the cut points for the polynomial x^2-2x-2. We use the sign analysis (see the figure below) with the test points given. Note: $1+\sqrt{3}\approx 2.7$ & $1\sqrt{3}\approx -0.7$.

Sign of x^2-2x-2	$+$	0	$-$	0	$+$
	$x=-10$	$1-\sqrt{3}$	$x=1$	$1+\sqrt{3}$	$x=100$

Substituting the test values – 10, 1, and 100 for x in the expression x^2-2x-2, we find that x^2-2x-2 is negative only when x is in the interval $(1-\sqrt{3},1+\sqrt{3})$.

Exercise 2.2

1. Solve the linear equations

   a) $2-3(x-4)=2(x-1)$ d) $\dfrac{2}{x+3}+4=\dfrac{5-x}{x+3}$

   b) $3x-[2+3(2-x)]=5-(3-x)$ e) $\dfrac{6}{x^2-3x}=\dfrac{12}{x}+\dfrac{1}{x-3}$

   c) $\frac{3}{4}(2x-3)=\frac{2}{3}x+5$

2. Solve the linear inequalities

   a) $4x+\frac{2}{3}\le 2x-(3x+1)$ b) $5x-2>3x-(x-\frac{1}{5})$ c) $\dfrac{5x-2}{3}\ge\dfrac{x+3}{4}$

3. A truck carries a load of 50 boxes; some are 20 kg boxes and the rest are 25 kg boxes. If the total weight of all boxes is 1175 kg, how many of each type are there?
4. The product of two numbers is 5. If their sum $\frac{9}{2}$, find the numbers.
5. Solve

   a) $2x^2-7x=15$ c) $x^2+2x-4=0$ e) $3x^2-6x+5=0$

   b) $x-3=\dfrac{1}{x+3}$ d) $\dfrac{1}{x-5}+\dfrac{3}{x+2}=4$

6. Solve the quadratic inequalities

   a) $x^2+2x-24>0$ d) $2x^2-x-2\ge 0$

   b) $x^2-5x\le 24$ e) $x^2\le 16$

   c) $x^2-3x-3<0$

7. A student was given the inequality: $\dfrac{3}{x-2}>4$. The first step the student took in solving this inequality was to transform it into $3>4(x-2)$. Explain what the student did wrong.

2.3. **Review of relations and functions**

After completing this section, the student should be able to:

- define Cartesian product of two sets
- understand the notion of relation and function
- know the difference between relation and function
- determine the domain and range of relations and functions
- find the inverse of a relation

The student is familiar with the phrase ordered pair. In the ordered pair $(2,3),(-2,4)$ and (a,b); $2,-2$ and a are the first coordinates while $3,4$ and b are the second coordinates.

- **Cartesian Product**

Given sets $A=\{3,4\}$ and $B=\{4,5,9\}$. Then, the set $\{(3,4),(3,5),(3,9),(4,4),(4,5),(4,9)\}$ is the Cartesian product of A and B, and it is denoted by $A\times B$.

Definition 2.3: Suppose A and B are sets. The Cartesian product of A and B, denoted by $A\times B$, is the set which contains every ordered pair whose first coordinate is an element of A and second coordinate is an element of B, i.e.

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

Example 2.15: For $A=\{2,4\}$ and $A=\{-1,3\}$, we have

a) $A\times B=\{(2,-1),(2,3),(4,-1),(4,3)\}$, and
b) $B\times A=\{(-1,2),(-1,4),(3,2),(3,4)\}$.

From this example, we can see that $A\times B$ and $B\times A$ are not equal. Recall that two sets are equal if one is a subset of the other and vice versa. To check equality of Cartesian products we need to define equality of ordered pairs.

Definition 2.4: (Equality of ordered Pairs)
Two ordered pairs (a,b) and (c,d) are equal if and only if $a=c$ and $b=d$.

Example 2.16: Let $A=\{1,2,3\}$ and $B=\{a,b,c\}$. Then,

$$A\times B=\{(1,a),(1,b),(1,c),(2,a),(2,b),(2,c),(3,a),(3,b),(3,c)\}.$$

Definition 2.5: (Relation)
If A and B are sets, any subset of $A\times B$ is called a relation from A into B.

Suppose R is a relation from a set A to a set B. Then, $R \subseteq A \times B$ and hence for each $(a,b) \in A \times B$, we have either $(a,b) \in R$ or $(a,b) \notin R$. If $(a,b) \in R$, we say "a is R-related (or simply related) to b", and write aRb. If $(a,b) \notin R$, we say that "a is not related to b".
In particular if R is a relation from a set A to itself, then we say that R is a relation on A.

Example 2.17:

1. Let $A = \{1,3,5,7\}$ and $B = \{6,8\}$. Let R be the relation "less than" from A to B. Then, $R = \{(1,6),(1,8),((3,6),(3,8),(5,6),(5,8),(7,8)\}$.
2. Let $A = \{1,2,3,4,5\}$ and $B = \{a,b,c\}$.
   a) The following are relations from A into B;
      i) $R_1 = \{(1,a)\}$
      ii) $R_2 = \{(2,b),(3,b),(4,c),(5,a)\}$
      iii) $R_3 = \{(1,a),((2,b),(3,c)\}$
   b) The following are relations from B to A;
      i) $R_4 = \{(a,3),(b,1)\}$
      ii) $R_5 = \{(b,2),(c,4),(a,2),(b,3)\}$
      iii) $R_6 = \{(b,5)\}$

Definition 2.6: Let R be a relation from A into B. Then,

a) the domain of R, denoted by $Dom(R)$, is the set of first coordinates of the elements of R, i.e
$$Dom(R) = \{a \in A : (a,b) \in R\}$$
b) the range of R, denoted by $Range(R)$, is the set of second coordinates of elements of R, i.e
$$Range(R) = \{b \in B : (a,b) \in R\}$$

Remark: If R is a relation form the set A to the set B, then the set B is called the codomain of the relation R. The range of relation is always a subset of the codomain.

Example 2.18:

1. The set $R = \{(4,7),(5,8),(6,10)\}$ is a relation from the set $A = \{1,2,3,4,5,6\}$ to the set $B = \{6,7,8,9,10]$. The domain of R is $\{4,5,6\}$, the range of R is $\{7,8,10\}$ and the codomain of R is $\{6,7,8,9,10\}$.
2. The set of ordered pairs $R = \{(8,2),(6,-3),(5,7),(5,-3)\}$ is a relation between the sets $\{5,6,8\}$ and $\{2,-3,7\}$, where $\{5,6,7\}$ is the domain and $\{2,-3,7\}$ is the range.

Remark:

1. If $(a,b) \in R$ for a relation R, we say a is related to (or paired with) b. Note that a may also be paired with an element different from b. In any case, b is called the image of a while a is called the pre-image of b.
2. If the domain and/or range of a relation is infinite, we cannot list each element assignment, so instead we use set builder notation to describe the relation. The situation we will encounter most frequently is that of a relation defined by an equation or formula. For example,

   $$R = \{(x, y) : y = 2x - 3,\ x, y \in IR\}$$

   is a relation for which the range value is 3 less than twice the domain value. Hence, $(0,-3), (0.5,-2)$ and $(-2,-7)$ are examples of ordered pairs that are of the assignment.

Example 2.19:

1. Let $A = \{1, 2, 3, 4, 6\}$

   Let R be the relation on A defined by $R = \{(a,b) : a, b \in A, a \text{ is a factor of } b\}$. Find the domain and range of R.

Solution: We have

$R = \{(1,1),(1,2),(1,3),(1,4),(1,6),(2,2),(2,4),(2,6),(3,3),(3,6),(4,4),(6,6)\}$.

Then, $Dom(R) = \{1,2,3,4,6\}$ and $Range(R) = \{1,2,3,4,6\}$.

2. Let $A = \{1,2,3,4,5\}$ and $B = \{1, 2, 3, \cdots, 67\}$.

   Let $R = \{(x, y) \in A \times B : x \text{ is cube root of } y\}$. Find a) R b) $Dom(R)$ c) $Range(R)$

Solution: We have $1 = \sqrt[3]{1}, 2 = \sqrt[3]{8}, 3 = \sqrt[3]{27}, 4 = \sqrt[3]{64}, 5 = \sqrt[3]{125}$ and 1,8,27 and 64 are in B whereas 125 is not in B. Thus, $R = \{(1,1),(2,8),(3,27),(4,64)\}$, $Dom(R) = \{1,2,3,4\}$ and $R = \{1,8,27,64\}$.

Remark:

1. A relation R on a set A is called
   i) a universal relation if $R = A \times A$
   ii) identity relation if $R = \{(a,a) : a \in A\}$
   iii) void or empty relation if $R = \phi$
2. If R is a relation from A to B, then the inverse relation of R, denoted by R^{-1}, is a relation from B to A and is defined as:

   $$R^{-1} = \{(y,x) : (x,y) \in R\}.$$

Observe that $Dom(R) = Range(R^{-1})$ and $Range(R) = Dom(R^{-1})$. For instance, if $R = \{(1,4),(9,15),(10,2)\}$ is a relation on a set $A = \{1,2,3,\cdots,20\}$, then $R^{-1} = \{(4,1),(15,9),(2,10)\}$

Example 2.20: Let R be a relation defined on IN by $R = \{(a,b) : a,b \in IN,\ a + 2b = 11\}$.
Find a) R b) $Dom(R)$ c) $Range(R)$ d) R^{-1}

Solution: The smallest natural number is 1.

$$b = 1 \Rightarrow a + 2(1) = 11 \Rightarrow a = 9$$
$$b = 2 \Rightarrow a + 2(2) = 11 \Rightarrow a = 7$$
$$b = 3 \Rightarrow a + 2(3) = 11 \Rightarrow a = 5$$
$$b = 4 \Rightarrow a + 2(4) = 11 \Rightarrow a = 3$$
$$b = 5 \Rightarrow a + 2(5) = 11 \Rightarrow a = 1$$
$$b = 6 \Rightarrow a + 2(6) = 11 \Rightarrow a = -1 \notin IN$$

Therefore, $R = \{(9,1),(7,2),(5,3),(3,4),(1,5)\}$, $Dom(R) = \{1,3,5,7,9\}$, $Range(R) = \{1,2,3,4,5\}$ and $R^{-1} = \{(1,9),(2,7),(3,5),(4,3),(5,1)\}$.

- **Functions**

Mathematically, it is important for us to distinguish among the relations that assign a unique range element to each domain element and those that do not.

Definition 2.7: (Function)
A function is a relation in which each element of the domain corresponds to exactly one element of the range.

Example 2.21: Determine whether the following relations are functions.

a) $R = \{(5,-2),(3,5),(3,7)\}$ b) {(2,4),(3,4),(6,-4)}

Solution:

a) Since the domain element 3 is assigned to two different values in the range, 5 and 7, it is not a function.
b) Each element in the domain, $\{2,3,6\}$, is assigned no more than one value in the range, 2 is assigned only 4, 3 is assigned only 4, and 6 is assigned only – 4. Therefore, it is a function.

Remark: Map or mapping, transformation and correspondence are synonyms for the word function. If f is a function and $(x, y) \in f$, we say *x* is mapped to *y*.

Definition 2.8: A relation f from *A* into *B* is called a function from *A* into *B*, denoted by

$$f : A \to B \quad \text{or} \quad A \xrightarrow{f} B$$

if and only if

(i) $Dom(f) = A$

(ii) No element of *A* is mapped by f to more than one element in *B*, i.e. if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$.

Remark: 1. If to the element *x* of A corresponds $y(\in B)$ under the function f, then we write $f(x) = y$ and y is called the image of *x* under y and *x* is called a pre-image of y under f.

2. The symbol $f(x)$ is read as "f of *x*" but not "f times *x*".

3. In order to show that a relation f from A into B is a function, we first show that the domain of f is *A* and next we show that f well defined or single-valued, i.e. if $x = y$ in *A*, then $f(x) = f(y)$ in *B* for all $x, y \in A$.

Example 2.22:

1. Let $A = \{1,2,3,4\}$ and $B = \{1,6,8,11,15\}$. Which of the following are functions from A to B.
   a) f defined by $f(1) = 1, f(2) = 6, f(3) = 8, f(4) = 8$
   b) f defined by $f(1) = 1, f(2) = 6, f(3) = 15$
   c) f defined by $f(1) = 6, f(2) = 6, f(3) = 6, f(4) = 6$
   d) f defined by $f(1) = 1, f(2) = 6, f(2) = 8, f(3) = 8, f(4) = 11$
   e) f defined by $f(1) = 1, f(2) = 8, f(3) = 11, f(4) = 15$

Solution:

a) f is a function because to each element of *A* there corresponds exactly one element of B.

b) f is not a function because there is no element of *B* which correspond to *4*($\in$*A*).

c) f is a function because to each element of *A* there corresponds exactly one element of *B*. In the given function, the images of all element of *A* are the same.

d) f is not a function because there are two elements of B which are corresponding to 2. In other words, the image of 2 is not unique.

e) f is a function because to each element of A there corresponds exactly one element of B.

As with relations, we can describe a function with an equation. For example, *y=2x+1* is a function, since each *x* will produce only one y.

2. Let $f = \{(x, y): y = x^2\}$. Then, f maps:

1 to 1	-1 to 1
2 to 4	-2 to 4
3 to 9	-3 to 9

More generally any real number x is mapped to its square. As the square of a number is unique, f maps every real number to a unique number. Thus, f is a function from $\Re$ into $\Re$.
We will find it useful to use the following vocabulary: The independent variable refers to the variable representing possible values in the domain, and the dependent variable refers to the variable representing possible values in the range. Thus, in our usual ordered pair notation (x, y), x is the independent variable and y is the dependent variable.

- **Domain, Codomain and range of a function**

For the function $f : A \to B$
(i) The set A is called the domain of f
(ii) The set B is called the codomain of f
(iii) The set $\{f(x): x \in A\}$ of all image of elements of A is called the range of f

Example 2.23:

1. Let $A = \{1,2,3\}$ and $B = \{1,2,3,\cdots,10\}$. Let $f : A \to B$ be the correspondence which assigns to each element in A, its square. Thus, we have $f(1) = 1, f(2) = 4, f(3) = 9$. Therefore, f is a function and $Dom(f) = \{1,2,3\}$, $Range(f) = \{1,4,9\}$ and codomain of f is $\{1,2,3,\cdots,10\}$.

2. Let $A = \{2,4,6,7,9\}, B = IN$. Let x and y represent the elements in the sets A and B, respectively. Let $f : A \to B$ be a function defined by $f(x) = 15x + 17, x \in A$.

   The variable x can take values 2, 4, 6, 7, 9. Thus, we have
   $$f(2) = 15(2) + 17 = 47, f(4) = 77, f(6) = 107, f(7) = 122, f(9) = 152.$$
   This implies that $Dom(f) = \{2,4,6,7,9\}, Range(f) = \{47,77,107,122,152\}$ and codomain of f is IN.

3. Let f be the subset of $Q \times Z$ defined by $f = \left\{\left(\frac{p}{q}, p\right): p, q \in Z, q \neq 0\right\}$. Is f a function?
   Solution: First we note that $Dom(f) = Q$. Then, f satisfies condition (i) in the definition of a function. Now, $\left(\frac{2}{3}, 2\right) \in f$, $\left(\frac{4}{6}, 4\right) \in f$ and $\frac{2}{3} = \frac{4}{6}$ but $f\left(\frac{2}{3}\right) = 2 \neq 4 = f\left(\frac{4}{6}\right)$. Thus f is not well defined. Hence, f is not a function from Q to Z.

4. Let f be the subset of $Z \times Z$ defined by $f = \{(mn, m+n) : m, n \in Z\}$. Is f a function?
Solution: First we show that f satisfies condition (i) in the definition. Let x be any element of Z. Then, $x = x \cdot 1$. Hence, $(x, x+1) = (x \cdot 1, x+1) \in f$. This implies that $x \in Dom(f)$. Thus, $Z \subseteq Dom(f)$. However, $Dom(f) \subseteq Z$ and so $Dom(f) = Z$. Now, $4 \in Z$ and $4 = 4 \cdot 1 = 2 \cdot 2$. Thus, $(4 \cdot 1, 4+1)$ and $(2 \cdot 2, 2+2)$ are in f. Hence we find that $4 \cdot 1 = 2 \cdot 2$ and $f(4 \cdot 1) = 5 \neq 4 = f(2 \cdot 2)$. This implies that f is not well defined, i.e, f does not satisfy condition (ii). Hence, f is not a function from Z to Z.

5. Determine whether the following equations determine y as a function of x, if so, find the domain.

   a) $y = -3x + 5$ b) $y = \dfrac{2x}{3x-5}$ c) $y^2 = x$

Solution:

a) To determine whether $y = -3x + 5$ gives y as a function of x, we need to know whether each *x-value* uniquely determines a *y-value*. Looking at the equation $y = -3x + 5$, we can see that once x is chosen we multiply it by – 3 and then add 5. Thus, for each *x* there is a unique y. Therefore, $y = -3x + 5$ is a function.

b) Looking at the equation $y = \dfrac{2x}{3x-5}$ carefully, we can see that each *x-value* uniquely determines a *y-value* (one *x-value* can not produce two different *y-values*). Therefore, $y = \dfrac{2x}{3x-5}$ is a function.

   As for its domain, we ask ourselves. Are there any values of x that must be excluded? Since $y = \dfrac{2x}{3x-5}$ is a fractional expression, we must exclude any value of x that makes the denominator equal to zero. We must have

$$3x - 5 \neq 0 \quad \Leftrightarrow \quad x \neq \frac{5}{3}$$

   Therefore, the domain consists of all real numbers except for $\frac{5}{3}$. Thus, $Dom(f) = \{x : x \neq \frac{5}{3}\}$.

c) For the equation $y^2 = x$, if we choose $x = 9$ we get $y^2 = 9$, which gives $y = \pm 3$. In other words, there are two $y-$values associated with $x = 9$. Therefore, $y^2 = x$ is not a function.

6. Find the domain of the function $y=\sqrt{3x-x^2}$.

Solution: Since y is defined and real when the expression under the radical is non-negative, we need x to satisfy the inequality

$$3x-x^2 \geq 0 \quad \Leftrightarrow \quad x(3-x)\geq 0$$

This is a quadratic inequality, which can be solved by analyzing signs:

Sign of $3x-x^2$ — — — | + + + | — — —

0 3

Since we want $3x-x^2 = x(3-x)$ to be non-negative, the sign analysis shows us that the domain is $\{x: 0 \leq x \leq 3\}$ or $[0,3]$.

Exercise 2.3

1. Let R be a relation on the set $A=\{1,2,3,4,5,6\}$ defined by $R=\{(a,b): a+b\leq 9\}$.
   i) List the elements of R
   ii) Is $R=R^{-1}$
2. Let R be a relation on the set $A=\{1,2,3,4,5,6,7\}$ defined by $R=\{(a,b): 4$ divides $a-b\}$.
   i) List the elements of R
   ii) Find $Dom(R)$ & $Range(R)$
   iii) Find the elements of R^{-1}
   iv) Find $Dom(R^{-1})$ & $Range(R^{-1})$
3. Let $A=\{1,2,3,4,5,6\}$. Define a relation on A by $R=\{(x,y): y=x+1\}$. Write down the domain, codomain and range of R. Find R^{-1}.
4. Find the domain and range of the relation $\{(x,y): |x|+y\geq 2\}$.
5. Let $A=\{1,2,3\}$ and $B=\{3,5,6,8\}$. Which of the following are functions from A to B?
   a) $f=\{(1,3),(2,3),(3,3)\}$
   b) $f=\{(1,3),(2,5),(1,6)\}$
   c) $f=\{(1,8),(2,5)\}$
   d) $f=\{(1,6),(2,5),(3,3)\}$
6. Determine the domain and range of the given relation. Is the relation a function?
   a) $\{(-4,-3),(2,-5),(4,6),(2,0)\}$
   b) $\{(8,-2),(6,-\frac{3}{2}),(-1,5)\}$
   c) $\{(-\sqrt{3},3),(-1,1),(0,0),(1,1),(\sqrt{3},3)\}$
   d) $\{(-\frac{1}{2},\frac{1}{6}),(-1,1),(\frac{1}{3},\frac{1}{8})\}$
   e) $\{(0,5),(1,5),(2,5),(3,5),(4,5),(5,5)\}$
   f) $\{(5,0),(5,1),(5,2),(5,3),(5,4),(5,5)\}$

7. Find the domain and range of the following functions.

   a) $f(x) = 1 + 8x - 2x^2$ c) $f(x) = \sqrt{x^2 - 6x + 8}$

   b) $f(x) = \dfrac{1}{x^2 - 5x + 6}$ d) $f(x) = \begin{cases} 3x+4, & -1 \le x < 2 \\ 1+x, & 2 \le x \le 5 \end{cases}$

8. Given $f(x) = \begin{cases} 3x-5, & x < 1 \\ x^2 - 1, & x \ge 1 \end{cases}$.

   Find a) $f(-3)$ b) $f(1)$ c) $f(6)$

2.4 **Real Valued functions and their properties**

After completing this section, the student should be able to:

- perform the four fundamental operations on polynomials
- compose functions to get a new function
- determine the domain of the sum, difference, product and quotient of two functions
- define equality of two functions

Let f be a function from set A to set B. If B is a subset of real number system $\Re$, then f is called a real valued function, and in particular if A is also a subset of $\Re$, then $f: A \to B$ is called a real function.

Example 2.24: 1. The function $f: \Re \to \Re$ defined by $f(x) = x^2 + 3x + 7$, $x \in \Re$ is a real function.

2. The function $f: \Re \to \Re$ defined as $f(x) = |x|$ is also a real valued function.

- **Operations on functions**

Functions are not numbers. But just as two numbers a and b can be added to produce a new number $a + b$, so two functions f and g can be added to produce a new function $f + g$. This is just one of the several operations on functions that we will describe in this section.

Consider functions f and g with formulas $f(x) = \dfrac{x-3}{2}$, $g(x) = \sqrt{x}$. We can make a new function $f + g$ by having it assign to x the value $\dfrac{x-3}{2} + \sqrt{x}$, that is,

$$(f+g)(x) = f(x) + g(x) = \frac{x-3}{2} + \sqrt{x} .$$

Definition 2.9: Sum, Difference, Product and Quotient of two functions

Let $f(x)$ and $g(x)$ be two functions. We define the following four functions:

1. $(f+g)(x) = f(x)+g(x)$ The sum of the two functions
2. $(f-g)(x) = f(x)-g(x)$ The difference of the two functions
3. $(f\cdot g)(x) = f(x)g(x)$ The product of the two functions
4. $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$ The quotient of the two functions (provided $g(x)\neq 0$)

Since an $x-$value must be an inout into both f and g, the domain of $(f+g)(x)$ is the set of all x common to the domain of f and g. This is usually written as $Dom(f+g) = Dom(f)\bigcap Dom(g)$. Similar statements hold for the domains of the difference and product of two functions. In the case of the quotient, we must impose the additional restriction that all elements in the domain of g for which $g(x)=0$ are excluded.

Example 2.25:

1. Let $f(x) = 3x^2+2$ and $g(x) = 5x-4$. Find each of the following and its domain

a) $(f+g)(x)$ b) $(f-g)(x)$ c) $(f.g)(x)$ d) $\left(\frac{f}{g}\right)(x)$

Solution:

a) $(f+g)(x) = f(x)+g(x) = (3x^2+2)+(5x-4) = 3x^2+5x-2$

b) $(f-g)(x) = f(x)-g(x) = (3x^2+2)-(5x-4) = 3x^2-5x+6$

c) $(f\cdot g)(x) = (3x^2+2)(5x-4) = 15x^3-12x^2+10x-8$

d) $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} = \frac{3x^2+2}{5x-4}$

We have

$$Dom(f+g) = Dom(f-g) = Dom(fg) = Dom(f)\bigcap Dom(g) = \Re\bigcap\Re = \Re$$

$$Dom\left(\frac{f}{g}\right) = \left[Dom(f)\bigcap Dom(g)\right]\setminus\{x : g(x)=0\} = \Re\setminus\left\{\frac{5}{4}\right\}$$

2. Let $f(x) = \sqrt[4]{x+1}$ and $g(x) = \sqrt{9-x^2}$, with respective domains $[-1,\infty)$ and $[-3,3]$. Find formulas for $f+g, f-g, f\cdot g, \frac{f}{g}$ and f^3 and give their domains.

Solution:

Formula	Domain
$(f+g)(x) = f(x) + g(x) = \sqrt[4]{x+1} + \sqrt{9-x^2}$	$[-1,3]$
$(f-g)(x) = f(x) - g(x) = \sqrt[4]{x+1} - \sqrt{9-x^2}$	$[-1,3]$
$(f \cdot g)(x) = f(x) \cdot g(x) = \sqrt[4]{x+1} \cdot \sqrt{9-x^2}$	$[-1,3]$
$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} = \frac{\sqrt[4]{x+1}}{\sqrt{9-x^2}}$	$[-1,3)$
$f^3(x) = (f(x))^3 = \left(\sqrt[4]{x+1}\right)^3 = (x+1)^{\frac{3}{4}}$	$[-1,\infty)$

There is yet another way of producing a new function from two given functions.

Definition 2.10: (Composition of functions)

Given two functions $f(x)$ and $g(x)$, the composition of the two functions is denoted by $f \circ g$ and is defined by:

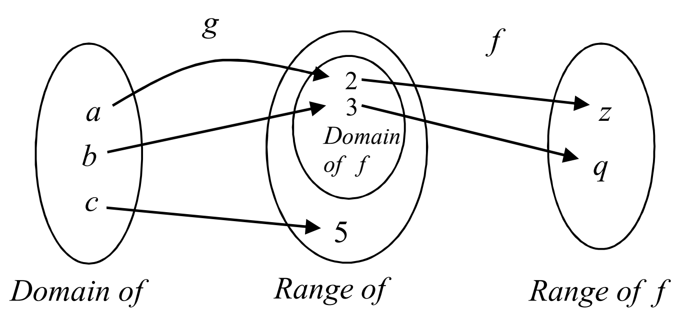
$$(f \circ g)(x) = f[g(x)].$$

$(f \circ g)(x)$ is read as "f composed with g of x". The domain of $f \circ g$ consists of those x's in the domain of g whose range values are in the domain of f, i.e. those x's for which $g(x)$ is in the domain of f.

Example 2.26:

1. Suppose $f = \{(2,z),(3,q)\}$ and $g = \{(a,2),(b,3),(c,5)\}$. The function $(f \circ g)(x) = f(g(x))$ is found by taking elements in the domain of g and evaluating as follows: $(f \circ g)(a) = f(g(a)) = f(2) = z,\quad (f \circ g)(b) = f(g(b)) = f(3) = q$

If we attempt to find $f(g(c))$ we get $f(5)$, but 5 is not in the domain of $f(x)$ and so we cannot find $(f \circ g)(c)$. Hence, $f \circ g = \{(a,z),(b,q)\}$. The figure below illustrates this situation.



2. Given $f(x)=5x^2-3x+2$ and $g(x)=4x+3$, find
   a) $(f\circ g)(-2)$ b) $(g\circ f)(2)$ c) $(f\circ g)(x)$ d) $(g\circ f)(x)$

Solution:

a) $(f\circ g)(-2)=f(g(-2))$ …… First evaluate $g(-2)=4(-2)+3=-5$
$$=f(-5)$$
$$=5(-5)^2-3(-5)+2=142$$

b) $(g\circ f)(2)=g(f(2))$ …….First evaluate $f(2)=5(2)^2-3(2)+2=16$
$$=g(16)$$
$$=4(16)+3=67$$

c) $(f\circ g)(x)=f(g(x))$ ……. But $g(x)=4x+3$
$$=f(4x+3)$$
$$=5(4x+3)^2-3(4x+3)+2$$
$$=80x^2+108x+38$$

d) $(g\circ f)(x)=g(f(x))$ ……. But $f(x)=5x^2-3x+2$
$$=g(5x^2-3x+2)$$
$$=4(5x^2-3x+2)+3$$
$$=20x^2-12x+11$$

3. Given $f(x)=\dfrac{x}{x+1}$ and $g(x)=\dfrac{2}{x-1}$, find
   a) $(f\circ g)(x)$ and its domain b) $(g\circ f)(x)$ and its domain

Solution: a) $(f\circ g)(x)=f\left(\dfrac{2}{x-1}\right)=\dfrac{\frac{2}{x-1}}{\frac{2}{x-1}+1}=\dfrac{2}{x+1}$. Thus, $Dom(f\circ g)=\{x:x\neq\pm1\}$.

b) $(g\circ f)(x)=g(f(x))=\dfrac{2}{\frac{x}{x+1}-1}=-2x-2$. Since x must first be an input into $f(x)$ and so must be in the domain of f, we see that $Dom(g\circ f)=\{x:x\neq-1\}$.

4. Let $f(x)=\dfrac{6x}{x^2-9}$ and $g(x)=\sqrt{3x}$. Find $(f\circ g)(12)$ and $(g\circ f)(x)$ and its domain.

Solution: We have $(f\circ g)(12)=f(g(12))=f(\sqrt{36})=f(6)=\frac{36}{27}=\frac{4}{3}$.
$$(f\circ g)(x)=f(g(x))=f(\sqrt{3x})=\frac{6\sqrt{3x}}{(\sqrt{3x})^2-9}=\frac{6\sqrt{3x}}{3x-9}=\frac{2\sqrt{3x}}{x-3}.$$
The domain of $f\circ g$ is $[0,3)\cup(3,\infty)$.

We now explore the meaning of equality of two functions. Let $f : A \to B$ and $g : A \to B$ be two functions. Then, f and g are subsets of $A \times B$. Suppose $f = g$. Let x be any element of A. Then, $(x, f(x)) \in f = g$ and thus $(x, f(x)) \in g$. Since g is a function and $(x, f(x))$, $(x, g(x)) \in g$, we must have $f(x) = g(x)$. Conversely, assume that $g(x) = f(x)$ for all $x \in A$. Let $(x, y) \in f$. Then, $y = f(x) = g(x)$. Thus, $(x, y) \in g$, which implies that $f \subseteq g$. Similarly, we can show that $g \subseteq f$. It now follows that $f = g$. Thus two functions $f : A \to B$ and $g : A \to B$ are equal if and only if $f(x) = g(x)$ for all $x \in A$. In general we have the following definition.

Definition 2.11: (Equality of functions)
Two functions are said to be equal if and only if the following two conditions hold:
- i) The functions have the same domain;
- **ii)** Their functional values are equal at each element of the domain.

Example 2.27:

1. Let $f : Z \to Z^+ \bigcup \{0\}$ and $g : Z \to Z^+ \bigcup \{0\}$ be defined by $f = \{(n, n^2) : n \in Z\}$ and $g = \{(n, |n|^2) : n \in Z\}$. Now, for all $n \in Z$, $f(n) = n^2 = |n|^2 = g(n)$. Thus, $f = g$.
2. Let $f(x) = \dfrac{x^2 - 25}{x - 5}, x \in \Re \setminus \{5\}$, and $g(x) = x + 5, x \in \Re$. The function f and g are not equal because $Dom(f) \neq Dom(g)$.

Exercise 2.4

1. For $f(x) = x^2 + x$ and $g(x) = \dfrac{2}{x+3}$, find each value:

   a) $(f - g)(2)$ c) $g^2(3)$ e) $(g \circ f)(1)$

   b) $\left(\dfrac{f}{g}\right)(1)$ d) $(f \circ g)(1)$ f) $(g \circ g)(3)$

2. If $f(x) = x^3 + 2$ and $g(x) = \dfrac{2}{x-1}$, find a formula for each of the following and state its domain.

   a) $(f + g)(x)$ c) $\left(\dfrac{g}{f}\right)(x)$

   b) $(f \circ g)(x)$ d) $(g \circ f)(x)$

3. Let $f(x) = x^2$ and $g(x) = \sqrt{x}$.

   a) Find $(f \circ g)(x)$ and its domain.

   b) Find $(g \circ f)(x)$ and its domain

c) Are $(f \circ g)(x)$ and $(g \circ f)(x)$ the same functions? Explain.

4. Let $f(x) = 5x - 3$. Find $g(x)$ so that $(f \circ g)(x) = 2x + 7$.
5. Let $f(x) = 2x + 1$. Find $g(x)$ so that $(f \circ g)(x) = 3x - 1$.
6. If f is a real function defined by $f(x) = \frac{x-1}{x+1}$. Show that $f(2x) = \frac{3f(x)+1}{f(x)+3}$.
7. Find two functions f and g so that the given function $h(x) = (f \circ g)(x)$, where

   a) $h(x) = (x+3)^3$ c) $h(x) = \frac{1}{x} + 6$

   b) $h(x) = \sqrt{5x-3}$ d) $h(x) = \frac{1}{x+6}$

8. Let $f(x) = 4x - 3$, $g(x) = \frac{1}{x}$ and $h(x) = x^2 - x$. Find

   a) $f(5x+7)$ c) $f(g(h(3)))$ e) $f(x+a)$

   b) $5f(x)+7$ d) $f(1) \cdot g(2) \cdot h(3)$ f) $f(x)+a$

2.5 **Types of functions**

After completing this section, the student should be able to:

- define one to oneness and ontoness of a function
- check invertibility of a function
- find the inverse of an invertible function

In this section we shall study some important types of functions.

- **One to One functions**

Definition 2.12: A function $f : A \to B$ is called **one to one**, often written 1 – 1, if and only if for all $x_1, x_2 \in A$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$. In words, no two elements of A are mapped to one element of B.

Example 2.28:

1. If we consider the sets $A = \{1,2,3,\cdots,6\}$ and $B = \{7,a,b,c,d,8,e\}$ and if $f = \{(1,7), (2,a), (3,b), (4,b), (5,c), (6,8)\}$ and $g = \{(1,7),(2,a),(3,b),(4,c),(5,8),(6,d)\}$, then both f and g are functions from A into B. Observe that f is not a 1 – 1 function because $f(3) = f(4)$ but $3 \neq 4$. However, g is a 1 – 1 function.

2. Let $A = \{1,2,3,4\}$ and $B = \{1,4,7,8\}$. Consider the functions

i) $f:A\to B$ defined as $f(1)=1, f(2)=4, f(3)=4, f(4)=8$

ii) $g:A\to B$ defined as $f(1)=4, f(2)=7, f(3)=1, f(4)=8$

Then, f is not 1 – 1, but g is a 1 – 1 function.

- **Onto functions**

Definition 2.13: Let f be a function from a set A into a set B. Then f is called an **onto function(or f maps onto B)** if every element of B is image of some element in A, i.e, $Range(f)=B.$

Example 2.29:

1. Let $A=\{1,2,3\}$ and $B=\{1,4,5\}$. The function $f:A\to B$ defined as $f(1)=1$, $f(2)=5$, $f(3)=1$ is not onto because there is no element in A, whose image under f is 4. The function $g:A\to B$ given by $g=\{(1,4),(2,5),(3,1)\}$ is onto because each element of B is the image of at least one element of A.

   Note that if A is a non-empty set, the function $i_A:A\to A$ defined by $i_A(x)=x$ for all $x\in A$ is a 1 – 1 function from A onto A. i_A is called the **identity map** on A.

2. Consider the relation f from Z into Z defined by $f(n)=n^2$ for all $n\in Z$. Now, domain of f is Z. Also, if $n=n'$, then $n^2=(n')^2$, i.e. $f(n)=f(n')$. Hence, f is well defined and a function. However, $f(1)=1=f(-1)$ and $1\neq -1$, which implies that f is not 1 – 1. For all $n\in Z$, $f(n)$ is a non-negative integer. This shows that a negative integer has no preimage. Hence, f is not onto. Note that f is onto $\{0,1,4,9,\cdots\}$.
3. Consider the relation f from Z into Z defined by $f(n)=2n$ for all $n\in Z$. As in the previous example, we can show that f is a function. Let $n,n'\in Z$ and suppose that $f(n)=f(n')$. Then $2n=2n'$ and thus $n=n'$. Hence, f is 1 – 1. Since for all $n\in Z$, $f(n)$ is an even integer; we see that an odd integer has no preimage. Therefore, f is not onto.

- **1 – 1 Correspondence**

Definition 2.14: A function $f:A\to B$ is said to be a 1 – 1 correspondence if f is both 1 – 1 and onto.

Example 2.30:

1. Let $A = \{0, 1, 2, 3, 4, 5\}$ and $B = \{0, 5, 10, 15, 20, 25\}$. Suppose $f : A \to B$ given by $f(x) = 5x$ for all $x \in A$. One can easily see that every element of B has a preimage in A and hence f is onto. Moreover, if $f(x) = f(y)$, then $5x = 5y$, i.e. $x = y$. Hence, f is 1 – 1. Therefore, f is a 1 – 1 correspondence between A and B.
2. Let A be a finite set. If $f : A \to A$ is onto, then it is one to one.

Solution: Let $A = \{a_1, a_2, \cdots, a_n\}$. Then $Range(f) = \{f(a_1), f(a_2), \cdots, f(a_n)\}$. Since f is onto we have $Range(f) = A$.Thus, $A = \{f(a_1), f(a_2), \cdots, f(a_n)\}$, which implies that $f(a_1)$, $f(a_2)$, $\cdots$, $f(a_n)$ are all distinct. Hence, $a_i \neq a_j$ implies $f(a_i) \neq f(a_j)$ for all $1 \leq i, j \leq n$. Therefore, f is 1 – 1.

- **Inverse of a function**

Since a function is a relation , the inverse of a function f is denoted by f^{-1} and is defined by:

$$f^{-1} = \{(y, x) : (x, y) \in f\}$$

For instance, if $f = \{(2,4), (3,6), (1,7)\}$, then $f^{-1} = \{(4,2), (6,3), (7,1)\}$. Note that the inverse of a function is not always a function. To see this consider the function $f = \{(2,4), (3,6), (5,4)\}$. Then, $f^{-1} = \{(4,2), (6,3), (4,5)\}$, which is not a function.

As we have seen above not all functions have an inverse, so it is important to determine whether or not a function has an inverse before we try to find the inverse. If the function does not have an inverse, then we need to realize that it does not have an inverse so that we do not waste our time trying to find something that does not exist.

A one to one function is special because only one to one functions have inverse. If a function is one to one, to find the inverse we will follow the steps below:

1. Interchange x and y in the equation $y = f(x)$
2. Solving the resulting equation for y, we will obtaining the inverse function.

Note that the domain of the inverse function is the range of the original function and the range of the inverse function is the domain of the original function.

Example 2.31:

1. Given $y = f(x) = x^3$. Find f^{-1} and its domain.

Solution: We begin by interchanging x and y, and we solve for y.

$y = x^3$	Interchange x and y
$x = y^3$	Take the cube root of both sides

$\sqrt[3]{x} = y$ This is the inverse of the function

Thus, $f^{-1}(x) = \sqrt[3]{x}$. The domain of f^{-1} is the set of all real numbers.

2. Let $y = f(x) = \dfrac{x}{x+2}$. Find $f^{-1}(x)$.

Solution: Again we begin by interchanging x and y, and then we solve for y.

$y = \dfrac{x}{x+2}$ Interchange x and y

$x = \dfrac{y}{y+2}$ Solving for y

$$x(y+2) = y \Leftrightarrow xy + 2x = y \Leftrightarrow 2x = y(1-x) \Leftrightarrow y = \frac{2x}{1-x}$$

Thus, $f^{-1}(x) = \dfrac{2x}{1-x}$.

Remark: Even though, in general, we use an exponent of -1 to indicate a reciprocal, inverse function notation is an exception to this rule. Please be aware that $f^{-1}(x)$ is not the reciprocal of f. That is,

$$f^{-1}(x) \neq \frac{1}{f(x)}$$

If we want to write the reciprocal of the function $f(x)$ by using a negative exponent, we must write

$$\frac{1}{f(x)} = [f(x)]^{-1}.$$

Exercise 2.5

1. Consider the function $f = \{(x, x^2) : x \in S\}$ from $S = \{-3,-2,-1,0,1,2,3\}$ into Z. Is f one to one? Is it onto?

2. Let $A = \{1,2,3\}$. List all one to one functions from A onto A.

3. Let $f : A \to B$. Let f^* be the inverse relation, i.e. $f^* = \{(y, x) \in B \times A : f(x) = y\}$.
   a) Show by an example that f^* need not be a function.
   b) Show that f^* is a function from $Range(f)$ into A if and only if f is 1 – 1.
   c) Show that f^* is a function from B into A if and only if f is 1 – 1 and onto.
   d) Show that if f^* is a function from B into A, then $f^{-1} = f^*$.

4. Let $A = \{x \in \Re : 0 \le x \le 1\}$ and $B = \{x \in \Re : 5 \le x \le 8\}$. Show that $f : A \to B$ defined by $f(x) = 5 + (8-5)x$ is a 1 – 1 function from A onto B.

5. Which of the following functions are one to one?
   a) $f : \Re \to \Re$ defined by $f(x) = 4, x \in \Re$
   b) $f : \Re \to \Re$ defined by $f(x) = 6x - 1, x \in \Re$
   c) $f : \Re \to \Re$ defined by $f(x) = x^2 + 7, x \in \Re$
   d) $f : \Re \to \Re$ defined by $f(x) = x^3, x \in \Re$
   e) $f : \Re \setminus \{7\} \to \Re$ defined by $f(x) = \dfrac{2x+1}{x-7}, x \in \Re \setminus \{7\}$

6. Which of the following functions are onto?
   a) $f : \Re \to \Re$ defined by $f(x) = 115x + 49, x \in \Re$
   b) $f : \Re \to \Re$ defined by $f(x) = |x|, x \in \Re$
   c) $f : \Re \to \Re$ defined by $f(x) = \sqrt{x^2}, x \in \Re$
   d) $f : \Re \to \Re$ defined by $f(x) = x^2 + 4, x \in \Re$

7. Find $f^{-1}(x)$ if

   a) $f(x) = 7x - 6$
   b) $f(x) = \dfrac{2x-9}{4}$
   c) $f(x) = 1 - \dfrac{3}{x}$
   d) $f(x) = \dfrac{4-x}{3x}$
   e) $f(x) = \dfrac{5x+3}{1-2x}$
   f) $f(x) = \sqrt[3]{x+1}$
   g) $f(x) == -(x+2)^2 - 1$
   h) $f(x) = \dfrac{2x}{1+x}$

2.6 **Polynomials, zeros of polynomials, rational functions and their graphs**

After completing this section, the student should be able to:

- define polynomial and rational functions
- apply the theorems on polynomials to find the zeros of polynomial functions
- use the division algorithm to find quotient and remainder
- apply theorems on polynomials to solve related problems
- sketch and analyze the graphs of rational functions

The functions described in this section frequently occur as mathematical models of real-life situations. For instance, in business the demand function gives the price per item, p, in terms of the number of items sold, x. Suppose a company finds that the price p (in Birr) for its model GC-5 calculator is related to the number of calculators sold, x (in millions), and is given by the demand function $p = 80 - x^2$.

The manufacturer's revenue is determined by multiplying the number of items sold (x) by the price per item (p). Thus, the revenue function is

$$R = xp = x(80 - x^2) = 80x - x^3$$

These demand and revenue functions are examples of polynomial functions. The major aim of this section is to better understand the significance of applied functions (such as this demand function). In order to do this, we need to analyze the domain, range, and behavior of such functions.

- **Polynomial functions**

Definition 2.15: A polynomial function is a function of the form

$$y = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \; a_n \neq 0.$$

Each a_i is assumed to be a real number, and n is a non-negative integer, a_n is called the leading coefficient. Such a polynomial is said to be of degree n.

Remark:

1. The domain of a polynomial function is always the set of real numbers.
2. (Types of polynomials)
   - A polynomial of degree 1 is called a linear function.
   - A polynomial of degree 2 is called quadratic function.
   - A polynomial of degree 3 is called a cubic function.

     i.e $p(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0, \; a_3 \neq 0.$

Example 2.32: $p(x) = 2x^2 + 1$, $q(x) = \sqrt{3}x^4 + 2x - \pi$ and $f(x) = 2x^3$ are examples of polynomial functions.

- **Properties of polynomial functions**

1. The graph of a polynomial is a smooth unbroken curve. The word smooth means that the graph does not have any sharp corners as turning points.
2. If p is a polynomial of degree n, then it has at most n zeros. Thus, a quadratic polynomial has at most 2 zeros.
3. The graph of a polynomial function of degree n can have at most $n-1$ turning points. Thus, the graph of a polynomial of degree 5 can have at most 4 turning points.

4. The graph of a polynomial always exhibits the characteristic that as $|x|$ gets very large, $|y|$ gets very large.

- **Zeros of a polynomial**

The zeros of a polynomial function provide valuable information that can be helpful in sketching its graph. One can find the zeros by factorizing the polynomial. However, we have no general method for factorizing polynomials of degree greater than 2. In this subsection, we turn our attention to methods that will allow us to find zeros of higher degree polynomials. To do this, we first need to discuss about the division algorithm.

Division Algorithm
Let $p(x)$ and $d(x)$ be polynomials with $d(x) \neq 0$, and with the degree of $d(x)$ less than or equal to the degree of $p(x)$. Then there are polynomials $q(x)$ and $R(x)$ such that $\underbrace{p(x)}_{dividend} = \underbrace{d(x)}_{divisor}.\underbrace{q(x)}_{quotient} + \underbrace{R(x)}_{remainder}$, where either $R(x) = 0$ or the degree of $R(x)$ is less than degree of $d(x)$.

Example 2.33: Divide $\dfrac{x^4 - 1}{x^4 + 2x}$.

Solution: Using long division we have

$$\begin{array}{r l}
 & x^2 - 2x + 4 \\
x^2 + 2x \,\big) & x^4 + 0x^3 + 0x^2 + 0x + 1 \\
 & \underline{-(x^4 + 2x^3)} \\
 & \quad -2x^3 + 0x^2 \\
 & \quad \underline{-(-2x^3 - 4x^2)} \\
 & \qquad 4x^2 + 0x \\
 & \qquad \underline{-(4x^2 + 8x)} \\
 & \qquad\qquad -8x - 1
\end{array}$$

This long division means $\underbrace{x^4 - 1}_{dividend} = \underbrace{(x^2 + 2x)}_{divisor}.\underbrace{(x^2 - 2x + 4)}_{quotient} + \underbrace{(-8x - 1)}_{remainder}$.

With the aid of the division algorithm, we can derive two important theorems that will allow us to recognize the zeros of polynomials.

If we apply the division algorithm where the divisor, $d(x)$, is linear (that is of the form $x-r$), we get

$$p(x)=(x-r)q(x)+R$$

Note that since the divisor is of the first degree, the remainder R, must be a constant. If we now substitute $x=r$, into this equation, we get

$$P(r)=(r-r)q(r)+R=0\cdot q(r)+R$$

Therefore, $p(r)=R$.

The result we just proved is called the remainder theorem.

The Remainder Theorem
When a polynomial $p(x)$ of degree at least 1 is divided by $x-r$, then the remainder is $p(r)$.

Example 2.34: The remainder when $P(x)=x^3-x^2+3x-1$ is divided by $x-2$ is $p(2)=9$.

As a consequence of the remainder theorem, if $x-r$ is a factor of $p(x)$, then the remainder must be 0. Conversely, if the remainder is 0, then $x-r$, is a factor of $p(x)$. This is known as the Factor Theorem.

The Factor Theorem
$x-r$ is a factor of $p(x)$ if and only if $p(r)=0$.

The next theorem, called location theorem, allows us to verify that a zero exists somewhere within an interval of numbers, and can also be used to zoom in closer on a value.

Location theorem
Let f be a polynomial function and a and b be real numbers such that $a<b$. If $f(a)f(b)<0$, then there is at least one zero of f between a and b.

The Factor and Remainder theorems establish the intimate relationship between the factors of a polynomial $p(x)$ and its zeros. Recall that a polynomial of degree *n* can have at most *n* zeros.

Does every polynomial have a zero? Our answer depends on the number system in which we are working. If we restrict ourselves to the set of real number system, then we are already familiar with the fact that the polynomial $p(x)=x^2+1$ has no real zeros. However, this polynomial does have two zeros in the complex number system. (The zeros are i and $-i$). Carl Friedrich Gauss (1777-1855), in his doctoral dissertation, proved that within the complex number system, every polynomial of degree ≥ 1 has at least one zero. This fact is usually referred to as the Fundamental theorem of Algebra.

Fundamental Theorem of Algebra
If $p(x)$ is a polynomial of degree $n > 0$ whose coefficients are complex numbers, then $p(x)$ has at least one zero in the complex number system.

Note that since all real numbers are complex numbers, a polynomial with real coefficients also satisfies the Fundamental theorem of Algebra. As an immediate consequence of the Fundamental theorem of Algebra, we have

The linear Factorization Theorem
If $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $n \geq 1$ and $a_n \neq 0$, then $p(x) = a_n (x - r_1)\ (x - r_2) \cdots (x - r_n)$, where the r_i are complex numbers (possible real and not necessarily distinct).

From the linear factorization theorem, it follows that every polynomial of degree $n \geq 1$ has exactly n zeros in the complex number system, where a root of multiplicity k counted k times.

Example 2.35: Express each of the polynomials in the form described by the Linear Factorization Theorem. List each zero and its multiplicity.

a) $p(x) = x^3 - 6x^2 - 16x$

b) $q(x) = 3x^2 - 10x + 8$

c) $f(x) = 2x^4 + 8x^3 + 10x^2$

Solution:

a) We may factorize $p(x)$ as follows:

$$\begin{aligned} p(x) = x^3 - 6x^2 - 16x &= x(x^2 - 6x - 16) \\ &= x(x-8)(x+2) \\ &= x(x-8)(x-(-2)) \end{aligned}$$

The zeros of $p(x)$ are 0, 8, and – 2 each of multiplicity one.

b) We may factorize $q(x)$ as follows:

$$\begin{aligned} q(x) = 3x^2 - 10x + 8 &= (3x-4)\ (x-2) \\ &= 3(x - \frac{4}{3})(x-2) \end{aligned}$$

Thus, the zeros of $q(x)$ are $\frac{4}{3}$ and 2, each of multiplicity one.

c) We may factorize $f(x)$ as follows:

$$f(x) = 2x^4 + 8x^3 + 10x^2 = 2x^2(x^2 + 4x + 5)$$
$$= 2x^2(x - (-2+i))(x - (-2-i))$$

Thus, the zeros of f(x) are 0 with multiplicity two and $-2+i$ and $-2-i$ each with multiplicity one.

Example 2.36:

1. Find a polynomial $p(x)$ with exactly the following zeros and multiplicity.

zeros	multiplicity
-1	3
2	4
5	2

   Are there any other polynomials that give the same roots and multiplicity?

2. Find a polynomial $f(x)$ having the zeros described in part (a) such that $f(1) = 32$.

Solution:

1. Based on the Factor Theorem, we may write the polynomial as:

   $$p(x) = (x-(-1))^3 (x-2)^4 (x-5)^2 = (x+1)^3 (x-2)^4 (x-5)^2$$

   which gives the required roots and multiplicities.

   Any polynomial of the form $kp(x)$, where k is a non-zero constant will give the same roots and multiplicities.

2. Based on part (1), we know that $f(x) = k(x+1)^3 (x-2)^4 (x-5)^2$. Since we want $f(x) = 32$, we have

   $$f(1) = k(1+1)^3 (1-2)^4 (1-5)^2$$
   $$32 = k(8)(1)(16) \quad \Rightarrow k = \tfrac{1}{4}$$

   Thus, $f(x) = \tfrac{1}{4}(x+1)^3 (x-2)^4 (x-5)^2$.

Our experience in using the quadratic formula on quadratic equations with real coefficients has shown us that complex roots always appear in conjugate pairs. For example, the roots of $x^2 - 2x + 5 = 0$ are $1+2i$ and $1-2i$. In fact, this property extends to all polynomial equations with real coefficients.

Conjugate Roots Theorem
Let $p(x)$ be a polynomial with real coefficients. If complex number $a+bi$ (where a and b are real numbers) is a zero of $p(x)$, then so is its conjugate $a-bi$.

Example 2.37: Let $r(x) = x^4 + 2x^3 - 9x^2 + 26x - 20$. Given that $1 - \sqrt{3}\ i$ is a zero, find the other zero of $r(x)$.

Solution: According to the Conjugate Roots Theorem, if $1 - \sqrt{3}\ i$ is a zero, then its conjugate, $1 + \sqrt{3}\ i$ must also be a zero. Therefore, $x - (1 - \sqrt{3}\ i)$ and $x - (1 + \sqrt{3}\ i)$ are both factors of $r(x)$, and so their product must be a factor of $r(x)$. That is, $[x - (1 - \sqrt{3}\ i)]\,[x - (1 + \sqrt{3}\ i)] = x^2 - 2x + 4$ is a factor of $r(x)$. Dividing $r(x)$ by $x^2 - 2x + 4$, we obtain

$$r(x) = (x^2 - 2x + 4)(x^2 + 4x - 5) = (x^2 - 2x + 4)\,(x + 5)\,(x - 1).$$

Thus, the zeros of $r(x)$ are $1 - \sqrt{3}\,i$, $1 + \sqrt{3}\,i$, -5 and 1.

The theorems we have discussed so far are called existence theorems because they ensure the existence of zeros and linear factors of polynomials. These theorems do not tell us how to find the zeros or the linear factors. The Linear Factorization Theorem guarantees that we can factor a polynomial of degree at least one into linear factors, but it does not tell us how.
We know from experience that if $p(x)$ happens to be a quadratic function, then we may find the zeros of $p(x) = Ax^2 + Bx + C$ by using the quadratic formula to obtain the zeros

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

The rest of this subsection is devoted to developing some special methods for finding the zeros of a polynomial function.

As we have seen, even though we have no general techniques for factorizing polynomials of degree greater than 2, if we happen to know a root, say r, we can use long division to divide $p(x)$ by $x - r$ and obtain a quotient polynomial of lower degree. If we can get the quotient polynomial down to a quadratic, then we are able to determine all the roots. But how do we find a root to start the process? The following theorem can be most helpful.

The Rational Root Theorem
Suppose that $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $n \ge 1$, $a_n \ne 0$ is an n^{th} degree polynomial with integer coefficients. If $\frac{p}{q}$ is a rational root of $f(x) = 0$, where p and q have no common factor other than ± 1, then p is a factor of a_0 and q is a factor of a_n.

To get a feeling as to why this theorem is true, suppose $\frac{3}{2}$ is a root of

$a_3x^3 + a_2x^2 + a_1x + a_0 = 0$.

Then, $a_3\left(\frac{3}{2}\right)^3 + a_2\left(\frac{3}{2}\right)^2 + a_1\left(\frac{3}{2}\right) + a_0 = 0$ which implies that

$$\frac{27a_3}{8} + \frac{9a_2}{4} + \frac{3a_1}{2} + a_0 = 0 \qquad \text{multiplying both sides by 8}$$

$$27a_3 + 18a_2 + 12a_1 = -8a_0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$27a_3 = -18a_2 - 12a_1 - 8a_0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

If we look at equation (1), the left hand side is divisible by 3, and therefore the right hand side must also be divisible by 3. Since 8 is not divisible by 3, a_0 must be divisible by 3. From equation (2), a_3 must be divisible by 2.

Example 2.38: Find all the zeros of the function $p(x) = 2x^3 + 3x^2 - 23x - 12$.

Solution: According to the Rational Root Theorem, if $\frac{p}{q}$ is a rational root of the given equation, then p must be a factor of -12 and q must be a factor of 2. Thus, we have

possible values of p: $\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12$

possible values of q: $\pm 1, \pm 2$

possible rational roots $\frac{p}{q}$: $\pm 1, \pm\frac{1}{2}, \pm 2, \pm 3, \pm\frac{3}{2}, \pm 4, \pm 6, \pm 12$

We may check these possible roots by substituting the value in $p(x)$. Now $p(1) = -30$ and $p(-1) = 12$. Since $p(1)$ is negative and $p(-1)$ is positive, by intermediate value theorem, $p(x)$ has a zero between -1 and 1. Since $P\left(-\frac{1}{2}\right) = 0$, then $\left(x + \frac{1}{2}\right)$ is a factor of $p(x)$. Using long division, we obtain

$$p(x) = 2x^3 + 3x^2 - 23x - 12 = (x + \tfrac{1}{2})(2x^2 + 2x - 24)$$
$$= 2(x + \tfrac{1}{2})(x + 4)(x - 3)$$

Therefore, the zeros of $p(x)$ are $-\frac{1}{2}$, -4 and 3.

- **Rational Functions and their Graphs**

A rational function is a function of the form $f(x) = \frac{n(x)}{d(x)}$ where both $n(x)$ and $d(x)$ are polynomials and $d(x) \neq 0$.

Example 2.39: The functions $f(x)=\frac{3}{x+5}$, $f(x)=\frac{x-1}{x^2-4}$ and $f(x)=\frac{x^5+2x^3-x+1}{x \ +5x}$ are examples of rational function.

Note that the domain of the rational function $f(x)=\frac{n(x)}{d(x)}$ is $\{x : d(x) \neq 0\}$

Example 2.40: Find the domain and zeros of the function $f(x)=\frac{3x-5}{x^2-x-12}$.

Solution: The values of x for which $x^2-x-12=0$ are excluded from the domain of f. Since $x^2-x-12=(x-4)(x+3)$, we have $Dom(f)=\{x : x \neq -3,4\}$. To find the zeros of $f(x)$, we solve the equation

$$\frac{n(x)}{d(x)}=0 \Leftrightarrow n(x)=0 \ \& \ q(x) \neq 0$$

Therefore, to find the zeros of $f(x)$, we solve $3x-5=0$, giving $x=\frac{5}{3}$. Since $\frac{5}{3}$ does not make the denominator zero, it is the only zero of $f(x)$.

The following terms and notations are useful in our next discussion.

Given a number *a*,

- x approaches a from the right means x takes any value near and near to a but $x > a$. This is denoted by: $x \to a^+$ (read: '*x* approaches *a* from the right').

  For instance, $x \to 1^+$ means *x* can be 1.001, 1.0001, 1.00001, 1.000001, etc.

- x approaches a from the left means x takes any value near and near to a but $x < a$.

  This is denoted by: $x \to a^-$ (read: '*x* approaches *a* from the left').

  For instance, $x \to 1^-$ means *x* can be 0.99, 0.999, 0.9999, 0.9999, etc.

- $x \to \infty$ (read: '*x* approaches or tends to *infinity*') means the value of x gets indefinitely larger and larger in magnitude (keep increasing without bound). For instance, *x* can be 10^6, 10^{10}, 10^{12}, etc.

- $x \to -\infty$ (read: '*x* approaches or tends to negative *infinity*') means the value of *x* is negative and gets indefinitely larger and larger negative in magnitude (keep decreasing without bound). For instance, *x* can be -10^6, -10^{10}, -10^{12}, etc.

The same meanings apply also for the values of a function *f* if we wrote $f(x) \to \infty$ or $f(x) \to -\infty$. The following figure illustrates these notion and notations.

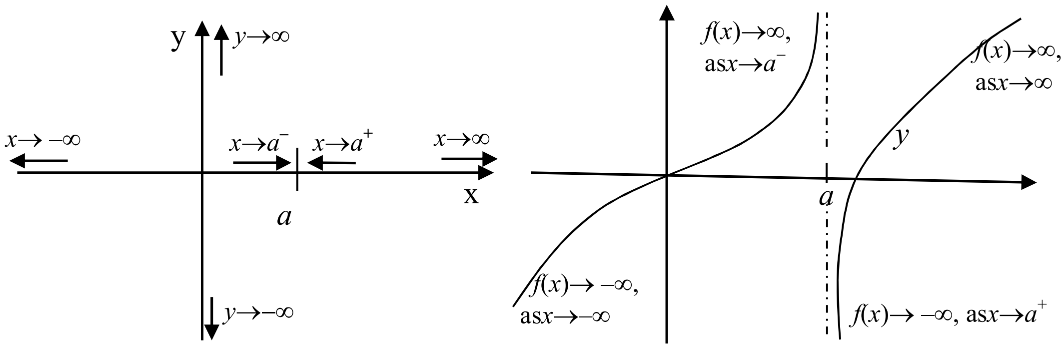


Fig. 2.1. Graphical illustration of the idea of $x \to a^+$, $f(x) \to \infty$, etc.

We may also write $f(x) \to b$ (read: '$f(x)$ approaches b') to mean the function values, $f(x)$, becomes arbitrarily closer and closer to b (i.e., approximately b) but not exactly equal to b. For instance, if $f(x) = \frac{1}{x}$, then $f(x) \to 0$ as $x \to \infty$; i.e., $\frac{1}{x}$ is approximately 0 when x is arbitrarily large.

The following steps are usually used to sketch (or draw) the graph of a rational function $f(x)$.

1. Identify the domain and simplify it.
2. Find the intercepts of the graph whenever possible. Recall the following:
   - y–intercept is the point on y-axis where the graph of $y = f(x)$ intersects with the y-axis. At this point x=0. Thus, $y = f(0)$, or $(0, f(0))$ is the y-intercept if $0 \in \text{Dom}(f)$.
   - x–intercept is the point on x-axis where the graph of $y = f(x)$ intersects with the x-axis. At this point y=0. Thus, x=a or $(a, 0)$ is x-intercept if $f(a)$=0.
3. Determine the asymptotes of the graph. Here, remember the following.

- Vertical Asymptote: The vertical line x=a is called a vertical asymptote(VA) of $f(x)$ if
  i) $a \notin \text{dom}(f)$, i.e., f is not defined at x=a; and
  ii) $f(x) \to \infty$ or $f(x) \to -\infty$ when $x \to a^+$ or $x \to a^-$. In this case, the graph of f is almost vertically rising upward (if $f(x) \to \infty$) or sinking downward (if $f(x) \to -\infty$) along with the vertical line x=a when x approaches a either from the right or from the left.

Example 2.41: Consider $f(x) = \frac{1}{(x-a)^n}$, where $a \geq 0$ and n is a positive integer.

Obviously $a \notin \text{Dom}(f)$. Next, we investigate the trend of the values of $f(x)$ near a. To do this, we consider two cases, when n is even or odd:

Suppose n is even: In this case $(x - a)^n > 0$ for all $x \in \Re \setminus \{a\}$; and since $(x - a)^n \to 0$ as $x \to a^+$ or $x \to a^-$. Hence, $f(x) = \frac{1}{(x-a)^n} \to \infty$ as $x \to a^+$ or $x \to a^-$. Therefore, x=a is a VA of $f(x)$. Moreover, $y = 1/a^n$ or $(0, 1/a^n)$ is its y-intercept since $f(0) = 1/a^n$. However, it has no x-intercept since $f(x) > 0$ for all x in its domain (See, Fig. 2.2 (A)).

Suppose n is odd: In this case $(x-a)^n > 0$ for all $x>a$ and $1/(x-a)^n \to \infty$ when $x \to a^+$ as in the above case. Thus, $x=a$ is its VA. However, $1/(x-a)^n \to -\infty$ when $x \to a^-$ since $(x-a)^n < 0$ for $x<a$. Moreover, $y=-1/a^n$ or $(0, -1/a^n)$ is its y-intercept since $f(0)=-1/a^n$. However, it has no x-intercept also in this case. (See, Fig. 2.2 (B)).

Note that in both cases, $f(x)=\dfrac{1}{(x-a)^n} \to 0$ as $x \to \infty$ or $x \to -\infty$.

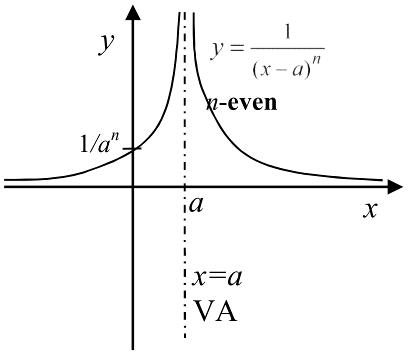


Fig. 2.2 (A)

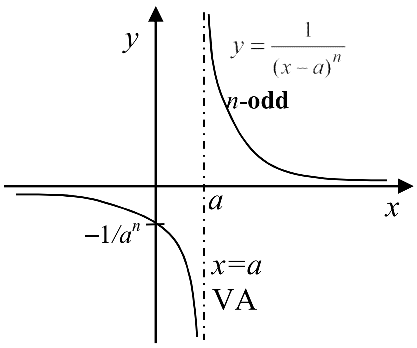


Fig. 2.2 (B)

Remark: Let $f(x)=\dfrac{n(x)}{d(x)}$ be a rational function. Then,

1. if $d(a)=0$ and $n(a)\neq 0$, then $x=a$ is a VA of f.
2. if $d(a)=0=n(a)$, then $x=a$ may or may **not** be a VA of f. In this case, simplify $f(x)$ and look for VA of the simplest form of f.

- Horizontal Asymptote: A horizontal line $y=b$ is called horizontal asymptote (HA) of $f(x)$ if the value of the function becomes closer and closer to b (i.e., $f(x) \to b$) as $x \to \infty$ or as $x \to -\infty$.

  In this case, the graph of f becomes almost a horizontal line along with (or near) the line $y=b$ as $x \to \infty$ and as $x \to -\infty$. For instance, from the above example, the HA of $f(x)=\dfrac{1}{(x-a)^n}$ is $y=0$ (the x-axis), for any positive integer n (See, Fig. 2.2).

Remark: A rational function $f(x)=\dfrac{n(x)}{d(x)}$ has a HA only when $degree(n(x)) \le degree(d(x))$.

In this case, (i) If $degree(n(x)) < degree(d(x))$, then $y=0$ (the x-axis) is the HA of f.

(ii) If $degree(n(x)) = degree(d(x)) = n$, i.e., $f(x)=\dfrac{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0}{b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0}$,

then $y=\dfrac{a_n}{b_n}$ is the HA of f.

- Oblique Asymptote: The oblique line $y=ax+b$, $a\neq 0$, is called an oblique asymptote (OA) of f if the value of the function, $f(x)$, becomes closer and closer to $ax+b$(i.e., $f(x)$ becomes approximately $ax+b$) as either $x\to\infty$ or $x\to -\infty$. In this case, the graph of f becomes almost a straight line along with (or near) the oblique line $y=ax+b$ as $x\to\infty$ and as $x\to -\infty$.

Note: A rational function $f(x)=\dfrac{n(x)}{d(x)}$ has an OA only when $degree(n(x)) = degree(d(x)) + 1$. In this case, using long division, if the quotient of $n(x) \div d(x)$ is $ax+b$, then $y=ax+b$ is the OA of f.

Example 2.42: Sketch the graphs of (a) $f(x)=\dfrac{x+2}{x-1}$ (b) $g(x)=\dfrac{x^2+3x+2}{x^2-1}$

Solution: (a) Since $x-1=0$ at $x=1$, dom(f) = $\Re\backslash\{1\}$.

- Intercepts: y-intercept: $x=0 \Rightarrow y=f(0)=-2$. Hence, $(0,-2)$ is y-intercept.
  x-intercept: $y=0 \Rightarrow x+2=0 \Rightarrow x=-2$. Hence, $(-2, 0)$ is x-intercept.
- Asymptotes:
  - VA: Since $x-1=0$ at $x=1$ and $x+2\neq 0$ at $x=1$, $x=1$ is VA of f. In fact, if $x\to 1^+$, then $x+2\approx 3$ but the denominator $x-1$ is almost 0 (but positive).
    Consequently, $f(x)\to\infty$ as $x\to 1^+$.
    Moreover, $f(x)\to -\infty$ as $x\to 1^-$ (since , if $x\to 1^-$ then $x-1$ is almost 0 but negative) .

    (So, the graph of f rises up to $+\infty$ at the right side of $x=1$, and sink down to $-\infty$ at the left side of $x=1$)
  - HA: Note that if you divide $x+2$ by $x-1$, the quotient is 1 and remainder is 3. Thus, $f(x)=\dfrac{x+2}{x-1}=1+\dfrac{3}{x-1}$. Thus, if $x\to\infty$ (or $x\to -\infty$), then $\dfrac{3}{x-1}\to 0$ so that $f(x)\to 1$.
    Hence, $y=1$ is the HA of f.

Using these information, you can sketch the graph of f as displayed below in Fig. 2.3 (A).

(b) Both the denominator and numerator are 0 at $x=1$. So, first factorize and simplify them: $x^2+3x+2=(x+2)(x+1)$ and $x^2-1=(x-1)(x+1)$. Therefore,

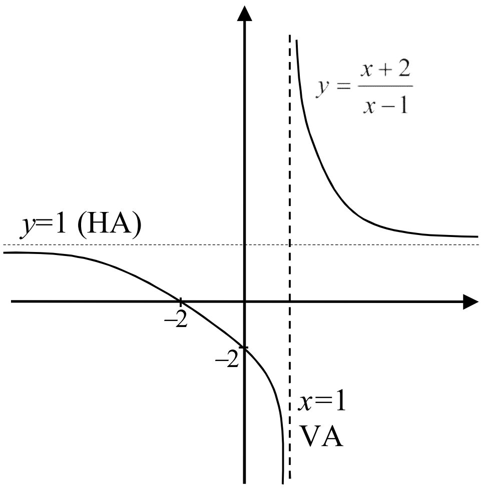
$$g(x)=\frac{x^2+3x+2}{x^2-1}=\frac{(x+2)\cancel{(x+1)}}{(x-1)\cancel{(x+1)}},\quad x\neq -1$$

$$=\frac{x+2}{x-1}.\qquad (\text{So, } \operatorname{dom}(g)=\Re\backslash\{1,-1\}\ )$$

This implies that only $x=1$ is VA.

Hence, the graph of $g(x)=\dfrac{x+2}{x-1}$, $x\neq -1$, is exactly the same as that of $f(x)=\dfrac{x+2}{x-1}$ except that $g(x)$ is not defined at $x=-1$. Therefore, the graph of g and its VA are the same as that of f

except that there should be a 'hole' at the point corresponding to $x = -1$ on the graph of g as shown on Fig. 2.3(B) below.



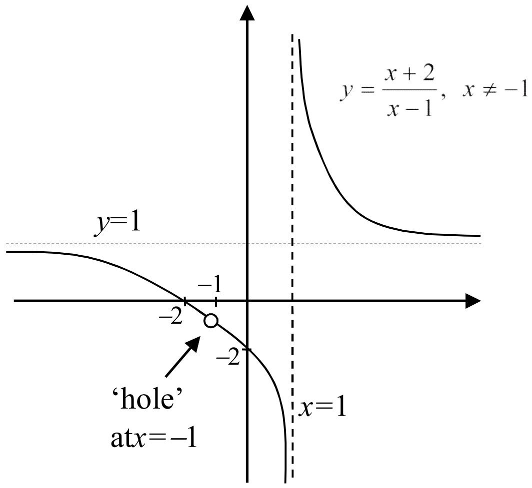


Fig 2.3 (A) $f(x) = \frac{x+2}{x-1}$ (B) $g(x) = \frac{x^2+3x+2}{x^2-1} = \frac{x+2}{x-1},\ x \neq -1$

Exercise 2.6

1. Perform the requested divisions. Find the quotient and remainder and verify the Remainder Theorem by computing $p(a)$.
   a) Divide $p(x) = x^2 - 5x + 8$ *by* $x + 4$
   b) Divide $p(x) = 2x^3 - 7x^2 + x + 4$ *by* $x - 4$
   c) Divide $p(x) = 1 - x^4$ *by* $x - 1$
   d) Divide $p(x) = x^5 - 2x^2 - 3$ *by* $x + 1$
2. Given that $p(4) = 0$, factor $p(x) = 2x^3 - 11x^2 + 10x + 8$ as completely as possible.
3. Given that $r(x) = 4x^3 - x^2 - 36x + 9$ and $r(\frac{1}{4}) = 0$, find the remaining zeros of $r(x)$.
4. Given that 3 is a double zero of $p(x) = x^4 - 3x^3 - 19x^2 + 87x - 90$, find all the zeros of $p(x)$.
5. a) Write the general polynomial $p(x)$ whose only zeros are 1, 2 and 3, with multiplicity 3, 2 and 1 respectively. What is its degree?
   b) Find $p(x)$ described in part (a) if $p(0) = 6$.
6. If $2 - 3i$ is a root of $p(x) = 2x^3 - 5x^2 + 14x + 39$, find the remaining zeros of p(x).
7. Determine the rational zeros of the polynomials
   a) $p(x) = x^3 - 4x^2 - 7x + 10$
   b) $p(x) = 2x^3 - 5x^2 - 28x + 15$

c) $p(x) = 6x^3 + x^2 - 4x + 1$

8. Find the domain and the real zeros of the given function.

a) $f(x) = \frac{3}{x^2 - 25}$ b) $g(x) = \frac{x-3}{x^2 4x - 12}$ c) $f(x) = \frac{(x-3)^2}{x^3 - 3x^2 + 2x}$ d) $f(x) = \frac{x^2 - 16}{x^2 + 4}$

9. Sketch the graph of

a) $f(x) = \frac{1-x}{x-3}$ b) $f(x) = \frac{x^2+1}{x}$ c) $f(x) = \frac{1}{x} + 2$ d) $f(x) = \frac{x^2}{x^2 - 4}$

10. Determine the behavior of $f(x) = \frac{x^3 - 8x - 3}{x - 3}$ when x is near 3.

11. The graph of any rational function in which the degree of the numerator is exactly one more than the degree of the denominator will have an oblique (or slant) asymptote.

a) Use long division to show that

$$y = f(x) = \frac{x^2 - x + 6}{x - 2} = x + 1 + \frac{8}{x - 2}$$

b) Show that this means that the line $y = x + 1$ is a slant asymptote for the graph and sketch the graph of $y = f(x)$.

2.7 Definition and basic properties of logarithmic, exponential, and trigonometric functions and their graphs

After completing this section, the student should be able to:

- define exponential, logarithmic and trigonometric functions
- understand the relationship between exponential and logarithmic functions
- sketch the graph of exponential, logarithmic, and trigonometric functions
- use basic properties of logarithmic, exponential and trigonometric functions to solve problems

- **Exponents and radicals**

Definition 2.16: For a natural number n and a real number x, the power x^n, read " the n^{th} power of x" or "x raised to n", is defined as follows:

$$x^n = \underbrace{x \cdot x \cdot \cdots \cdot x}_{n \text{ factors each equal to } x}$$

In the symbol x^n, x is called the base and n is called the exponent.

For example, $2^5 = 2 \times 2 \times 2 \times 2 \times 2 = 32$.

Based of the definition of x^n, n must be a natural number. It does not make sense for n to be negative or zero. However, we can extend the definition of exponents to include 0 and negative exponents.

Definition 2.17: (Zero and Negative Exponents)	
Definition of zero Exponent	*Definition of Negative Exponent*
$x^0 = 1 \ (x \neq 0)$	$x^{-n} = \dfrac{1}{x^n} \quad (x \neq 0)$
Note: 0^0 is undefined.	

As a result of the above definition, we have $\dfrac{1}{x^{-n}} = x^n$. We have the following rules of exponents for integer exponents:

Rules for Integer Exponents

1. $x^n \cdot x^m = x^{n+m}$
2. $(x^n)^m = x^{nm}$
3. $\left(\dfrac{x}{y}\right)^n = \dfrac{x^n}{y^n} \quad (y \neq 0)$
4. $(xy)^n = x^n y^n$
5. $\dfrac{x^n}{x^m} = x^{n-m}$

Next we extend the definition of exponents even further to include rational number exponents. To do this, we assume that we want the rules for integer exponents also to apply to rational exponents and then use the rules to show us to define a rational exponent. For example, how do we define $a^{\frac{1}{2}}$? Consider $9^{\frac{1}{2}}$.

If we apply rule 2 and square $9^{\frac{1}{2}}$, we get $\left(9^{\frac{1}{2}}\right)^2 = 9^{\frac{2}{2}} = 9$. Thus, $9^{\frac{1}{2}}$ is a number that, when squared, yields 9. There are two possible answers: 3 and – 3, since squaring either number will yield 9. To avoid ambiguity, we define $a^{\frac{1}{2}}$ (called the principal square root of a) as the non-negative quantity that, when squared, yield a. Thus, $9^{\frac{1}{2}} = 3$.

We will arrive at the definition of $a^{\frac{1}{3}}$ in the same way as we did for $a^{\frac{1}{2}}$. For example, if we cube $8^{\frac{1}{3}}$, we get $\left(8^{\frac{1}{3}}\right)^3 = 8^{\frac{3}{3}} = 8$. Thus, $8^{\frac{1}{3}}$ is the number that, when cubed, yields 8. Since $2^3 = 8$ we have $8^{\frac{1}{3}} = 2$. Similarly, $(-27)^{\frac{1}{3}} = -3$. Thus, we define $a^{\frac{1}{3}}$ (called the cube root of a) as the quantity that, when cubed yields a.

Definition 2.18: (Rational Exponent $a^{\frac{1}{n}}$)

If n is an odd positive integer, then $a^{\frac{1}{n}} = b$ if and only if $b^n = a$

If n is an even positive integer and $a \geq 0$, then $a^{\frac{1}{n}} = |b|$ if and only if $b^n = a$

We call $a^{\frac{1}{n}}$ the principal n^{th} root of a. Hence, $a^{\frac{1}{n}}$ is the real number (nonnegative when n is even) that, when raised to the n^{th} power, yields a. Therefore,

$$(16)^{\frac{1}{2}} = 4 \text{ since } 4^2 = 16$$

$$(-125)^{\frac{1}{3}} = -5 \text{ since } (-5)^3 = -125$$

$$\left(\frac{1}{81}\right)^{\frac{1}{4}} = \frac{1}{3} \text{ since } \left(\frac{1}{3}\right)^4 = \frac{1}{81}$$

$$27^{\frac{1}{3}} = 3 \text{ since } 3^3 = 27$$

$$(-16)^{\frac{1}{4}} \text{ is not a real number}$$

Thus far, we have defined $a^{\frac{1}{n}}$, where n is a natural number. With the help of the second rule for exponent, we can define the expression $a^{\frac{m}{n}}$, where m and n are natural numbers and $\frac{m}{n}$ is reduced to lowest terms.

Definition 2.19: (Rational Exponent $a^{\frac{m}{n}}$)

If $a^{\frac{1}{n}}$ is a real number, then $a^{\frac{m}{n}} = \left(a^{\frac{1}{n}}\right)^m$ (i.e. the n^{th} root of a raised to the m^{th} power)

We can also define negative rational exponents:

$$a^{-\frac{m}{n}} = \frac{1}{a^{\frac{m}{n}}} \quad (a \neq 0)$$

Example 2.43: Evaluate the following

a) $27^{\frac{2}{3}}$ b) $36^{-\frac{1}{2}}$ c) $(-32)^{-\frac{3}{5}}$

Solution: We have

a) $27^{\frac{2}{3}} = \left(27^{\frac{1}{3}}\right)^2 = 3^2 = 9$

b) $36^{-\frac{1}{2}} = \frac{1}{36^{\frac{1}{2}}} = \frac{1}{6}$

c) $(-32)^{-\frac{3}{5}} = \frac{1}{(-32)^{\frac{3}{5}}} = \frac{1}{\left((-32)^{\frac{1}{5}}\right)^3} = \frac{1}{(-2)^3} = -\frac{1}{8}$

Radical notation is an alternative way of writing an expression with rational exponents. We define for real number a, the n^{th} root of a as follows:

Definition 2.20 (n^{th} root of a): $\sqrt[n]{a} = a^{\frac{1}{n}}$, where n is a positive integer.

The number $\sqrt[n]{a}$ is also called the principal n^{th} root of a. If the n^{th} root of a exists, we have:

For a a real number and n a positive integer,

$$\sqrt[n]{a^n} = \begin{cases} |a|, & \text{if } n \text{ is even} \\ a, & \text{if } n \text{ is odd} \end{cases}$$

For example, $\sqrt[3]{5^3} = 5$ and $\sqrt[4]{(-3)^4} = 3$.

- **Exponential Functions**

In the previous sections we examined functions of the form $f(x) = x^n$, where n is a constant. How is this function different from $f(x) = n^x$.

Definition 2.21: A function of the form $y = f(x) = b^x$, where $b > 0$ and $b \neq 1$, is called an exponential function.

Example 2.44: The functions $f(x) = 2^x$, $g(x) = 3^x$ and $h(x) = \left(\frac{1}{2}\right)^x$ are examples of exponential functions.

As usual the first question raised when we encounter a new function is its domain. Since rational exponents are well defined, we know that any rational number will be in the domain of an exponential function. For example, let $f(x) = 3^x$. Then as x takes on the rational values $x = 4$, -2, $\frac{1}{2}$ and $\frac{4}{5}$, we have

$$f(4) = 3^4 = 3 \cdot 3 \cdot 3 \cdot 3 = 81 \qquad f(-2) = 3^{-2} = \tfrac{1}{3^2} = \tfrac{1}{9}$$

$$f(\tfrac{1}{2}) = 3^{\frac{1}{2}} = \sqrt{3} \qquad f(\tfrac{4}{5}) = 3^{\frac{4}{5}} = \sqrt[5]{3^4} = \sqrt[5]{81}$$

Note that even though we do not know the exact values of $\sqrt{3}$ and $\sqrt[5]{81}$, we do know exactly what they mean. However, what about $f(x)$ for irrational values of x? For instance, $f(\sqrt{2}) = 3^{\sqrt{2}} = ?$

We have not defined the meaning of irrational exponents. In fact, a precise formal definition of b^x where x is irrational requires the ideas of calculus. However, we can get an idea of what $3^{\sqrt{2}}$ should be by using successive rational approximations to $\sqrt{2}$. For example, we have

$$1.414 < \sqrt{2} < 1.415$$

Thus, it would seem reasonable to expect that $3^{1.414} < 3^{\sqrt{2}} < 3^{1.415}$. Since 1.414 and 1.415 are rational numbers, $3^{1.414}$ and $3^{1.415}$ are well defined, even though we cannot compute their values by hand. Using a calculator, we get $4.7276950 < 3^{\sqrt{2}} < 4.7328918$. If we use better approximations to $\sqrt{2}$, we get $3^{1.4142} < 3^{\sqrt{2}} < 3^{1.4143}$. Using a calculator again, we get $4.7287339 < 3^{\sqrt{2}} < 4.7292535$. Computing $3^{\sqrt{2}}$ directly on a calculator gives $3^{\sqrt{2}} \approx 4.7288044$. This numerical evidence suggests that as x approaches $\sqrt{2}$, the values of 3^x approach a unique real number that we designate by $3^{\sqrt{2}}$, and so we will accept without proof, the fact that the domain of the exponential function is the set of real numbers.

The exponential function $y = b^x$, where $b > 0$ and $b \neq 1$, is defined for all real values of x. In addition all the rules for rational exponents hold for real number exponents as well.

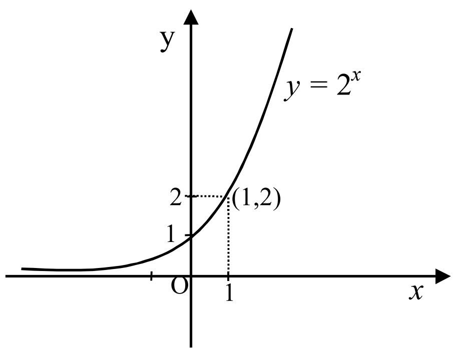
Before we state some general facts about exponential functions , let's see if we can determine what the graph of an exponential function will look like.

Example 2.45:

1. Sketch the graph of the function $y = 2^x$ and identify its domain and range.

   Solution: To aid in our analysis, we set up a short table of values to give us a frame of reference.

x	y
-3	$2^{-3} = \frac{1}{8}$
-2	$2^{-2} = \frac{1}{4}$
-1	$2^{-1} = \frac{1}{2}$
0	$2^0 = 1$
1	$2^1 = 2$
2	$2^2 = 4$
3	$2^3 = 8$

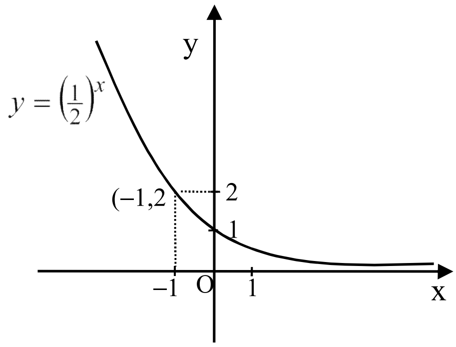


With these points in hand, we draw a smooth curve through the points obtaining the graph appearing above. Observe that the domain of $y = 2^x$ is IR, the graph has no $x-$intercepts, as

$x \to +\infty$, the y values are increasing very rapidly, whereas as $x \to -\infty$, the y values are getting closer and closer to 0. Thus, $x-axis$ is a horizontal asymptote, the $y-$intercept is 1 and the range of $y = 2^x$ is the set of positive real numbers.

2. Sketch the graph of $y = f(x) = \left(\frac{1}{2}\right)^x$.

Solution: It would be instructive to compute a table of values as we did in example 1 above (you are urged to do so). However, we will take a different approach. We note that $y = f(x) = \left(\frac{1}{2}\right)^x = \frac{1}{2^x} = 2^{-x}$. If $f(x) = 2^x$, then $f(-x) = 2^{-x}$. Thus by the graphing principle for $f(-x)$, we can obtain the graph of $y = 2^{-x}$ by reflecting the graph of $y = 2^x$ about the $y-axis$.



Here again the $x-axis$ is a horizontal asymptote, there is no $x-$intercept, 1 is $y-$intercept and the range is the set of positive real numbers. However, the graph is now decreasing rather than increasing.

The following box summarizes the important facts about exponential functions and their graphs.

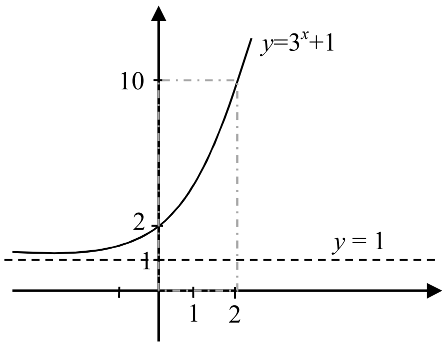
The Exponential function $y = f(x) = b^x$
1. The domain of the exponential function is the set of real numbers 2. The range of the exponential function is the set of positive real numbers 3. The graph of $y = b^x$ exhibits exponential growth if $b > 1$ or exponential decay if $0 < b < 1$. 4. The $y-$intercept is 1. 5. The $x-$intercept is a horizontal asymptote 6. The exponential function is 1 – 1. Algebraically if $b^x = b^y$, then $x = y$

Example 2.46: Sketch the graph of each of the following. Find the domain, range, intercepts, and asymptotes.

a) $y = 3^x + 1$ b) $y = 3^{x+1}$ c) $y = -9^{-x} + 3$

Solution:

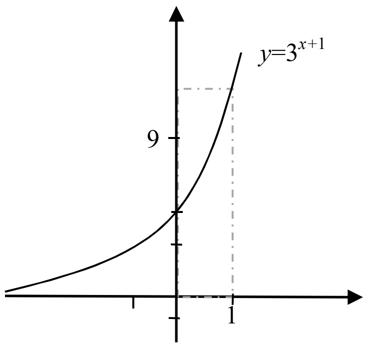
a) To get the graph of $y = 3^x + 1$. We start with the graph of $y = 3^x$, which is the basic exponential growth graph, and shift it up 1 unit.



From the graph we see that

- $Dom(f) = \Re$
- $Range(f) = (1, \infty)$
- The $y-$intercept is 2
- The line $y = 1$ is a horizontal asymptote

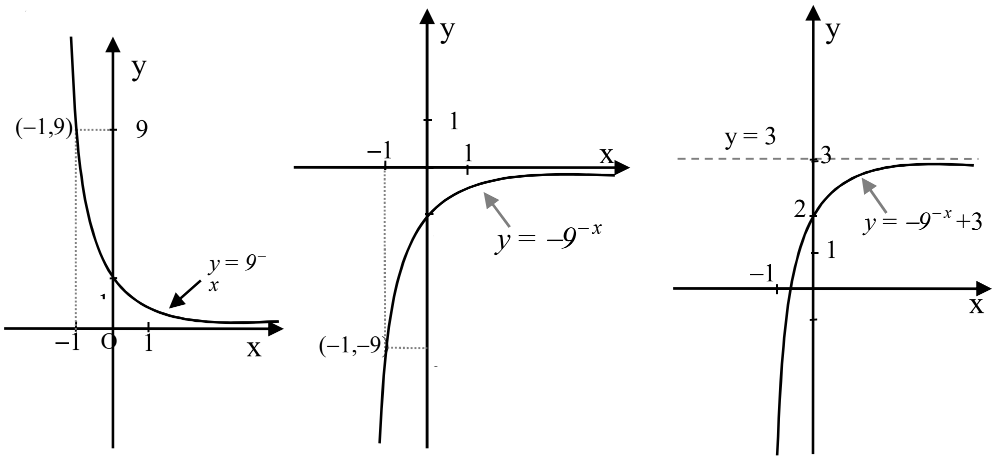
b) To get the graph of $y = 3^{x+1}$, we start with the graph of $y = 3^x$, and shift 1 unit to the left.



From the graph we see that

- $Dom(f) = \Re$
- $Range(f) = (0, \infty)$
- The $y-$intercept is 3
- The line $y = 0$ is a horizontal asymptote

c) To get the graph of $y = -9^{-x} + 3$, we start with the basic exponential decay $y = 9^{-x}$. We then reflect it with respect to the $x-axis$, which gives the graph of $y = -9^{-x}$. Finally, we shift this graph up 3 units to get the required graph of $y = -9^{-x} + 3$.



From the graph of $y=-9^{-x}+3$, we can see that $Dom(h)=\Re$, $Range(h)=(-\infty,3)$, the line $y=3$ is a horizontal asymptote, 2 is the $y-$intercept and $x=-\frac{1}{2}$ is the $x-$intercept.

Remark: When the base b of the exponential function $f(x)=b^x$ equals to the number e, where $e=2.7182\cdots$, we call the exponential function the natural exponential function.

- **Logarithmic Functions**

In the previous subsection we noted that the exponential function $f(x)=b^x$ (where $b>0$ and $b\neq 1$) is one to one. Thus, the exponential function has an inverse function. What is the inverse of $f(x)=b^x$?

To find the inverse of $f(x)=b^x$, let's review the process for finding an inverse function by comparing the process for the polynomial function $y=x^3$ and the exponential function $y=3^x$. Keep in mind that x is our independent variable and y is the dependent variable and so whenever possible we want a function solved explicitly for y.

To find the inverse of $y=x^3$		**To find the inverse of** $y=3^x$	
$y=x^3$	Interchange x and y	$y=3^x$	Interchange x and y
$x=y^3$	solve for y	$x=3^y$	solve for y
$y=\sqrt[3]{x}$		$y=??$	

There is no algebraic procedure we can use to solve $x=3^y$ for y. By introducing radical notations we could express the inverse of $y=x^3$ explicitly in the form $y=\sqrt[3]{x}$. In words, $y^3=x$ and $y=\sqrt[3]{x}$ both mean exactly the same thing: y is the number whose cube is x. Similarly, if we want to express $x=3^y$ explicitly as a function of x, we need to invent a special notation for this. The key idea is to take the equation $x=3^y$ and express it verbally.

$x=3^y$ means y is the exponent to which 3 must be raised to yield x

We introduce the following notation, which expresses this idea in a much more compact form.

Definition 2.22: For $b>0$ and $b\neq 1$, we write $y=\log_b x$ to mean y is the exponent to which b must be raised to yield x. In other words,

$$x=b^y \Leftrightarrow y=\log_b x$$

We read $y = \log_b x$ as "y equals the logarithm of x to the base b".

REMEMBER: $y = \log_b x$ is an alternative way of writing $x = b^y$

When an expression is written in the form $x = b^y$, it is said to be in exponential form. When an expression is written in the form $y = \log_b x$, it is said to be in logarithmic form. The table below illustrates the equivalence of the exponential and logarithmic forms.

Exponential form	Logarithmic form
$4^2 = 16$	$\log_4 16 = 2$
$2^4 = 16$	$\log_2 16 = 4$
$5^{-3} = \frac{1}{125}$	$\log_5 \frac{1}{125} = -3$
$6^{\frac{1}{2}} = \sqrt{6}$	$\log_6 \sqrt{6} = \frac{1}{2}$
$7^0 = 1$	$\log_7 1 = 0$

Example 2.47:

1. Write each of the following in exponential form.

   a) $\log_3 \frac{1}{9} = -2$ b) $\log_{16} 2 = \frac{1}{4}$

   Solution: We have a) $\log_3 \frac{1}{9} = -2$ means $3^{-2} = \frac{1}{9}$.

   b) $\log_{16} 2 = \frac{1}{4}$ means $16^{\frac{1}{4}} = 2$

2. Write each of the following in logarithmic form.

   a) $10^{-3} = 0.001$ b) $27^{\frac{2}{3}} = 9$

   Solution: We have a) $10^{-3} = 0.001$ means $\log_{10} 0.001 = -3$

   b) $27^{\frac{2}{3}} = 9$ means $\log_{27} 9 = \frac{2}{3}$

3. Evaluate each of the following.

   a) $\log_3 81$ b) $\log_8 \frac{1}{64}$

   Solution:

   a) To evaluate $\log_3 81$, we let $t = \log_3 81$, and then rewrite the equation in exponential form, $3^t = 81$. Now, if we can express both sides in terms of the same base, we can solve the resulting exponential equation, as follows:

   Let $t = \log_3 81$ Rewrite in exponential form

   $3^t = 81$ Express both sides in terms of the same base

$3^t = 3^4$ Since the exponential function is 1 – 1

$t = 4$

Therefore, $\log_3 81 = 4$.

b) We apply the same procedure as in part (a).

Let $t = \log_8 \frac{1}{64}$ Rewrite in exponential form

$8^t = \frac{1}{64}$ Express both sides in terms of the same base

$8^t = 8^{-2}$ Since the exponential function is 1 – 1

$t = -2$

Therefore, $\log_8 \frac{1}{64} = -2$.

As was pointed out at the beginning of this subsection, logarithm notation was invented to express the inverse of the exponential function. Thus, $\log_b x$ is a function of x. We usually write $f(x) = \log_b x$ rather than writing $f(x) = \log_b(x)$ and use parenthesis only when needed to clarify the input to the log function. For example,

If $f(x) = \log_5(4-x)$, then $f(-1) = \log_5(4-(-1)) = \log_5 5 = 1$, whereas if $f(x) = 4 - \log_5 x$, then $f(-1) = 4 - \log_5(-1)$, which is undefined.

Example 2.48: Given $f(x) = \log_5 x$, find

a) $f(25)$ b) $f(\frac{1}{25})$ c) $f(0)$ d) $f(-125)$

Solution:

a) $f(25) = \log_5 25 = 2$ (since $5^2 = 25$)

b) $f(\frac{1}{25}) = \log_5 \frac{1}{25} = -2$ (since $5^{-2} = \frac{1}{25}$)

c) $f(0) = \log_5 0$ is not defined (what power of 5 will yield 0?). We say that 0 is not in the domain of f.

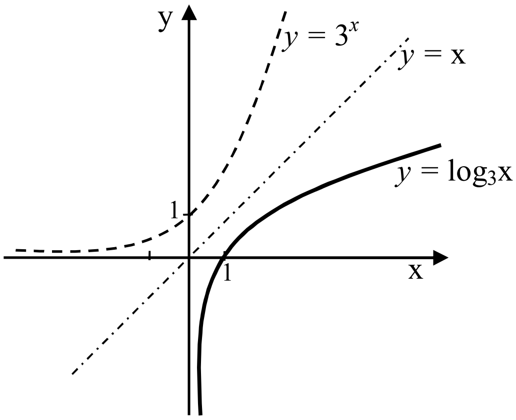
d) $f(-125) = \log_5(-125)$ is not defined (what power of 5 will yield -125?). We say that -125 is not in the domain of f.

Acknowledging that the logarithmic and exponential functions are inverses, we can derive a great deal of information about the logarithmic function and its graph from the exponential function and its graph.

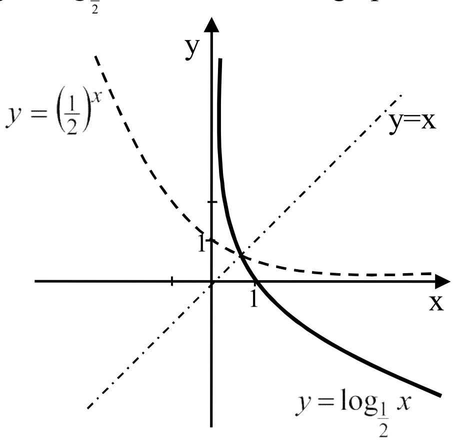
Example 2.49: Sketch the graph of the following functions. Find the domain and range of each.

a) $y = \log_3 x$ b) $y = \log_{\frac{1}{2}} x$

Solution: a) Since $y = \log_3 x$ is the inverse of $y = 3^x$, we can obtain the graph of $y = \log_3 x$ by reflecting the graph of $y = 3^x$ about the line $y = x$, as shown below.



b) To get the graph of $y = \log_{\frac{1}{2}} x$, we reflect the graph of $y = \left(\frac{1}{2}\right)^x$ about the line $y = x$ as shown below.



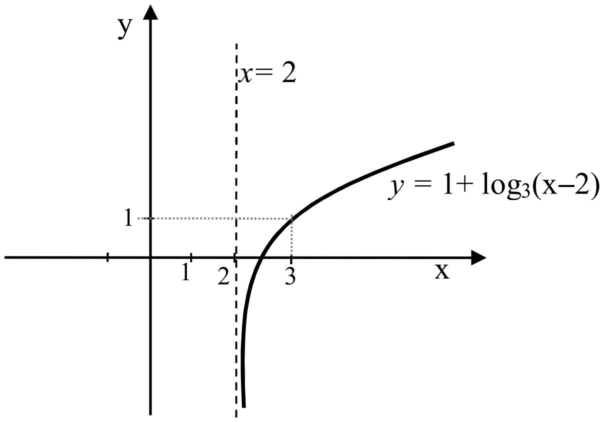
Taking note of the features of the two graphs we have the following important informations about the graph of the log function:

The Logarithmic Function $y = \log_b x$
1. Its domain is the set of positive real numbers 2. Its range is the set of real numbers. 3. Its graph exhibits logarithmic growth if $b > 1$ and logarithmic decay if $0 < b < 1$. 4. The $x-$ intercept is 1. There is no $y-$intercept. 5. The $y-axis$ is a vertical asymptote.

Example 2.50:

1. Sketch the graph of $f(x) = 1 + \log_3(x-2)$. Find the domain, range, asymptote and intercepts.

   Solution: We can obtain the graph of $y = 1 + \log_3(x-2)$ by applying the graphing principle to shift the basic logarithmic growth graph 2 units to the right and 1 unit up.



We have $Dom(f)=\{x: x>2\}$, $Range(f)=\Re$ and the graph has the line $x=2$ as a vertical asymptote. To find the intercept, we set $y=0$ and solve for x. Setting $y=0$ and solving for x, we will obtain $x=\frac{7}{3}$. Thus, the $x-$intercept is $\frac{7}{3}$.

2. Find the inverse function for
   a) $y=f(x)=3^x+4$ b) $y=g(x)=\log_3(x-2)$

   Solution: Following the procedure for finding an inverse function, we have

(a) $y=3^x+4$ Interchange x and y
$x=3^y+4$ solve explicitly for y
$x-4=3^y$ Write in logarithmic form
$y=\log_3(x-4)$
Thus, $f^{-1}(x)=\log_3(x-4)$

(b) $y=\log_3(x-2)$ Interchange x and y
$x=\log_3(y-2)$ Write in logarithmic form
$y-2=3^x$ solve explicitly for y
$y=3^x+2$
Thus, $g^{-1}(x)=3^x+2$

The following table contains the basic properties of logarithm:

Properties of logarithm

Assume that b,u and v are positive and $b\neq 1$. Then

1. $\log_b(uv)=\log_b u+\log_b v$
   In words, logarithm of a product is equal to the sum of the logs of the factors.
2. $\log_b(\frac{u}{v})=\log_b u-\log_b v$
   In words, the log of a quotient is the log of the numerator minus the log of the denominator.
3. $\log_b\left(u^r\right)=r\log_b u$
   In words, the log of a power is the exponent times the log.
4. $\log_b(b^x)=x\log_b b=x$
5. $b^{\log_b x}=x$

Example 2.51:

1. Express in terms of simpler logarithms.

   a) $\log_b(x^3 y)$ b) $\log_b(x^3 + y)$ c) $\log_b\left(\frac{\sqrt{xy}}{z^3}\right)$

   Solution:

   a) $\log_b(x^3 y) = \log_b x^3 + \log_b y = 3\log_b x + \log_b y$

   b) Examining the properties of logarithms, we can see that they deal with log of a product, quotient and power. Thus, $\log_3(x^3 + y)$ which is the log of a sum cannot be simplified using log properties.

   c) We have

   $$\log_b\left(\frac{\sqrt{xy}}{z^3}\right) = \log_b \sqrt{xy} - \log_b(z^3) = \log_b(xy)^{\frac{1}{2}} - 3\log_b z = \frac{1}{2}(\log_b x + \log_b y) - 3\log_b z\,.$$

2. Show that $\log_b \frac{1}{2} = -\log_b 2$.

   Solution: We have $\log_b \frac{1}{2} = \log_b 1 - \log_b 2 = 0 - \log_b 2 = -\log_b 2$.

The logarithmic function was introduced without stressing the particular base chosen. However, there are two bases of special importance in science and mathematics, namely, $b = 10$ and $b = e$.

Definition 2.23: (Common Logarithm)
$f(x) = \log_{10} x$ is called the common logarithm function. We write $\log_{10} x = \log x$.

The inverse of the natural exponential function is called the natural logarithmic function and has its own special notation.

Definition 2.24: (Natural Logarithm)
$f(x) = \log_e x$ is called the natural logarithmic function. We write $\log_e x = \ln x$.

Example 2.52:

1. Evaluate $\log 1000$

   Solution: Let $a = \log 1000$. Then, $a = \log_{10} 1000 = \log_{10}(10^3) = 3$.

2. Find the inverse function of $f(x) = e^x + 1$.

   Solution: Let

$y = e^x + 1$	Interchange x and y
$x = e^y + 1$	Solve for y
$x - 1 = e^y$	Rewrite in logarithmic form
$y = \ln(x-1)$	

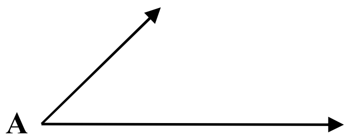
   Thus, $f^{-1}(x) = \ln(x-1)$.

- **Trigonometric functions and their graphs**

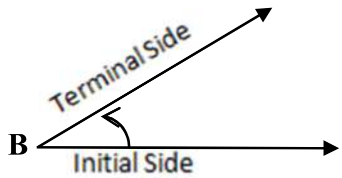
For the functions we have encountered so far, namely polynomial, rational and exponential functions, as the independent variable goes to infinity the graph of each of these three functions either goes to infinity(very quickly) for exponential functions or approaches a finite horizontal asymptote. None of these functions can model the regular periodic patterns that play an important role in the social, biological, and physical sciences: business cycles, agricultural seasons, heart rhythms, and hormone level fluctuations, and tides and planetary motions. The basic functions for studying regular periodic behaviour are the trigonometric functions. The domain of the trigonometric functions is more naturally the set of all geometric angles.

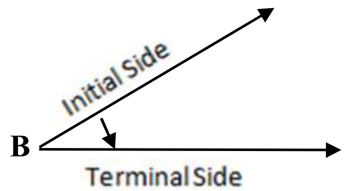
Angle Measurement

An angle is the figure formed by two half-lines or rays with a common end point. The common end point is called the vertex of the angle.



In forming the angle, one side remains fixed and the other side rotates. The fixed side is called the initial side and the side that rotates is called the terminal side. If the terminal side rotates in a counter clockwise direction, we call the angle positive angle, and if the terminal side rotates in a clockwise direction, we call the angle negative angle.

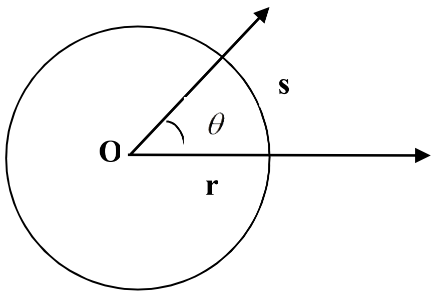




What attribute of an angle are we trying to measure when we measure the size of an angle? A moment of thought will lead us to the conclusion that when we measure an angle we are trying to answer the question: Through what part of a complete rotation has the terminal side rotated?

We will use degree (°) as the unit of measurement for angles. Recall that the measure of a full round angle (full circle) is 360°, straight angle is 180°, and right angle is 90°.

An alternative unit of measure for angles which will indicate their size is the radian measure. To see the connection between the degree measure and radian measure of an angle, let us consider an angle θ and draw a circle of radius r with the vertex of θ at its center O. Let s represent the length of the arc of the circle intercepted by $\angle\theta$ (as shown below).



Basic geometry tells us that the central angle θ will be the same fractional part of one complete rotation as s will be of the circumference of the circle. For example, if θ is $\frac{1}{10}$ of a complete rotation, then s will be $\frac{1}{10}$ of the circumference of the circle. In other words, we can set up the following proportion:

$$\frac{\theta}{1\, complete\, rotation} = \frac{s}{circumference\ of\ circle} = \frac{s}{2\pi r}$$

Thus, we have the following conversion formula:

$$\frac{\theta\ in\ degrees}{180^\circ} = \frac{\theta\ in\ radians}{\pi}$$

Example 2.53:

1. Convert each of the following radian measures to degrees.
   a) $\frac{\pi}{6}$ b) $\frac{3\pi}{5}$

Solution: a) By the conversion formula, we have $\frac{\theta}{180^\circ} = \frac{\frac{\pi}{6}}{\pi}$, which implies that $\theta = 30^\circ$.

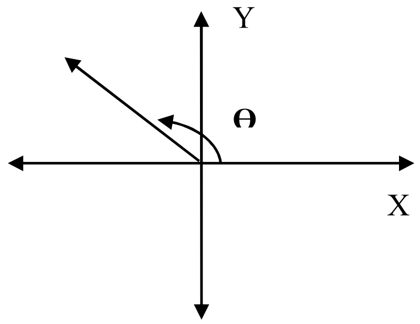
b) Again using the conversion formula, we get $\frac{\theta}{180^\circ} = \frac{\frac{3\pi}{5}}{\pi}$, which implies that $\theta = 108^\circ$.

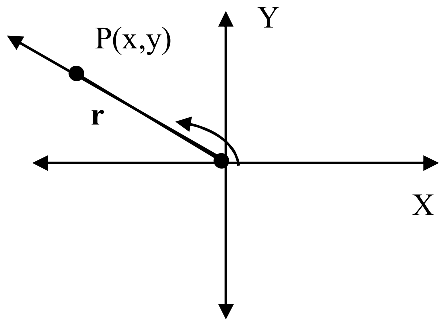
2. Convert to radian measures
   a) 90° b) 270°

Solution: a) Let θ represent the radian measure of 90°. Using the conversion formula, we obtain: $\frac{\theta}{\pi} = \frac{90^\circ}{180^\circ}$, which implies that $\theta = \frac{\pi}{2}$.

b) Rather than using the conversion formula, we notice that $270^\circ = 3(90^\circ)$. In part (a) we found that $90^\circ = \frac{\pi}{2}$, and so we have $270^\circ = \frac{3\pi}{2}$.

To define the trigonometric functions, we will view all angles in the context of a Cartesian coordinate system: that is, given an angle θ, we begin by putting θ in standard position, meaning that the vertex of θ is placed at the origin and initial side of θ is placed along the positive $x-axis$. Thus the location of the terminal side of θ will, of course, depend on the size of θ.





We then locate a point (other than the origin) on the terminal side of θ and identify its coordinates (x, y) and its distance to the origin, dented by r. Then, r is positive.

With θ in standard position, we define the six trigonometric functions of θ as follows:

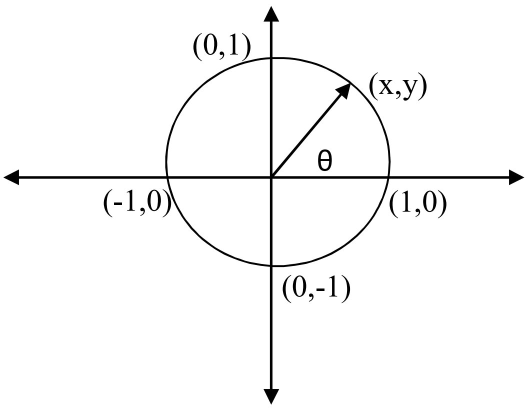
Definition 2.25

Name of function	**Abbreviation**	**Definition**
Sine θ	$\sin\theta$	$\sin\theta = \frac{y}{r}$
Cosine θ	$\cos\theta$	$\cos\theta = \frac{x}{r}$
Tangent θ	$\tan\theta$	$\tan\theta = \frac{y}{x}$
Cosecant θ	$\csc\theta$	$\csc\theta = \frac{r}{y}$
Secant θ	$\sec\theta$	$\sec\theta = \frac{r}{x}$
Cotangent θ	$\cot\theta$	$\cot\theta = \frac{x}{y}$

Recall that the radian measure of an angle is defined as $\theta = \frac{s}{r}$, where θ is angle in radians s is the length of the arc intercepted by θ and r is the length of the radius. Since s and r are both lengths, the quotient $\frac{s}{r}$ is a pure number without any units attached. Thus, any angle can be interpreted as a real number. Conversely, any real number can be interpreted as an angle. Thus, we can describe the domains of the trigonometric functions in the frame work of the real number systems. If we let $f(\theta)=\sin\theta$, the domain consists of all real numbers θ for which $\sin\theta$ is defined. Since $\sin\theta=\frac{y}{r}$ and r is never equal to zero, the domain for $\sin\theta$ is the set of all real numbers. Similarly, the domain of $f(\theta)=\cos\theta=\frac{x}{r}$ is also the set of all real numbers.

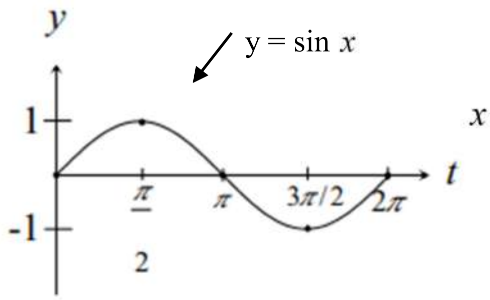
- **The graph of** $y=\sin\theta$

To analyze $f(\theta)=\sin\theta$, we keep in mind that once we choose a real number θ, we draw the angle, in standard position, that corresponds to θ. To simplify our analysis, we choose the point (x,y) on the terminal side so that $r=1$. That is, (x,y) is a point on the unit circle $x^2+y^2=1$. Note that $\sin\theta=\frac{y}{1}=y$.

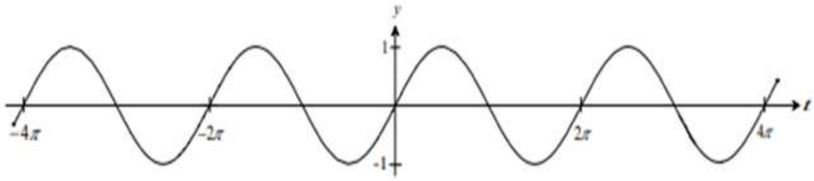


As the terminal side of θ moves through the first quadrant, y increases from 0 (when $\theta=0$) to 1(when $\theta=\frac{\pi}{2}$). Thus, as $_\theta$ increases from 0 to $\frac{\pi}{2}$, $y=\sin\theta$ steadily increases from 0 to 1. As θ increases from $\frac{\pi}{2}$ to π, $y=\sin\theta$ decreases form 1 to 0. A similar analysis reveals that as θ increases from π to $\frac{3\pi}{2}$, $\sin\theta$ decreases from 0 to – 1; and as θ increases from $\frac{3\pi}{2}$ to 2π, $\sin\theta$ increases from – 1 to 0.

Based on this analysis, we have the graph of $f(\theta)=\sin\theta$ in the interval $[0,2\pi]$ as show below.



Since the values of $f(\theta) = \sin\theta$ depend only on the position of the terminal side, adding or subtracting multiples of 2π to θ will leave the value of $f(\theta) = \sin\theta$ unchanged. Thus, the values of $f(\theta) = \sin\theta$ will repeat every 2π units. The complete graph of $f(\theta) = \sin\theta$ appears below.



The graph of $y = \sin x$, which is called the basic sine curve.

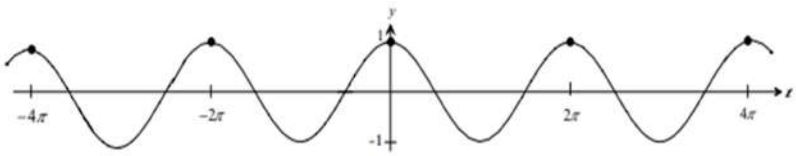
- **The graph of** $y = \cos\theta$

Applying the same type of analysis to $f(\theta) = \cos\theta$, we will able to get a good idea of what its graph looks like. The figure below shows the angle corresponding to θ as it increases through quadrant I, II, III and IV.

Keeping in mind that $\cos\theta = \frac{x}{1} = x$, we have the following:

1. As θ increases from 0 to $\frac{\pi}{2}$, $x = \cos\theta$ decreases from 1 to 0.
2. As θ increases from $\frac{\pi}{2}$ to π, $x = \cos\theta$ decreases from 0 to – 1.
3. As θ increases from π to $\frac{3\pi}{2}$, $x = \cos\theta$ increases from – 1 to 0.
4. As θ increases from $\frac{3\pi}{2}$ to 2π, $x = \cos\theta$ increases from 0 to 1.

Based on this analysis, we have the graph of $f(\theta) = \cos\theta$ as shown below:



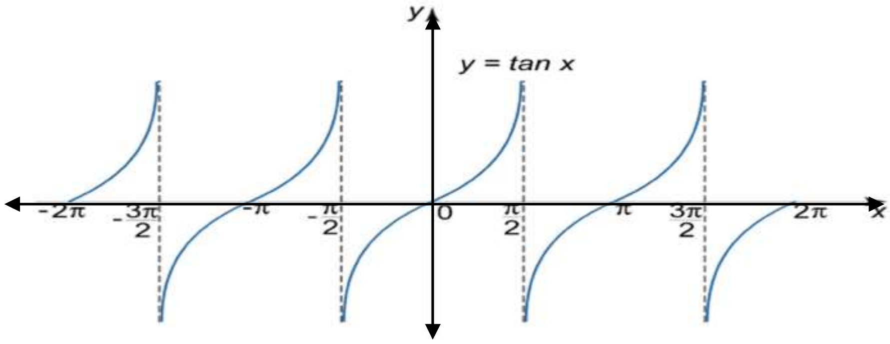
- **The graph of** $y = \tan\theta$

Since $\tan\theta = \frac{y}{x}$ is undefined whenever $x = 0$, $\tan\theta$ is undefined whenever the terminal side of the angle corresponding to θ falls on the $y-axis$. This happens for $\theta = \frac{\pi}{2}$, to which we can

add or subtract any multiple of π that will again bring the terminal side back to the $y-axis$.
Thus, domain of $\tan\theta$ is $\{\theta : \theta \neq \frac{\pi}{2} + n\pi\}$, where n is an integer.

1. As θ increases from 0 to $\frac{\pi}{2}$, x decreases from 1 to 0 and y increases from 0 to 1; therefore, $\tan\theta = \frac{y}{x}$ increases from 0 to ∞.
2. As θ increases from $\frac{\pi}{2}$ to π, x decreases from 0 to – 1 and y decreases from 1 to 0; therefore, $\tan\theta = \frac{y}{x}$ increases from $-\infty$ to 0.
3. As θ increases from π to $\frac{3\pi}{2}$, x increases from – 1 to 0 and y decreases from 0 to – 1; therefore, $\tan\theta = \frac{y}{x}$ increases from 0 to ∞.
4. As θ increases from $\frac{3\pi}{2}$ to 2π, x increases from 0 to 1 and y increases from – 1 to 0; therefore, $\tan\theta = \frac{y}{x}$ increases from $-\infty$ to 0.

You may want to add some more specific values to this analysis. In any case, we get the following as the graph of the tangent function.



Definition 2.26: (Periodic function)
A function $y = f(x)$ is called periodic if there exists a number p such that $f(x+p) = f(x)$ for all x in the domain of f. The smallest such number p is called the period of the function.

A periodic function keeps repeating the same set of $y-values$ over and over again. The graph of a periodic function shows the same basic segment of its graph being repeated. In the case of sine and cosine functions, the period is 2π. The period of the tangent function is π.

Definition 2.27: (Amplitude of a periodic function)
The amplitude of a periodic function $f(x)$ is

$$A = \frac{1}{2}[\text{ maximum value of } f(x) - \text{minimum value of } f(x)]$$

Thus, the amplitude of the basic sine and cosine function is 1.

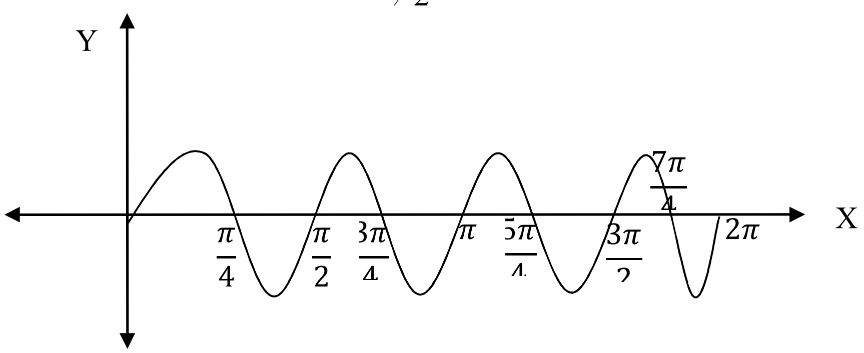
The portion of the graph of a sine or cosine function over one period is called a complete cycle of the graph. In other words, the minimal portion of a sine or cosine graph that keeps repeating itself is called a complete cycle of the graph.

Definition 2.28: (Frequency of a periodic function)

The number of complete cycles a sine or cosine graph makes on an interval of length equal to 2π is called its frequency.

The frequency of the basic sine curve $y = \sin x$ and the basic cosine curve $y = \cos x$ is 1, because each graph makes 1 complete cycle in the interval $[0,2\pi]$.

If a sine function has period of $\frac{\pi}{2}$ (see the figure below), then the number of complete cycles its graph will make in an interval of length 2π is $\frac{2\pi}{\pi/2} = 4$.



Thus if a sine function has a period of $\frac{\pi}{2}$, its frequency is 4 and its graph will make 4 complete cycles in an interval of length 2π.

Example 2.54: Sketch the graph of $y = \sin 2x$ and find its amplitude, period and frequency.

Solution: We can obtain this graph by applying our knowledge of the basic sine graph. For the basic curve, we have

$\sin 0 = 0 \qquad \sin\frac{\pi}{2} = 1 \qquad \sin\pi = 0 \qquad \sin\frac{3\pi}{2} = -1 \qquad \sin 2\pi = 0$

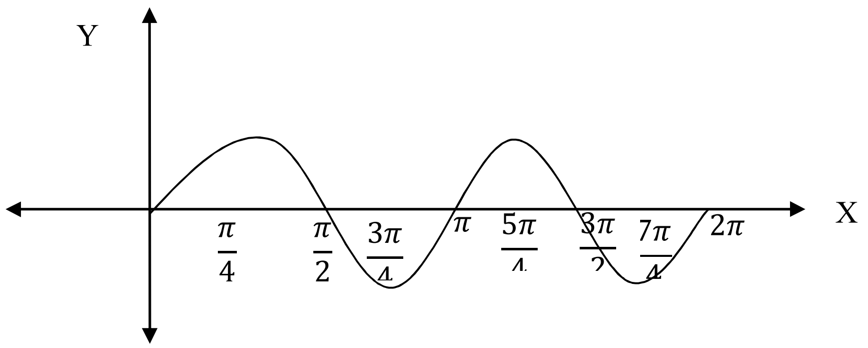
These quadrantal values serve as guide points, which help us draw the graph. To obtain similar guide points for $y = \sin 2x$, we ask for what values of x is

$2x = 0 \qquad 2x = \frac{\pi}{2} \qquad 2x = \pi \qquad 2x = \frac{3\pi}{2} \qquad 2x = 2\pi$

and we get

$x = 0 \qquad x = \frac{\pi}{4} \qquad x = \frac{\pi}{2} \qquad x = \frac{3\pi}{2} \qquad x = \pi$

Thus, $y = \sin 2x$ will have the values 0, 1, 0, -1, 0 at $x = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$, and π, respectively. The graph of $y = \sin 2x$ will thus complete one cycle in the interval $[0, \pi]$, and will repeat the same values in the interval $[\pi, 2\pi]$.



From this graph we see that $y = \sin 2x$ has an amplitude of 1, a period π, and a frequency of 2.

For convenience we summarize our discussion on the domains of the trigonometric functions in the table.

1. $f(x) = \sin x$	Domain = All real numbers
2. $f(x) = \cos x$	Domain = All real numbers
3. $f(x) = \tan x$	Domain = $\{x : x \neq \frac{\pi}{2} + n\pi\}$
4. $f(x) = \csc x$	Domain = $\{ x : x \neq n\pi \}$
5. $f(x) = \sec x$	Domain = $\{x : x \neq \frac{\pi}{2} + n\pi\}$
6. $f(x) = \cot x$	Domain = $\{x : x \neq n\pi\}$
	where n is an integer

We have the following trigonometric identities

1. $\sin^2 x + \cos^2 x = 1$
2. $\tan^2 x + 1 = sce^2 x$
3. $1 + \cot^2 x = \csc^2 x$

Exercise 2.7

1. Find the domain of the given function.

   a) $f(x) = \dfrac{1}{6^x}$ b) $g(x) = \sqrt{3^x + 1}$ c) $h(x) = \sqrt{2^x - 8}$ d) $f(x) = \dfrac{1}{2^{3x} - 2}$

2. Sketch the graph of the given function. Identify the domain, range, intercepts, and asymptotes.

   a) $y = 5^{-x}$ b) $y = 9 - 3^x$ c) $y = 1 - e^{-x}$ d) $y = e^{x-2}$

3. Solve the given exponential equation.

a) $2^{x-1}=8$ b) $3^{2x}=243$ c) $8^{x}=\sqrt{2}$ d) $16^{3a-2}=\frac{1}{4}$

4. Let $f(x)=2^{x}$. Show that $f(x+3)=8f(x)$.

5. Let $g(x)=5^{x}$. Show that $g(x-2)=\dfrac{1}{25}g(x)$.

6. Let $f(x)=3^{x}$. Show that $\dfrac{f(x+2)-f(2)}{2}=4(3^{x})$.

7. Evaluate the given logarithmic expression (where it is defined).

a) $\log_2 32$ c) $\log_3(-9)$ e) $\log_5(\log_3 243)$

b) $\log_{\frac{1}{3}} 9$ d) $\log_6 \frac{1}{\sqrt{6}}$ f) $2^{\log_2 \sqrt{5}}$

8. If $f(x)=\log_2(x^2-4)$, find $f(6)$ and the domain of f.

9. If $g(x)=\log_3(x^2-4x+3)$, find $f(4)$ and the domain of g.

10. Show that $\log_{\frac{1}{6}} x=-\log_6 x$

11. Sketch the graph of the given function and identify the domain, range, intercepts and asymptotes.

a) $f(x)=\log_2(x-3)$ b) $f(x)=-3+\log_2 x$ c) $f(x)=-\log_3(-x)$ d) $f(x)=3\log_5 x$

12. Find the inverse of $f(x)=e^{(3x-1)}$.

13. Let $f(x)=e^{\sqrt{x}}$. Find a function so that $(f\circ g)(x)=(g\circ f)(x)=x$.

14. Convert the given angle from radians to degrees

a) $\frac{\pi}{3}$ b) $-\frac{5\pi}{2}$ c) $-\frac{4\pi}{3}$

15. Convert the given angle from degrees to radians

a) 315° b) -40° c) 330°

16. Sketch the graph of

a) $f(\theta)=\sec\theta$ c) $f(\theta)=\csc\theta$ e) $f(\theta)=\cot\theta$

b) $f(x)=1+\cos x$ d) $f(x)=\sin(x+\frac{\pi}{2})$ f) $f(x)=\tan 2x$

17. Verify the following identities:

a) $(\sin x-\cos x)(\csc x+\sec x)=\tan x-\cot x$

b) $\sec^2 x-\csc^2 x=\tan^2 x-\cot^2 x$

18. Given $\tan\theta=\frac{1}{2}$ and $\sin\theta<0$, find $\cos\theta$.

Chapter Three

Matrices, Determinant and Systems of Linear Equation

Matrices, which are also known as rectangular arrays of numbers or functions, are the main tools of linear algebra. Matrices are very important to express large amounts of data in an organized and concise form. Furthermore, since matrices are single objects, we denote them by single letters and calculate with them directly. All these features have made matrices very popular for expressing scientific and mathematical ideas. Moreover, application of matrices are found in most scientific fields; such as economics, finance, probability theory and statistics, computer science, engineering, physics, geometry, and other areas.

Main Objectives of this Chapter
At the end of this chapter, students will be able to:-

- Understand the notion of matrices and determinants
- Use matrices and determinants to solve system of linear equations
- Apply matrices and determinants to solve real life problems

3.1 Definition of Matrix

Consider an automobile company that manufactures two types of vehicles, Trucks and Passenger cars in two different colors, red and blue. The company's sales for the month of January are 15 Trucks and 20 Passenger cars in red color, and 10 Trucks and 16 Passenger cars in blue color. This data is presented in Table 1.

Table 1

	Trucks	Passenger Cars
Red	15	20
Blue	10	16

The information in the table can be given in the form of rectangular arrays of numbers as

$$\begin{array}{c c} & \begin{array}{cc} C_1 & C_2 \end{array} \\ \begin{array}{c} R_1 \\ R_2 \end{array} & \begin{bmatrix} 15 & 20 \\ 10 & 16 \end{bmatrix} \end{array}.$$

In this arrangement, the horizontal and vertical lines of numbers are called **rows** (R_1, R_2) and **columns** (C_1, C_2), respectively. The columns C_1 and C_2 represent the Trucks and Passenger cars, respectively, which are sold in January. And the rows R_1 and R_2 represent the red and blue colored vehicles, respectively. An arrangement of this type is called a **matrix**. Note that the above matrix has two rows and two columns. This shows us the usefulness of matrix to organize information.

Definition 3.1 (**Definition of Matrix**). *If m and n are positive integers, then by a **matrix** of size m by n, or an $m \times n$ matrix, we shall mean a rectangular array consisting of mn numbers, or symbols, or expressions in a boxed display consisting of m rows and n columns. This can be denoted by*

$$\begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_m \end{array} \begin{array}{c} \begin{array}{ccccc} C_1 & C_2 & C_3 & & C_n \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \end{array}$$

where $(R_1, R_2, R_3, ..., R_m)$ and $(C_1, C_2, C_3, ..., C_n)$ represent the m rows and n columns, respectively.

Remark.

1. Note that the first suffix denotes the number of a row (or position) and the second suffix that of a column, so that a_{ij} appears at the intersection of the i-th row and the j-th column.

2. Matrix A of size $m \times n$ may also be expressed by

$$A = [a_{ij}]_{m \times n},$$

where a_{ij} represents the (i, j)-th entry of the matrix $[a_{ij}]$.

Example 3.1. The following are matrices of different size.

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a 2×2 matrix $\quad B = \begin{bmatrix} a & b & c \\ b & c & d \\ c & d & e \end{bmatrix}$ is 3×3 matrix

$C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{bmatrix}$ is 4×2 matrix $\quad D = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ is 4×1 matrix

$E = \begin{bmatrix} a & b & c & d \\ b & c & d & e \end{bmatrix}$ is 2×4 matrix, $\quad F = \begin{bmatrix} b & c & d & e \end{bmatrix}$ is 1×4 matrix

Definition 3.2. *Matrices which are $n \times 1$ or $1 \times n$ are called **vectors**. Thus, the $n \times 1$ matrix*

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{2n} \end{bmatrix}$$

is called a column vector, and the $1 \times n$ matrix

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \end{bmatrix}$$

is called a row vector.

Definition 3.3 (Submatrix). *Let A be an $m \times n$ matrix. A submatrix of matrix A is any matrix of size $r \times s$ with $r \leq m$ and $s \leq n$, which is obtained by deleting any collection of rows and/or columns of matrix A.*

Example 3.2. For the given matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$,

(i) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$ is a submatrix of A, which is obtained by deleting the third row of A.

(ii) $\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 5 \end{bmatrix}$ is a submatrix of A, which is obtained by deleting the second column of A.

(iii) $\begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}$ is a submatrix of A, which is obtained by deleting the first column and first row of A.

Definition 3.4 (Equality of Matrices). *Two matrices of the same size, $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{m \times n}$, are said to be equal (and write $A = B$) if and only if*

$$a_{ij} = b_{ij}, \quad \textit{for all } ij.$$

Example 3.3.

(a) Determine the values of a, b, c and d for which the matrices A and B are equal:

$$A = \begin{bmatrix} 5 & 4 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Solution: By Definition 3.4, we have $a_{11} = b_{11}$ implies $a = 5$, $a_{12} = b_{12}$ implies $b = 4$, $a_{21} = b_{21}$ implies $c = 0$ and $a_{22} = b_{22}$ implies $d = 2$.

(b) Find the values of α and β for which the given matrices A and B are equal.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} \alpha - \beta & 2 \\ \alpha & -1 \end{bmatrix}$$

Solution: Similarly, we have $a_{11} = b_{11}$ implies $\alpha - \beta = 1$, $a_{21} = b_{21}$ implies $\alpha = 3$, and hence $\beta = 2$.

Definition 3.5 (Zero Matrix). *An $m \times n$ matrix $A = [a_{ij}]$ is said to be the zero matrix if $a_{ij} = 0$ for all ij.*

Example 3.4. The following are zero matrices.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Exercise 3.1.

1. Write out the matrix of size 3×3 whose entries are given by $x_{ij} = i + j$.

2. Write out the matrix of size 4×4 whose entries are given by

$$x_{ij} = \begin{cases} 1 & if\ i > j \\ 0 & if\ i = j \\ -1 & if\ i < j. \end{cases}$$

3. For the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$, give all the submatrices of size 2×2.

3.2 Matrix Algebra

In this section, we discuss addition of matrices, scalar multiplication, and matrix multiplication.

3.2.1 Addition and Scalar Multiplication

Addition and scalar multiplication are the basic matrix operations. To see the usefulness of these operations, let us observe the following simple application.

Consider again an automobile company that manufactures two types of vehicles, Trucks and Passenger cars in two different colors, red and blue. If the sales for the months of January and February, respectively, are given by

$$J = \begin{bmatrix} 15 & 20 \\ 10 & 16 \end{bmatrix} \text{ and } F = \begin{bmatrix} 12 & 28 \\ 20 & 14 \end{bmatrix},$$

then the total sales for two months can be given as follows. The total number of red Trucks sold in two months is $15 + 12 = 27$. Similarly, the total number of blue Trucks, red Passenger cars and blue Passenger cars sold in the two months are given by $10 + 20 = 30$, $20 + 28 = 48$ and $16 + 14 = 30$, respectively.

The preceding computations are examples of matrix addition. We can write the sum of two 2×2 matrices indicating the sales of January and February as

$$J + F = \begin{bmatrix} 15 & 20 \\ 10 & 16 \end{bmatrix} + \begin{bmatrix} 12 & 28 \\ 20 & 14 \end{bmatrix} = \begin{bmatrix} 15+12 & 20+28 \\ 10+20 & 14+16 \end{bmatrix} = \begin{bmatrix} 27 & 48 \\ 30 & 30 \end{bmatrix}.$$

Definition 3.6. *Let $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{m\times n}$ be two matrices of the same size. Then the sum of A and B, denoted by $A + B$, is the $m \times n$ matrix defined by the formula*

$$A + B = [a_{ij} + b_{ij}].$$

The sum of two matrices of different sizes is undefined.

Example 3.5. For the given matrices A, B, C, D compute $A + B$ and $C + D$.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} w & x \\ y & z \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 4 \\ -1 & 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & 3 \end{bmatrix}$$

Solution: Using Definition 3.6, we have

$$A + B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a+w & b+x \\ c+y & d+z \end{bmatrix}$$

and

$$C + D = \begin{bmatrix} 2 & 0 & 4 \\ -1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 4 \\ -1 & -1 & 4 \end{bmatrix}.$$

Theorem 3.1 (Laws of Matrix Addition). *Let A, B, C be matrices of the same size $m \times n$, **0** the $m \times n$ zero matrix. Then*

1. ***Closure Law of Addition:*** *$A + B$ is an $m \times n$ matrix.*
2. ***Associative Law:*** *$(A + B) + C = A + (B + C)$.*
3. ***Commutative Law :*** *$A + B = B + A$.*
4. ***Identity Law :*** *$A + \mathbf{0} = A$.*
5. ***Inverse Law :*** *$A + (-A) = \mathbf{0}$.*

Definition 3.7 (Scalar Multiplication). *Let $A = [a_{ij}]$ be an $m \times n$ matrix, and α a scalar. Then the product of the scalar α with matrix A, denoted by αA, is defined by*

$$\alpha A = [\alpha a_{ij}]_{m \times n}.$$

Example 3.6. Consider the automobile manufacturing company once again. Suppose the company's sales for the months of January and March, respectively, are given by

$$J = \begin{bmatrix} 15 & 20 \\ 10 & 16 \end{bmatrix}, \text{ and } M = \begin{bmatrix} 18 & 22 \\ 14 & 20 \end{bmatrix}.$$

(a) If the sales of January is to be doubled in February, then the sales of February should be

$$2J = \begin{bmatrix} 2(15) & 2(20) \\ 2(10) & 2(16) \end{bmatrix} = \begin{bmatrix} 30 & 40 \\ 20 & 32 \end{bmatrix}.$$

(b) If the sales of March is to be declined by 50% in April, then the sales of April should be

$$(\frac{1}{2})J = \begin{bmatrix} \frac{1}{2}(18) & \frac{1}{2}(22) \\ \frac{1}{2}(14) & \frac{1}{2}(20) \end{bmatrix} = \begin{bmatrix} 9 & 11 \\ 7 & 10 \end{bmatrix}.$$

Example 3.7. Given the matrices A and B, compute $4A$ and $A + (-1)B$.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$$

Solution: Using Definition 3.7, we have

$$4A = 4 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 4(1) & 4(2) \\ 4(3) & 4(4) \end{bmatrix} = \begin{bmatrix} 4 & 8 \\ 12 & 16 \end{bmatrix}.$$

And, from the definitions 3.6 and 3.7, we have

$$A + (-1)B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + (-1) \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -2 & -4 \\ -1 & -3 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 2 & 1 \end{bmatrix}.$$

From this example, we observe that the difference of two matrices A and B, which is denoted by $A - B$, can be defined by the formula

$$A - B = A + (-1)B = [a_{ij} - b_{ij}]_{m \times n}.$$

Theorem 3.2 (Laws of Scalar Multiplication). *Let A, B be matrices of the same size $m \times n$, and α and β scalars. Then*

1. ***Closure Law of Scalar Multiplication:*** *αA is an $m \times n$ matrix.*
2. ***Associative Law:*** *$\alpha(\beta A) = (\alpha\beta)A$.*
3. ***Distributive Law:*** *$\alpha(A + B) = \alpha A + \alpha B$.*
4. ***Distributive Law:*** *$(\alpha + \beta)A = \alpha A + \beta A$.*
5. ***Monoidal Law:*** *$1A = A$.*

Example 3.8. Let

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be the given matrices. Then,

$$2(A + B) = 2 \begin{bmatrix} 1+2 & 2+0 \\ 0+1 & 1+1 \end{bmatrix} = 2 \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} (2)3 & (2)2 \\ (2)1 & (2)2 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ 2 & 4 \end{bmatrix}$$

and

$$2A + 2B = \begin{bmatrix} (2)1 & (2)2 \\ (2)0 & (2)1 \end{bmatrix} + \begin{bmatrix} (2)2 & (2)0 \\ (2)1 & (2)1 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ 2 & 4 \end{bmatrix}.$$

Thus, we have $2(A + B) = 2A + 2B$.

Example 3.9. Solve for X in the matrix equation $2X + A = B$, where

$$A = \begin{bmatrix} 4 & 0 \\ -2 & 2 \end{bmatrix}, \text{ and } B = \begin{bmatrix} 6 & -4 \\ 8 & 0 \end{bmatrix}.$$

Solution: We begin by solving the equation for X to obtain

$$2X = B - A \text{ implies } X = (\frac{1}{2})(B - A).$$

Thus, we have the solution

$$X = \frac{1}{2} \begin{bmatrix} 6-4 & -4-0 \\ 8-(-2) & 0-2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & -4 \\ 10 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 5 & -1 \end{bmatrix}.$$

3.2.2 Matrix Multiplication

An other important matrix operation is matrix multiplication. To see the usefulness of this operation, consider the application below, in which matrices are helpful for organizing information.

A football stadium has three concession areas, located in South, North and West stands. The top-selling items are, peanuts, hot dogs and soda. Sales for one day are given in the first matrix below, and the prices (in dollar) of the three items are given in the second matrix (note that the price per Peanuts, Hot dogs and Soda are given by \$2.00, \$3.00 and \$2.75, respectively).

$$\begin{array}{r} \\ \text{South Stand} \\ \text{North Stand} \\ \text{West Stand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Peanuts} & \text{Hot dogs} & \text{Sodas} \end{array} \\ \begin{bmatrix} 120 & 250 & 305 \\ 207 & 140 & 419 \\ 29 & 120 & 190 \end{bmatrix} \end{array} \begin{bmatrix} 2.00 \\ 3.00 \\ 2.75 \end{bmatrix}.$$

To calculate the total sales of the three top-selling items at the South stand, multiply each entry in the first row of the matrix on the left by the corresponding entry in the price column matrix on the right and add the results. Thus, we have

$$120(2.00) + 250(3.00) + 305(2.75) = 1828.75\$ \text{ (South stand sales)}.$$

Similarly, the sales for the other two stands are given below:

$$207(2.00) + 140(3.00) + 419(2.75) = 1986.25\$ \text{ (North stand sales)}.$$

$$29(2.00) + 120(3.00) + 190(2.75) = 940.5\$ \text{ (West stand sales)}.$$

The preceding computations are examples of matrix multiplication. We can write the product of the 3×3 matrix indicating the number of items sold and the 3×1 matrix indicating the selling prices as shown below.

$$\begin{bmatrix} 120 & 250 & 305 \\ 207 & 140 & 419 \\ 29 & 120 & 190 \end{bmatrix} \begin{bmatrix} 2.00 \\ 3.00 \\ 2.75 \end{bmatrix} = \begin{bmatrix} 1828.75 \\ 1986.25 \\ 940.5 \end{bmatrix}$$

The product of these matrices is the 3×1 matrix giving the total sales for each of the three stands.

Definition 3.8 (Matrix Multiplication). *Let $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{n\times p}$ be two matrices. Then the product of A and B, denoted by AB, is an $m \times p$ matrix whose (i, j)-th entry is defined by the formula*

$$[AB]_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + ... + a_{in}b_{nj.}$$

In the other words, the (i, j)-th entry of the product AB is obtained by summing the products of the elements in the i-th row of A with corresponding elements in the j-th column of B.

The above definition can be understood as follows. If

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots a_{1n} \end{bmatrix}$$

has only one row (R_1), and

$$B = \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix}$$

has only one column (C_1), then product AB is given by

$$AB = [R_1C_1] = \begin{bmatrix} a_{11} & a_{12} & \dots a_{1n} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} = a_{11}b_{11} + a_{12}b_{21} + ... + a_{1n}b_{n1}.$$

If A has m rows $R_1, R_2, ..., R_m$, and B has n columns $C_1, C_2, ..., C_p$, then the product AB can be given by the formula

$$AB = \begin{bmatrix} R_1C_1 & R_1C_2 & \dots & R_1C_p \\ R_2C_1 & R_2C_2 & \dots & R_2C_p \\ \vdots & \vdots & \dots & \vdots \\ R_mC_1 & R_mC_2 & \dots & R_mC_p \end{bmatrix}.$$

That is, the (i, j)-th entry of AB is R_iC_j.

Remark. The product AB of two matrices A and B is defined only if the number of columns in A and the number of rows in B are equal.

Example 3.10. Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$ be two matrices. Clearly, the product AB is defined in this case, since the number of column of A and the number of rows of B are equal. Thus, we have

$$AB = \begin{bmatrix} R_1C_1 & R_1C_2 \\ R_2C_1 & R_2C_2 \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}.$$

In this example, the matrices A and B, respectively, are 2×3 and 3×2 matrices, whereas the product AB is a 2×2 matrix.

Example 3.11. Compute the product AB of the given matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix}.$$

Solution: The product AB is defined since the number of columns in matrix A and the number of rows in matrix B are equal. Thus, we have AB is given by

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} (1)(1) + (2)(1) + (3)(1) & (1)(1) + (2)(-1) + (3)(2) \end{bmatrix} = \begin{bmatrix} 6 & 5 \end{bmatrix}.$$

Note that the product BA is not defined in this case.

Example 3.12. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ be the given matrices. Then, we have

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } BA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

In this example, we observe that both the products AB and BA are defined. This is true in general i.e., the products AB and BA are defined for any two square matrices A and B of the same size. For the matrices A and B given above, we have $AB \neq BA$. Hence, matrix multiplication is **not commutative**.

Example 3.13. Consider the following diagonal matrices.

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}, \text{ and } B = \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}$$

The product AB is given by

$$AB = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & 0 & 0 \\ 0 & a_{22}b_{22} & 0 \\ 0 & 0 & a_{33}b_{33} \end{bmatrix}$$

Similarly, we have

$$BA = \begin{bmatrix} b_{11}a_{11} & 0 & 0 \\ 0 & b_{22}a_{22} & 0 \\ 0 & 0 & b_{33}a_{33} \end{bmatrix}.$$

In this case, we have $AB = BA$, and hence the given matrices A and B commute. More generally, if A and B are any two diagonal matrices of the same size, then $AB = BA$.

Theorem 3.3. *Matrix multiplication is associative, i.e., whenever the products are defined, we have* $A(BC) = (AB)C$.

From Theorem 3.3, we shall write ABC for either $A(BC)$ or $(AB)C$. Also, for every positive integer n, we shall write A^n for the product $AAA...A$ (n terms).

Theorem 3.4. *If all multiplications and additions make sense, the following hold for matrices, A, B, C and α, β scalars.*

1. $A(\alpha B + \beta C) = \alpha(AB) + \beta(AC)$.
2. $(\alpha B + \beta C)A = \alpha(BA) + \beta(CA)$.

Exercise 3.2.

1. Find your own examples:
   (i) 2×2 matrices A and B such that $A \neq 0$, $B \neq 0$ with $AB \neq BA$.
   (ii) 2×2 matrices A and B such that $A \neq 0$, $B \neq 0$ but $AB = 0$.
   (iii) 2×2 matrix A such that $A^2 = I_2$ and yet $A \neq I_2$ and $A \neq -I_2$.
2. Let $A = \begin{bmatrix} -1 & -1 \\ 3 & 3 \end{bmatrix}$. Find all 2×2 matrices, B such that $AB = 0$.
3. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 1 & c \end{bmatrix}$. Is it possible to choose c so that $AB = BA$? If so, what should be the value of c?
4. Given the matrices $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}$ and α a scalar
   i. Compute the products $A(BC)$, $(AB)C$, and verify that $A(BC) = (AB)C$.
   ii. Compute the products $\alpha(AB)$, $(\alpha A)B$, $A(\alpha B))$, and verify that

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

5. Consider the automobile producer whose agency's sales for the month of January were given by

$$J = \begin{bmatrix} 15 & 20 \\ 10 & 16 \end{bmatrix}.$$

Suppose that the price of a Truck is \$200 and that of a Passenger car is \$100. Use matrix multiplication to find the total values of the red and blue vehicles for the month of January.

3.3 Types of Matrices

There are certain types of matrices that are so important that they have acquired names of their own. In this section we are going to discuss some of these matrices and their properties.

Definition 3.9 (Square Matrix). *A matrix A is said to be square if it has the same number of rows and columns. If A has n-rows and n-columns, we call it a square matrix of size n.*

Example 3.14. The following are square matrices.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ (Square matrix of size 2)}$$

$$B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 4 & 2 & -2 \end{bmatrix} \text{ (Square matrix of size 3)}$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \text{ (Square matrix of size } n\text{)}$$

Definition 3.10 (Identity Matrix). *A square matrix $A = [a_{ij}]_{n\times n}$ is called an identity matrix if*

$$a_{ij} = \begin{cases} 1, & if\ i = j \\ 0, & \textit{otherwise} \end{cases}$$

and it is denoted by I_n.

Example 3.15. The following are identity matrices.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ (Identity matrix of size 2)}$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ (Identity matrix of size 3)}$$

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \text{ (Identity matrix of size } n\text{)}$$

Definition 3.11 (Diagonal Matrix). *A square matrix $D = [d_{ij}]_{n\times n}$ is said to be diagonal if $d_{ij} = 0$ whenever $i \neq j$. Less formally, D is said to be diagonal when all the entries off the main diagonal are 0.*

Example 3.16. The following are diagonal matrices.

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ (Diagonal matrix of size 2)}$$

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix} \text{ (Diagonal matrix of size 3)}$$

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix} \text{ (Diagonal matrix of size 3)}$$

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ (Diagonal matrix of size 3)}$$

Note that the identity matrix is the special case of diagonal matrix where all the entries in the main diagonal are 1.

Definition 3.12 (Scalar Matrix). *A diagonal matrix in which all diagonal entries are equal is called a* ***scalar matrix****.*

Example 3.17. The following are scalar matrices.

$$(a) \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \quad (b) \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad (c) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition 3.13 (Triangular Matrix). *A square matrix $A = [a_{ij}]_{n\times n}$ is said to be* ***lower triangular*** *if and only if $a_{ij} = 0$ whenever $i < j$. A is said to be* ***upper triangular*** *if and only if $a_{ij} = 0$ whenever $i > j$.*

Example 3.18.

(i) $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 7 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (Upper triangular matrices).

(ii) $\begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ -2 & 4 & 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (Lower triangular matrices).

Remark.

(a) In the lower triangular matrix all the entries above the main diagonal are zero, whereas in the upper triangular matrix all the entries below the main diagonal are zero.

(b) Any diagonal matrix is both upper and lower triangular.

Definition 3.14 (Transpose of Matrix). *Let $A = [a_{ij}]$ be an $m \times n$ matrix . Then by the transpose of A we mean the $n \times m$ matrix, denoted by A^t, whose (i, j)-th entry is the (j, i)-th entry of A. More precisely, if $A = [a_{ij}]_{m\times n}$, then $A^t = [a_{ji}]_{n\times m}$. That is,*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \text{ then } A^t = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \dots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nm} \end{bmatrix}.$$

Note that the k-th row of matrix A becomes k-th column of A^t, and the k-th column of A becomes k-th row of A^t.

Example 3.19. Compute the transposes of the following matrices.

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & -3 \\ 3 & -3 & 7 \end{bmatrix}$$

Solution: First let us consider matrix A. Now, row 1 of matrix A becomes column 1 of A^t, and row 2 of A becomes column 2 of A^t. Thus, we have

$$A^t = \begin{bmatrix} 1 & 1 \\ -1 & 2 \\ -1 & 3 \end{bmatrix}.$$

Similarly,

$$B^t = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & -3 \\ 3 & -3 & 7 \end{bmatrix}.$$

Definition 3.15 (Symmetric Matrix). *A square matrix A is said to be **Symmetric** if $A = A^t$.*

Example 3.20. Distinguish whether the given matrix is symmetric or not.

$$(a)\ \ A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & -3 \\ 3 & -3 & 7 \end{bmatrix} \quad (b)\ \ B = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 3 & 2 & 3 \end{bmatrix}$$

Solution:

(a) For the matrix $A = \begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & 2 \\ -3 & -2 & 0 \end{bmatrix}$, $A^t = \begin{bmatrix} 0 & -1 & -3 \\ 1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix}$. Thus, we have $A \neq A^t$, and hence A is not symmetric.

(b) For the matrix $B = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 3 & 2 & 3 \end{bmatrix}$, $B^t = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 3 & 2 & 3 \end{bmatrix}$. Thus, we have $B = B^t$, and hence B is symmetric.

Theorem 3.5 (Properties of Matrix Transpose). *When the relevant sums and products are defined, and α is a scalar. Then*

1. $(A^t)^t = A$.

2. $(A + B)^t = A^t + B^t$.

3. $(\alpha A)^t = \alpha(A^t)$.

3. $(AB)^t = B^t A^t$.

Exercise 3.3.

For the given matrices $A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}$, and $B = \begin{bmatrix} 3 & -2 \\ 0 & 1 \end{bmatrix}$:

(a) Show that $(A^t)^t = A$.

(b) Show that $(A + B)^t = A^t + B^t$.

(c) Show that $(4A)^t = 4(A^t)$.

(d) Show that$(AB)^t = B^t A^t$.

3.4 Elementary row operations

Elementary row operations are useful to find the rank of a matrix (see Section 3.6), to compute the determinants of matrices (see Section 3.7), and to find the inverse of a matrix (see Section 3.8). Furthermore, elementary row operations are widely used in solving systems of linear equations (see Section 3.9).

In this section, we introduce the elementary row operations and apply these operations to transform the given matrix into different form.

Definition 3.16 (Elementary Row Operations).
Let A be an $m \times n$ matrix. The following are known as elementary row operations.

1. ***Interchanging two rows:*** $R_i \leftrightarrow R_j$.*(Rule of Interchanging)*
2. ***Multiplying a row by a nonzero scalar:*** $R_i \to \alpha R_i$ *(α is a nonzero scalar). (Rule of Scaling)*
3. ***Adding a multiple of one row to another:*** $R_i \to R_i + \alpha R_j$ *(α is a nonzero scalar). (Rule of Replacement)*

Example 3.21. Use elementary row operations to transform the given matrix A into, (a) an upper triangular matrix, (b) an identity matrix.

$$A = \begin{bmatrix} 3 & 12 & 6 \\ 1 & 1 & -1 \\ 1 & 2 & 3 \end{bmatrix}$$

Solution: Consider the given matrix A:

(a) First let us transform the matrix A into an upper triangular. This can be done as follows:

$$A = \begin{bmatrix} 3 & 12 & 6 \\ 1 & 1 & -1 \\ 1 & 2 & 3 \end{bmatrix} \; R_1 \to (\tfrac{1}{3})R_1 \; \begin{bmatrix} 1 & 4 & 2 \\ 1 & 1 & -1 \\ 1 & 2 & 3 \end{bmatrix} \text{(Scaling } R_1\text{)}$$

$$R_2 \to R_2 + (-1)R_1, \; R_3 \to R_3 + (-1)R_1 \; \begin{bmatrix} 1 & 4 & 2 \\ 0 & -3 & -3 \\ 0 & -2 & 1 \end{bmatrix} \text{(Replacing } R_2 \text{ and } R_3\text{)}$$

$$R_2 \to (-\tfrac{1}{3})R_2 \; \begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 1 \\ 0 & -2 & 1 \end{bmatrix} \text{(Scaling } R_2\text{)}$$

$$R_3 \to R_3 + 2R_2 \; \begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} \text{(Replacing } R_3\text{)}$$

Hence, the matrix $\begin{bmatrix}1 & 4 & 2\\0 & 1 & 1\\0 & 0 & 3\end{bmatrix}$ is an upper triangular, which is obtained from A by elementary row operations.

(b) To transform the matrix A into a diagonal matrix, we simply change all the entries above the main diagonal into zeros and the entries in the main diagonal into 1. Let us denote the above upper triangular matrix by B. Then we have

$B = \begin{bmatrix}1 & 4 & 2\\0 & 1 & 1\\0 & 0 & 3\end{bmatrix}$ $R_3 \to (\frac{1}{3})R_3$ $\begin{bmatrix}1 & 4 & 2\\0 & 1 & 1\\0 & 0 & 1\end{bmatrix}$ (Scaling R_3)

$R_2 \to R_2 + (-1)R_3,\ R_1 \to R_1 + (-2)R_3$ $\begin{bmatrix}1 & 4 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ (Replacing R_1 and R_2)

$R_1 \to R_1 + (-4)R_2$ $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ (Replacing R_1). Thus, $I_3 = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ is the identity matrix obtained from A.

Definition 3.17. *Two matrices are said to be* ***raw equivalent*** *if one can be obtained from the other by a sequence of elementary row operations.*

Example 3.22. Let A, B, I_3 be the matrices in Example 3.21. Then, A is row equivalent to both B and the identity matrix I_3. Also the matrix B is row equivalent to the identity matrix I_3.

Exercise 3.4.

1. Given the matrix $A = \begin{bmatrix}4 & 3\\2 & 1\end{bmatrix}$, use elementary row operations to find the lower triangular matrix which are row equivalent to A.

2. Given the matrix $B = \begin{bmatrix}0 & 1 & 1\\1 & 0 & 1\\1 & 1 & 0\end{bmatrix}$, use elementary row operations to find an identity matrix which is row equivalent to B.

3.5 Row Echelon Form and Reduced Row Echelon Form of a Matrix

In order to find the rank, or to compute the inverse of a matrix, or to solve a linear system, we usually write the matrix either in its row echelon form or reduced row echelon form.

Definition 3.18. *An $m \times n$ matrix is said to be in* ***echelon form (or row echelon form)*** *if the following conditions are satisfied:*

1. *All nonzero rows are above any rows of all zeros.*
2. *Each leading entry of a row is in a column to the right of the leading entry of the row above it. (A* ***leading entry*** *refers to the left most nonzero entry in a nonzero row)*
3. *All entries in a column below a leading entry are zeros.*

If a matrix in row echelon form satisfies the following additional conditions, then it is in ***reduced echelon form (or reduced row echelon form)***

4. *The leading entry in each nonzero row is 1.*
5. *Each leading 1 is the only nonzero entry in its column.*

A matrix in ***row echelon*** *form is said to be in* ***reduced row echelon*** *when every column that has a leading 1 has zeros in every position above and below the leading entry.*

Example 3.23. The given matrices A, B, C, D are in row echelon form

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 5 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and the following are in reduced row echelon form.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Theorem 3.6 (Uniqueness of the Reduced Echelon Form). *Each matrix is row equivalent to one and only one reduced echelon matrix.*

Definition 3.19. *A* ***pivot position*** *in a matrix A is a location in A that corresponds to a leading 1 in the reduced row echelon form of A. A* ***pivot column*** *is a column of A that contains a pivot position. A* ***pivot element*** *is a nonzero number in a pivot position that is used as needed to create zeros via row operations.*

To write a matrix in reduced echelon form:

1. Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.
2. Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.
3. Use row replacement operations to create zeros in all positions below the pivot.
4. Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.
5. Beginning with the rightmost pivot column and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

Example 3.24. Find the reduced row echelon form of the matrix A.

$$A = \begin{bmatrix} 0 & 0 & 2 & 3 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 1 & 5 \end{bmatrix}.$$

Solution:
Step 1: Here, the left most nonzero column is the second column.
Step 2: By row interchanging rule, we can obtain the pivot position as follows;

$$\begin{bmatrix} 0 & 0 & 2 & 3 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 1 & 5 \end{bmatrix} \; R_1 \leftrightarrow R_3 \; \begin{bmatrix} 0 & 1 & 1 & 5 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{bmatrix}$$

Step 3:
Now, the leading entry is 1, and to create zeros in all positions below the pivot, we use the replacement rule:

$$R_2 \to R_2 + (-2)R_1 \begin{bmatrix} 0 & 1 & 1 & 5 \\ 0 & 0 & -2 & -9 \\ 0 & 0 & 2 & 3 \end{bmatrix}$$

Step 4:
Now we proceed to the second row. Here, the leading entry is -2. Using a scaling rule we obtain a leading 1:

$$R_2 \to (-\frac{1}{2})R_2 \quad \begin{bmatrix} 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & \frac{9}{2} \\ 0 & 0 & 2 & 3 \end{bmatrix}$$

And applying row replacement rule:

$$R_3 \to R_3 + (-2)R_2 \begin{bmatrix} 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & \frac{9}{2} \\ 0 & 0 & 0 & -6 \end{bmatrix}$$

And scaling R_3,

$$R_3 \leftrightarrow (-\frac{1}{6})R_3 \begin{bmatrix} 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & \frac{9}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Step 5: Beginning with the rightmost pivot column, we create zeros above each pivot element. That is, we start from the fourth column:

$$R_1 \to R_1 + (-5)R_3\,, R_2 \to R_2 + (-\frac{9}{2})R_3 \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

And using row replacement (to create zeros above the pivot element in the third column),

$$R_1 \to R_1 + (-1)R_2, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus, the required matrix in reduced row echelon form is given by

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Exercise 3.5.

1. Determine which matrices are in reduced row echelon form.

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 5 \\ 0 & 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

2. Give the row echelon form and also the reduced row echelon form of the following matrices.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & -2 \\ 3 & 0 & 0 \\ 3 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 1 & 3 \\ -3 & 2 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 1 & 2 & 2 \\ 1 & 1 & 0 & 3 \end{bmatrix}$$

3.6 Rank of matrix using elementary row operations

The ranks of matrices are useful in determining the number of solutions for linear systems.

Definition 3.20 (**Rank of Matrix**). *Rank of an $m \times n$ matrix A, denoted by $rank(A)$, is the number of nonzero rows of the reduced row echelon form of A.*

Example 3.25. Determine the ranks of the following matrices which, are in reduced row echelon form.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Solution: Clearly, all the matrices are in reduced row echelon form. Hence, by Definition 3.20, we have $rank(A) = 3$ (since the number of nonzero rows in matrix A is 3). Similarly, $rank(B) = 2$ (since the number of nonzero rows in matrix B is 2), $rank(C) = 2$ (since the number of nonzero rows in matrix C is 2), and $rank(D) = 1$ (since the number of nonzero rows in matrix D is 1).

Example 3.26. Find $rank(A)$, where $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 5 \\ 3 & 3 & 2 \end{bmatrix}$.

Solution: After a sequence of elementary row operations, we obtain the reduced echelon form of A, which is given by

$$\widetilde{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, $rank(A) = 2$.

Remark. The matrix A and its transpose A^t have the same rank. That is

$$rank(A) = rank(A^t).$$

Example 3.27. Verify that the given matrix A and its transpose A^t have the same rank.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0- \end{bmatrix}, \text{ and } A^t = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$

Solution: Observe that the matrix A is in its row echelon form, and hence its rank is 2. Now, we apply elementary row operations to reduce matrix A^t into its row echelon form, and and we get that

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $rank(A^t) = 2 = rank(A)$.

Exercise 3.6. Determine the rank of the following matrices.

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 5 \\ 0 & 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & -2 \\ 3 & 0 & 0 \\ 3 & 2 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 2 & 1 & 3 \\ -3 & 2 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 1 & 2 & 2 \\ 1 & 1 & 0 & 3 \end{bmatrix}$$

3.7 Determinant and its properties

The determinant is a function that takes a square matrix as an input and produces a scalar as an output. It has many beneficial properties for studying, matrices and systems of equations.

Definition 3.21 (Determinant of 2×2 matrix). *The determinant of a 2×2 matrix $A = \begin{bmatrix} a & c \\ d & b \end{bmatrix}$, denoted by $det(A)$ (or $|A|$), is defined by the formula*

$$det(A) = \begin{vmatrix} a & c \\ d & b \end{vmatrix} = ab - cd.$$

Example 3.28. Find the determinant of a matrix $A = \begin{bmatrix} 5 & 2 \\ 3 & 4 \end{bmatrix}$.

Solution: Using Definition 3.21, the determinant of matrix A is given by

$$det(A) = \begin{vmatrix} 5 & 2 \\ 3 & 4 \end{vmatrix} = (5)(4) - (3)(2) = 14.$$

The determinant of a 3×3 matrix can be defined using the determinants of 2×2 matrices.

Definition 3.22 (Determinant of 3×3 Matrix). *Let*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

be a 3×3 matrix, and A_{ij} (for $i, j = 1, 2, 3$) be the 2×2 submatrix of A obtained by deleting the i^{th}-raw and the j^{th}-column of A. Then determinant of A is denoted by $det(A)$ (or $|A|$), and is defined as:

$$\begin{aligned} |A| &= (-1)^{1+1}a_{11}|A_{11}| + (-1)^{1+2}a_{12}|A_{12}| + (-1)^{1+3}a_{13}|A_{13}| \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \end{aligned}$$

Example 3.29. Compute the determinant of a matrix $A = \begin{vmatrix} 2 & 4 & 0 \\ 3 & -1 & 2 \\ 2 & 1 & 1 \end{vmatrix}$.

Solution: Using Definition 3.22, the determinant is given by

$$det(A) = \begin{vmatrix} 2 & 4 & 0 \\ 3 & -1 & 2 \\ 2 & 1 & 1 \end{vmatrix} = 2\begin{vmatrix} -1 & 2 \\ 1 & 1 \end{vmatrix} - 4\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + 0\begin{vmatrix} 3 & -1 \\ 2 & 1 \end{vmatrix}$$

$$= 2(-1-2) - 4(3-4) + 0(3+2) = -6 + 4 + 0 = -2.$$

So far we discussed the determinants of 2×2 and 3×3 matrices. Next we define the determinant of an $n \times n$ matrix for each positive integer n.

Definition 3.23 (Minors and Cofactors).
Let $A = (a_{ij})_{n\times n}$, and A_{ij} be the submatrix of A obtained by deleting the i^{th}-raw and j^{th}-column of A for $i, j = 1, 2, 3, ..., n$. Then

(a) The minor for A at location (i, j), denoted by $M_{ij}(A)$, is the determinant of the submatrix A_{ij}. That is, $M_{ij}(A) = det(A_{ij})$.

(b) The cofactor, denoted by $C_{ij}(A)$, for A at location (i, j) is the sighed determinant of the submatrix A_{ij}. That is, $C_{ij}(A) = (-1)^{i+j} det(A_{ij})$.

Remark. In Definition 3.23, the cofactor $C_{ij}(A)$ at location (i, j) can be computed using the following formula:

$$C_{ij}(A) = \begin{cases} det(A_{ij}), \text{if } i+j \text{ is even} \\ -det(A_{ij}), \text{if } i+j \text{ is odd.} \end{cases}$$

Example 3.30. Compute the matrix of cofactors for the given matrix.

$$(a)\ A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \quad (b)\ B = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 2 & 0 & 1 \end{bmatrix}$$

Solution: (a) The minors of A are

$$M_{11}(A) = 2,\ \ M_{21}(A) = 1,\ \ M_{12}(A) = -1,\ \ M_{22}(A) = 1,$$

and the cofactors are

$$C_{11}(A) = (-1)^{1+1}M_{11}(A) = (1)(2) = 2,\ \ C_{21}(A) = (-1)^{2+1}M_{21}(A) = (-1)(1) = -1,$$

$$C_{12}(A) = (-1)^{1+2}M_{12}(A) = (-1)(-1) = 1,\ \ C_{22}(A) = (-1)^{2+2}M_{12}(A) = (1)(1) = 1.$$

Thus, the matrix of cofactors for A is

$$[C_{ij}(A)] = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}.$$

(b) The minors of B are

$M_{11}(B) = \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} = 1,\ M_{21}(B) = \begin{vmatrix} 0 & 2 \\ 0 & 1 \end{vmatrix} = 0,\ M_{31}(B) = \begin{vmatrix} 0 & 2 \\ 1 & 3 \end{vmatrix} = -2,$

$M_{12}(B) = \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} = -5, M_{22}(B) = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3,\ M_{32}(B) = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = 1,$

$M_{13}(B) = \begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} = -2,\ M_{23}(B) = \begin{vmatrix} 1 & 0 \\ 2 & 0 \end{vmatrix} = 0$ and $M_{33}(B) = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1,$

and the confactors are

$$\begin{aligned}
&C_{11}(B) = (-1)^{1+1}M_{11}(B) = 1,\ \ C_{21}(B) = (-1)^{2+1}M_{21}(B) = 0,\\
&C_{31}(B) = (-1)^{3+1}M_{31}(B) = -2,\ \ C_{12}(B) = (-1)^{1+2}M_{12}(B) = 5,\\
&C_{22}(B) = (-1)^{2+2}M_{22}(B) = -3,\ C_{32}(B) = (-1)^{3+2}M_{32}(B) = -1,\\
&C_{13}(B) = (-1)^{1+3}M_{13}(B) = -2,\ \ C_{23}(B) = (-1)^{2+3}M_{23}(B) = 0,\\
&\text{ and } \ C_{33}(B) = (-1)^{3+3}M_{33}(B) = 1.
\end{aligned}$$

Thus, the matrix of cofactors for B is

$$[C_{ij}(B)] = \begin{bmatrix} 1 & 5 & -2 \\ 0 & -3 & 0 \\ -2 & -1 & 1 \end{bmatrix}.$$

Definition 3.24 (Determinants of $n \times n$ Matrix). *The determinant of a square matrix $A = [a_{ij}]$ of size $n \times n$, denoted by $det(A)$ (or $|A|$), is defined recursively as follows: if $n = 1$ then $det(A) = a_{11}$; otherwise, we suppose that determinants are defined for all square matrices of size less than n and specify that*

$$det(A) = \sum_{k=1}^{n} a_{k1}C_{k1}(A) = a_{11}C_{11}(A) + a_{21}C_{21}(A) + ... + a_{n1}C_{n1}(A), \tag{3.1}$$

*where $C_{ij}(A)$ is the (i, j)-th cofactor of A. The formula (3.1) is called **a cofactor expansion across the 1^{st} column of** A.*

Example 3.31. Consider the matrices given in Example 3.30,

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 2 & 0 & 1 \end{bmatrix}.$$

The cofactors of matrices A and B, respectively, are given by

$$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 5 & -2 \\ 0 & -3 & 0 \\ -2 & -1 & 1 \end{bmatrix}.$$

Now, using Definition 3.24, we have

$$det(A) = a_{11}C_{11} + a_{21}C_{21} = (1)(2) + (-1)(-1) = 3, \text{ and}$$

$$det(B) = b_{11}C_{11} + b_{21}C_{21} + b_{31}C_{31} = (1)(1) + (1)(0) + (2)(-2) = -3.$$

Example 3.32. Compute the determinant of matrix A:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

(a) by expanding the cofactrs across the 1^{st} row

(b) by expanding the cofactrs across the 1^{st} column

Solution: We have the matrix of cofactors $C_{ij}(A)$, given by

$$[C_{ij}(A)] = \begin{bmatrix} -2 & 1 & -2 \\ 0 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

(a) Now, expanding the cofactors across the 1^{st} row, we have

$$det(A) = a_{11}C_{11}(A) + a_{12}C_{12}(A) + a_{13}C_{13}(A) = (1)(-2) + (1)(1) + (0)(-2) = -1.$$

(b) Similarly, expanding cofactors across the 1^{st} column, we have

$$det(A) = a_{11}C_{11}(A) + a_{21}C_{21}(A) + a_{31}C_{31}(A) = (1)(-2) + (0)(0) + (1)(1) = -1.$$

Observe that the determinant has the same value for expansions of cofactors across the 1^{st} row as well as the 1^{st} column. This is true in general, i.e., the determinant value is the same for the expansions of cofactors across any row or any column. This is briefly stated in the following theorem.

Theorem 3.7. *The determinant of an $n \times n$ matrix A can be computed by cofactor expansion across any row or any column. The expansion across i^{th} row is*

$$det(A) = \sum_{j=1}^{n} a_{ij}C_{ij}(A) = a_{i1}C_{i1}(A) + a_{i2}C_{i2}(A) + ... + a_{in}C_{in}(A)$$

$$= (-1)^{i+1}a_{i1}|A_{i1}| + (-1)^{i+2}a_{i2}|A_{i2}| + ... + (-1)^{i+n}a_{in}|A_{in}|$$

and the expansion across j^{th} column is

$$det(A) = \sum_{i=1}^{n} a_{ij}C_{ij}(A) = a_{1j}C_{1j}(A) + a_{2j}C_{2j}(A) + ... + a_{nj}C_{nj}(A)$$

$$= (-1)^{1+j}a_{1j}|A_{1j}| + (-1)^{2+j}a_{2j}|A_{2j}| + ... + (-1)^{n+j}a_{nj}|A_{nj}|$$

Remark. In Theorem 3.7, if the matrix A (for instance) is of size 3×3, then the determinants can be easily computed as follows.

(i) The expansion across 2^{nd} row is

$$|A| = -a_{21}|A_{21}| + a_{22}|A_{22}| + a_{23}|A_{23}|.$$

(ii) The expansion across 3^{rd} column is

$$|A| = a_{13}|A_{13}| - a_{23}|A_{23}| + a_{33}|A_{33}|.$$

(iii) The sign $+$ or $-$ can be determined using the pattern.

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

(iv) The computation of determinants becomes easier by expanding the cofactors across a row or column with the most zero entries.

Example 3.33. Compute the determinant of matrix A by expanding the cofactors across an appropriate row or column.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

Solution: Here, we observe that the 3^{rd} column has more number of zero entries than any other columns and rows. Thus, the determinant of A (by expanding the cofactors across the

3^{rd} column) is given by

$$det(A) = a_{13}|A_{13}| - a_{23}|A_{23}| + a_{33}|A_{33}| = 0 - 1 + 0 = -1.$$

Properties of determinats: Let A be the square matrix of size n.

1. If an entire row (or an entire column) consists of zeros, then $det(A) = 0$.
2. If two rows (or columns) are equal, then $det(A) = 0$.
3. If one row (or column) is a scalar multiple of another row (or column), then $det(A) = 0$.
4. If A, B and C, respectively, are the upper triangular, lower triangular, and diagonal matrices, given by

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}, \; B = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix}, \; D = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix},$$

then

$$det(A) = a_{11}a_{22}a_{33}, \;\; det(B) = b_{11}b_{22}b_{33}, \;\; \text{and} \;\; det(D) = d_{11}d_{22}d_{33}.$$

That is, the determinants of the triangular and diagonal matrices are simply the products of the entries in the main diagonal.

Example 3.34. Determine the determinants of the following matrices.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 2 & 0 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 1 & 0 \\ -2 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 1 & 1 & 3 \\ -1 & -1 & -3 \\ 1 & 2 & 0 \end{bmatrix}$$

Solution: We have, $det(A) = 0$ (since the entire second row of matrix A consists of zeros), $det(B) = 0$ (since the entire third column of matrix A consists of zeros), $det(C) = 0$ (since the first and third rows of C are equal), and $det(D) = 0$ (since the second row of D is a scalar multiple of the first row).

Example 3.35. Compute the determinants of the following matrices.

$$A = \begin{bmatrix} 4 & 3 & -6 \\ 0 & 2 & 9 \\ 0 & 0 & 3 \end{bmatrix}, \; B = \begin{bmatrix} 3 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 1 & 5 \end{bmatrix}, \; D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Solution: Using the properties of determinants, we have

$$det(A) = \begin{vmatrix} 4 & 3 & -6 \\ 0 & 2 & 9 \\ 0 & 0 & 3 \end{vmatrix} = (4)(2)(3) = 24, \;\; det(B) = \begin{vmatrix} 3 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 1 & 5 \end{vmatrix} = (3)(4)(5) = 60, \;\; \text{and}$$

$$det(D) = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{vmatrix} = (2)(3)(5) = 30.$$

Theorem 3.8. *For any square matrix A, $det(A) = det(A^t)$ (Transposition doesn't alter determinants).*

Example 3.36. For the given matrix A, verify that $det(A) = det(A^t)$.

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

Solution: The transpose of matrix A is given by

$$A^t = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 2 & 1 & 3 \end{bmatrix}.$$

Now, we have the determinants of A and A^t are

$$det(A) = 2, \text{ and } det(A^t) = 2.$$

Thus, $det(A) = det(A^t)$.

Theorem 3.9 (Effects of elementary row operations).

*I. If matrix B is obtained from a square matrix A by interchanging any two rows (i.e., $R_i \leftrightarrow R_j$), then $det(B) = -det(A)$. (**Interchanging**)*

*II. If matrix B is obtained from a square matrix A by multiplying the i^{th} row by a nonzero scalar α (i.e., $R_i \to \alpha R_i$), then $det(B) = \alpha det(A)$. (**Scaling**)*

*III If matrix B is obtained from a square matrix A by adding scalar multiple of one row to the other (i.e., $R_i \to R_i + \alpha R_j$), then $det(B) = det(A)$. (**Replacement**)*

Example 3.37. Let $A = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ be the given matrix with $det(A) = -2$.

(a) If a matrix B is obtained from A by interchanging the first and second rows (i.e., $R_1 \leftrightarrow R_2$), then we have

$$det(B) = \begin{vmatrix} 1 & 0 & 1 \\ 3 & 1 & 0 \\ 0 & 1 & -1 \end{vmatrix} = 2.$$

Thus, $det(B) = -det(A)$. Here, we observe that if the row interchanging has been made two times, then $det(B) = (-1)^2 det(A) = det(A)$. In general, if the row interchanging has been made n times, then $det(B) = (-1)^n det(A)$. Thus, $det(B) = det(A)$ if n is even, and $det(B) = -det(A)$ if n is odd.

(b) If a matrix B is obtained from A by multiplying the second row by 4 (i.e., $R_2 \to 4R_2$), then we have

$$B = \begin{vmatrix} 3 & 1 & 0 \\ 4 & 0 & 4 \\ 0 & 1 & -1 \end{vmatrix} = -8.$$

Thus, $det(B) = 4det(A)$. If each row of matrix A is multiplied by 4, then we have

$$det(B) = 4^3 det(A).$$

More generally, if A is an $n \times n$ matrix, and B is obtained by multiplying each row of A by a nonzero scalar c, then we have $det(B) = det(cA) = c^n det(A)$.

(c) If a matrix B is obtained by replacing row 2 (i.e., $R_2 \to R_2 + 2R_1$), then $det(B) = \begin{vmatrix} 3 & 1 & 0 \\ 7 & 2 & 1 \\ 0 & 1 & -1 \end{vmatrix} = 2$. Thus, $det(B) = det(A)$.

Remark. Property (III) of determinants in Theorem 3.9 is particularly more interesting, since it doesn't change the determinant of the original matrix. This property can be used to transform the given matrix into triangular matrix (upper or lower) for which the computation of determinants is much easier than computing the determinant of the original matrix directly, which is tedious and computationally inefficient.

Example 3.38. Compute the determinants of the matrices A and B using elementary row operations.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 3 & 5 & 6 \\ 1 & 3 & 5 & 3 \\ 1 & 1 & 3 & 6 \end{bmatrix}$$

Solution:

(a) Consider the given matrix A. Applying the row replacement; $R_2 \to R_2 - 2R_1$ and then $R_3 \to R_3 - R_2$, we obtain the following upper triangular matrix.

$$\tilde{A} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 7 \end{bmatrix}$$

Therefore, by Theorem 3.9 we have $det(A) = det(\tilde{A}) = (1)(1)(7) = 7$.

(b) Similarly, by applying the row replacement

$$R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - R_1, R_4 \to R_4 - R_1,$$

we obtain the following row equivalent matrix.

$$\tilde{B} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

Now, the determinant of the matrix $\tilde{B}$ (by expanding the cofactors across the 1^{st} column and using the determinant of matrix A computed above) is given by

$$\tilde{B} = \begin{vmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 1 & 4 \end{vmatrix} = 1 \begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \\ 0 & 1 & 4 \end{vmatrix} = (1)(7) = 7.$$

Therefore, by Theorem 3.9 we have $det(B) = det(\tilde{B}) = 7$.

Theorem 3.10 (Product Rule).
If A and B are two matrices for which the product AB is defined, then

$$det(AB) = det(A)det(B).$$

Example 3.39. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}$ be the given matrices. Then verify that

$$det(AB) = det(A)det(B).$$

Solution: Here, we have

$$AB = \begin{bmatrix} 4 & 8 \\ 5 & -4 \end{bmatrix},\ det(AB) = -56,\ det(A) = \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} = -7, \text{ and } det(B) = \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} = 8.$$

Thus,

$$det(A)det(B) = (-7)(8) = -56 = det(AB).$$

Definition 3.25 (Definition of rank using Determinant). *Let A be an $m \times n$ matrix. Then $rank(A) = r$, where r is the largest number such that some $r \times r$ submatrix of A has a nonzero determinant.*

Example 3.40. Compute the the rank of matrix $A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 2 & -3 & 2 & 0 \end{bmatrix}$ using determinants.

Solution: Observe that, the largest possible size of any square submatrix of A is 2×2. We have (say) a submatrix $\begin{bmatrix} 1 & 0 \\ 2 & -3 \end{bmatrix}$ (which is obtained by deleting the last two columns of A) with $\begin{vmatrix} 1 & 0 \\ 2 & -3 \end{vmatrix} = -3 \neq 0$. Therefore, $rank(A) = 2$.

Exercise 3.7.

1. Compute the determinants of the following matrices using elementary row operations.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -2 \\ 5 & -3 & -1 \\ -2 & 0 & 1 \end{bmatrix}$$

2. Compute the determinants of the following matrices by expanding cofactors across any appropriate row or column.

$$A = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 0 \\ 6 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 0 & 1 \\ -1 & 2 & 0 & 1 \\ 5 & 0 & 0 & 0 \\ 4 & 1 & 1 & 2 \end{bmatrix}, C = \begin{bmatrix} 0 & 3 & 1 & 0 & 2 \\ 0 & 2 & 1 & 0 & 2 \\ 5 & 1 & -1 & 3 & 3 \\ 0 & 0 & 1 & 0 & 0 \\ 4 & 1 & 1 & 0 & 1 \end{bmatrix}$$

3. Compute the matrix of cofactors for the given matrices.

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -2 \\ -1 & 1 & 4 \\ 2 & 0 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 2 & 1 & 1 \\ -1 & 0 & 2 & 0 \\ 4 & 1 & -1 & 0 \\ 3 & 0 & 1 & 0 \end{bmatrix}$$

4. Determine the ranks of the following matrices using determinants.

$$A = \begin{bmatrix} 1 & 2 & 3 & 0 & 1 \\ 2 & 1 & 3 & 2 & 4 \\ -1 & 2 & 1 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

3.8 Adjoint and Inverse of a Matrix

The inverses of matrices are useful to solve linear systems. In this section, we define the inverse of a matrix, we discuss different methods to compute an inverse, and the properties of inverses.

Definition 3.26 (Adjoint of a Matrix). *Let A be an $n \times n$ matrix. If $[C_{ij}(A)]$ denotes the matrix of cofactors for A, then the adjugate (or adjoint) matrix of A, denoted by $Adj(A)$, is defined by the formula*

$$Adj(A) = [C_{ij}(A)]^t$$

That is, adjoint of matrix A is the transpose of the matrix of cofactors for A.

Example 3.41. Compute the adjoints of the given matrices.

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}, \text{ and } B = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 3 \\ 2 & 0 & 1 \end{bmatrix}$$

Solution: The matrix of cofactors for A is

$$[C_{ij}(A)] = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}.$$

Thus, the adjoint of matrix A is

$$Adj(A) = [C_{ij}(A)]^t = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}.$$

The matrix of cofactors for B is given by

$$[C_{ij}(B)] = \begin{bmatrix} 1 & 5 & -2 \\ 0 & -3 & 0 \\ -2 & -1 & 1 \end{bmatrix}.$$

Thus, the adjoint of matrix B is

$$Adj(B) = [C_{ij}(B)]^t = \begin{bmatrix} 1 & 0 & -2 \\ 5 & -3 & -1 \\ -2 & 0 & 1 \end{bmatrix}.$$

Definition 3.27 (Inverse of a Matrix). *Let A be an $n \times n$ square matrix. The inverse of matrix A is an $n \times n$ matrix B such that*

$$AB = I_n = BA,$$

where I_n is the $n \times n$ identity matrix. If such a Matrix B exists, then the matrix A is said to be ***invertible*** *(or* ***nonsingular****), and its inverse is denoted by A^{-1} (i.e. $B = A^{-1}$). A matrix that does not have an inverse is said to be* ***noninvertible*** *(or* ***singular****).*

Example 3.42. Consider the following matrices:

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & -3 & -2 \\ -1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we have

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = BA$$

That is, the products AB and BA give us the identity matrix I_2. Therefore, matrix B is the inverse of A i.e., $A^{-1} = B$.

Similarly, we have

$$CD = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = DC.$$

Thus, the matrix D is the inverse of C i.e., $C^{-1} = D$.

Theorem 3.11. *Let A be an $n \times n$ matrix. If A is invertible (non singular) then $det(A) \neq 0$, and the inverse A^{-1} is given by the formula*

$$A^{-1} = \frac{1}{det(A)} Adj(A).$$

Example 3.43. Compute the inverse of the given matrix A.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Solution: We have, $det(A) = 6$,

$$[C_{ij}(A)] = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \text{ and } Adj(A) = [C_{ij}(A)]^t = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Therefore, by Theorem 3.11, we have

$$A^{-1} = \frac{1}{det(A)} Adj(A) = \frac{1}{6} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

Theorem 3.12 (Laws of Inverse). *Let A, B, C be matrices of appropriate sizes so that the following multiplications make sense, I is a suitably sized identity matrix, and α a nonzero scalar. Then*

i. *If the matrix A is invertible, then it has one and only one inverse, A^{-1}.*

ii. *If A is invertible matrix of size $n \times n$, then so is A^{-1} and hence, $(A^{-1})^{-1} = A$.*

iii *If any two of the three matrices A, B, AB are invertible, then so is the third, and moreover, $(AB)^{-1} = B^{-1}A^{-1}$.*

iv *If the matrix A is invertible, then so is αA. Moreover, $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.*

v *If the matrix A is invertible, then so is A^t. Moreover $(A^t)^{-1} = (A^{-1})^t$.*

vi *Suppose A is invertible. If $AB = AC$ or $BA = CA$, then $B = C$.*

Example 3.44. Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ be the given matrix. Then we have

$$A^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}, \text{ and } (A^{-1})^t = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}.$$

Now,

(a) $2A = \begin{bmatrix} 2 & -2 \\ 2 & 0 \end{bmatrix}$ and $(2A)^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} = \frac{1}{2}A^{-1}$. Thus, we have $(2A)^{-1} = \frac{1}{2}A^{-1}$.

(b) $A^t = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ and $(A^t)^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$. Thus, we have $(A^t)^{-1} = (A^{-1})^t$.

Computation of Inverse Using Elementary Row Operations: Gauss-Jordan Elimination

Let A be an $n \times n$ invertible matrix and I_n be the identity matrix of size $n \times n$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Then the inverse A^{-1} can be obtained using elementary row operations as follows.

Finding the Inverse of a Matrix by Gauss-Jordan Elimination

1. Write the $n \times 2n$ matrix that consists of A on the left and the $n \times n$ identity matrix I_n on the right to obtain $[A|I_n]$. This process is called adjoining matrix I_n to matrix A.
2. If possible, row reduce A to I_n using elementary row operations on the entire matrix $[A|I_n]$. The result will be the matrix $[I_n|A^{-1}]$. If this is not possible, then A is noninvertible (or singular).
3. Check your work by multiplying to see that $AA^{-1} = I_n = A^{-1}A$.

Example 3.45. Compute the inverses of the given matrices using Gauss-Jordan Elimination.

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Solution: Let $A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix}$. Then we have

$$[A|I_2] = \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{array}\right] R_2 \to R_2 + (-3)R_1 \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 5 & -3 & 1 \end{array}\right]$$

$$R_2 \to \frac{1}{5}R_2 \quad \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{3}{5} & \frac{1}{5} \end{array}\right] R_1 \to R_1 + R_2 \quad \left[\begin{array}{cc|cc} 1 & 0 & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & -\frac{3}{5} & \frac{1}{5} \end{array}\right].$$

Therefore, the transformed matrix is

$$[I_2|A^{-1}] = \left[\begin{array}{cc|cc} 1 & 0 & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & -\frac{3}{5} & \frac{1}{5} \end{array}\right]$$

and hence, the inverse of matrix A is given by $A^{-1} = \begin{bmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{3}{5} & \frac{1}{5} \end{bmatrix}$.

Similarly, for $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$,

$$[A|I_3] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{array}\right] R_2 \to \frac{1}{2}R_2 \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{array}\right]$$

$$R_3 \to \frac{1}{3}R_3 \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{3} \end{array}\right]$$

Therefore, the transformed matrix is

$$[I_3|A^{-1}] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{3} \end{array}\right]$$

Thus, $A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$.

Exercise 3.8.

1. For the given matrices A and B, compute the adjoint matrices.

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -2 \\ -1 & 1 & 4 \\ 2 & 0 & 3 \end{bmatrix}$$

2. Compute the inverse of the given matrix (if it exists).

$$A = \begin{bmatrix} 1 & 2 \\ 4 & -2 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \end{bmatrix}$$

3. Compute the inverse (if it exists) of the given matrix using elementary row operations.

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 4 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 1 & -2 \\ 0 & 3 & 2 & 1 \end{bmatrix}$$

3.9 System of Linear Equations

Consider an oil refinery that produces gasoline, kerosene and jet fuel form light crude oil and heavy crude oil. The refinery produces 0.3, 0.2 and 0.4 of gasoline, kerosene and jet fuel, respectively, per barrel of light crude oil. And it produces 0.2, 0.4 and 0.3 of gasoline, kerosene and jet fuel, respectively, per barrel of heavy crude oil. This is shown in Table 2. Note that 10% of each of the crude oil is lost during the refining process.

Table 2

	Gasoline	Kerosene	Jet fuel
Light crude oil	0.3	0.2	0.4
Heavy crude oil	0.2	0.4	0.3

Suppose that the refinery has contracted to deliver 550 barrels of gasoline, 500 barrels of kerosene, and 750 barrels of jet fuel. The problem is to find the number of barrels of each crude oil that satisfies the demand.

If l and h represent the number of barrels of light and heavy crude oil, respectively, then the given problem can be expressed as a system of linear equations

$$\begin{aligned} 0.3l + 0.2h &= 550 \\ 0.2l + 0.4h &= 500 \\ 0.4l + 0.3h &= 750 \end{aligned}$$

The given linear system has three equations and two unknowns. The matrix

$$\begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 0.4 \\ 0.4 & 0.4 \end{bmatrix}$$

is known as the **coefficient matrix** of the system, and the right side of the system is a matrix

$$\begin{bmatrix} 550 \\ 500 \\ 750 \end{bmatrix}.$$

With the column vector of unknowns $\begin{bmatrix} l \\ h \end{bmatrix}$, the above information can be organized in matrix form

$$\begin{bmatrix} 0.3 & 0.2 \\ 0.2 & 0.4 \\ 0.4 & 0.3 \end{bmatrix} \begin{bmatrix} l \\ h \end{bmatrix} = \begin{bmatrix} 550 \\ 500 \\ 750 \end{bmatrix}.$$

Example 3.46. Consider the following system of two equations and two unknowns x, y

$$\begin{aligned} ax + by &= b_1 \\ cx + dy &= b_2 \end{aligned}.$$

If we interpret (x, y) as coordinates in the xy-plane, then each of the two equations represents a straight line, and (x^*, y^*) is a solution if and only if the point P with coordinates x^*, y^* lies on both lines. In this case, there are three possible cases: there exists only one solution if the lines intersect (see Figure 1 a), there are infinitely many solutions if the lines coincide (see Figure 1 b) and the system has no solution if the lines are parallel (see Figure 1 c).

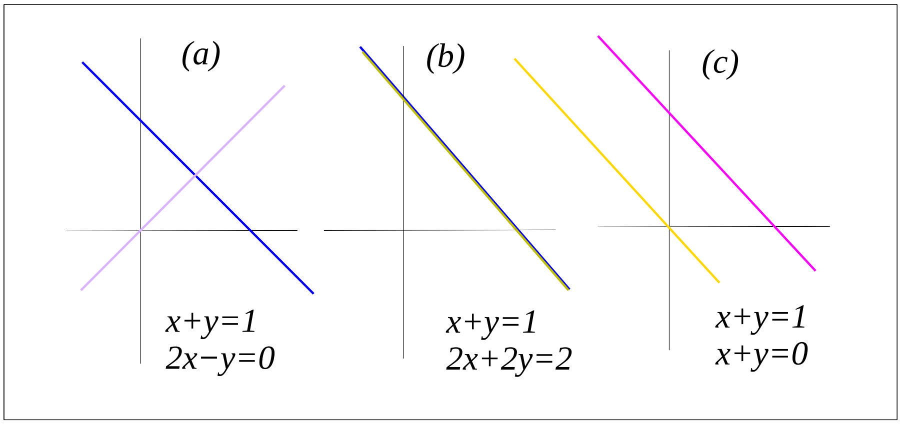


Figure 1: In this figure: (a) represents the case where the lines intersect (b) represents the case where the lines coincide (c) represents the case where the lines are parallel

Let us briefly discuss the three different cases: In part (a) the linear system is given by

$$x + y = 1$$
$$2x - y = 0.$$

This system has only solution, namely (x, y)=($\frac{1}{3}, \frac{2}{3}$).

In part (b) the linear system is given by

$$x + y = 1$$
$$2x + 2y = 2.$$

This system has infinitely many solutions. In fact, the point $(\alpha, 1 - \alpha)$ is a solution for each real number α.

And finally, in part (c) the linear system is given by

$$x + y = 1$$
$$x + y = 0,$$

which has no solutions, since the expressions in the left side of the two equations are the same, but different values in the right side of the two equations.

Definition 3.28. *A linear system (or system of linear equations) of m-**equations** in n-**unknowns** $x_1, x_2, x_3, ..., x_n$ is a set of equations of the form*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m, \end{aligned} \tag{3.2}$$

*where a_{ij}'s (for $i = 1, 2, 3, ..., m$ and $j = 1, 2, 3, ..., n$), are given numbers, called the **coefficients** of the system, and $b_1, b_2, b_3, ..., b_m$ on the right side are also numbers.*

A solution of (3.2) is a set of numbers $x_1, x_2, x_3, ..., x_n$ that satisfies all the m-equations simultaneously.

Matrix Form of a Linear System

From the definition of matrix multiplication, we see that the m-**equations** of (3.2) may be written as a single vector equation

$$Ax = b, \tag{3.3}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

are known as the coefficient matrix, the column vector of unknowns and the column vector of numbers, respectively. We assume that the coefficients a_{ij} are not all zero, so that A is not a zero matrix. Note that x has n components, whereas b has m components.

For the system of linear equations in (3.2), precisely one of the statements below is true:

1. **It admits a unique Solution**: There is one and only one vector $x = (x_1, x_2, x_3, ..., x_n)$ that satisfies all the m-equations simultaneously (the system is **consistent**).

2. **It has infinitely Many Solutions:** There are infinitely many different values of x that satisfy all the m-equations simultaneously (the system is said to be **consistent**).

3. **Has no Solution**: There is no vector x that satisfies all equations simultaneously, or the solution set is empty (the system is said to be **inconsistent**).

3.9.1 Gaussian Elimination

Gaussian elimination, also known as row reduction, is used for solving a system of linear equations. It is usually understood as a sequence of elementary row operations performed on the corresponding matrix of coefficients.

Consider the linear system given in (3.2). The augmented matrix which represents the system is given by

$$[A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ . & . & \dots & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

Then, the idea here is, we solve the linear system whose augmented matrix is in row echelon form, which is row equivalent to the original system. And, we have the following theorem on the row equivalent linear systems.

Theorem 3.13. *Row-equivalent linear systems have the same set of solutions.*

Thus, if the augmented matrix is initially in row echelon form, then we simply solve it by using back substitution. If it is not, then first rewrite it as a row equivalent system whose augmented matrix is in its row echelon form, and then apply Theorem 3.13.

Example 3.47. Rewrite the following linear system as a row equivalent system, and then solve it.

$$\begin{aligned} x_1 - x_2 &= 1 \\ x_1 + 2x_2 &= 4. \end{aligned}$$

Solution: Here, the augmented matrix of the given system is

$$[A|b] = \left[\begin{array}{cc|c} 1 & -1 & 1 \\ 1 & 2 & 4 \end{array}\right],$$

which has row echelon form (after a sequence of elementary operations)

$$\widetilde{[A|b]} = \left[\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 1 & 1 \end{array}\right].$$

Thus, the row equivalent system is

$$\begin{aligned} x_1 - x_2 &= 1 \\ x_2 &= 1. \end{aligned}$$

Clearly, solving the above linear system (whose augmented matrix is in row echelon form) is much easier than solving the original system. The only solution of the linear system (represented by an augmented matrix in row echelon form) is $(x_1, x_2) = (2, 1)$. And, hence by Theorem 3.13, a vector $(x_1, x_2) = (2, 1)$ also solves the original linear system.

Gaussian Elimination:

(a) Write the augmented matrix for the linear system.

(b) Use elementary row operations to rewrite the matrix in row echelon form.

(c) Write the system of linear equations corresponding to the matrix in row echelon form, and use back-substitution to find the solution.

Example 3.48. Consider an oil refinery's problem which is given as a system of linear equations

$$\begin{aligned} 0.3l + 0.2h &= 550 \\ 0.2l + 0.4h &= 500 \\ 0.4l + 0.3h &= 750 \end{aligned}$$

where l and h represent the number of barrels of light and heavy crude oil, respectively. The augmented matrix of the given linear system is

$$[A|b] = \left[\begin{array}{cc|c} 0.3 & 0.2 & 550 \\ 0.2 & 0.4 & 500 \\ 0.4 & 0.3 & 750 \end{array}\right],$$

where

$$A = \begin{bmatrix} 0.3 & 0.2 \\ 0.2 & 0.4 \\ 0.4 & 0.3 \end{bmatrix}, \text{ and } b = \begin{bmatrix} 550 \\ 500 \\ 750 \end{bmatrix}.$$

And the matrix in row echelon form is given by

$$\widetilde{[A|b]} = \left[\begin{array}{cc|c} 0.1 & 0.2 & 250 \\ 0 & 0.1 & 50 \\ 0 & 0 & 0 \end{array}\right].$$

Now, rewriting the given linear system as row equivalent system we have

$$\begin{aligned} 0.1l + 0.2h &= 250 \\ 0.1h &= 50. \end{aligned}$$

The only solution of the above system (in row echelon form) is $(l, h) = (1500, 500)$, which is also a solution for the original system. Thus, an oil refinery needs 1500 barrels of light crude oil and 500 barrels of heavy crude oil in order to satisfy the demand.

Example 3.49. Solve the given linear system by using the method of Gaussian elimination.

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ x_1 - x_2 - 2x_3 &= -1. \end{aligned}$$

Solution: The augmented matrix representing the given system is

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 1 & -1 & -2 & -1 \end{array}\right].$$

Now, by replacing R_2 (i.e., $R_2 \to R_2 - R_1$), we obtain

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & -3 & -3 & -3 \end{array}\right]$$

and by Scaling R_2 ($i.e., R_2 \to (-\frac{1}{3})R_2$), we have

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right].$$

The last matrix is in its row echelon form, and hence the row equivalent system is

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ x_2 + x_3 &= 1. \end{aligned}$$

In this case, the system has infinitely many solutions, and the set of solutions is be given by

$$\{(1 - \alpha, \alpha, 1 - \alpha) : \alpha \in R\}.$$

Example 3.50. Solve the following system of linear equations using the method of Gaussian elimination.

$$\begin{aligned} 4x_2 + 3x_3 &= 8 \\ 2x_1 - x_3 &= 2 \\ 3x_1 + 2x_2 &= 5 \end{aligned}$$

Solution: The augmented matrix of the given system is

$$[A|b] = \left[\begin{array}{ccc|c} 0 & 4 & 3 & 8 \\ 2 & 0 & -1 & 2 \\ 3 & 2 & 0 & 5 \end{array}\right]$$

Applying the following elementary row operations:
$R_1 \leftrightarrow R_3$ (Interchanging R_1 and R_3)

$$\left[\begin{array}{ccc|c} 3 & 2 & 0 & 5 \\ 2 & 0 & -1 & 2 \\ 0 & 4 & 3 & 8 \end{array}\right]$$

$R_2 \leftrightarrow R_3$ (Interchanging R_2 and R_3)

$$\left[\begin{array}{ccc|c} 3 & 2 & 0 & 5 \\ 0 & 4 & 3 & 8 \\ 2 & 0 & -1 & 2 \end{array}\right]$$

$R_3 \to R_3 + (-\frac{2}{3})R_1$ (Replacing R_3)

$$\left[\begin{array}{ccc|c} 3 & 2 & 0 & 5 \\ 0 & 4 & 3 & 8 \\ 0 & -\frac{4}{3} & -1 & -\frac{4}{3} \end{array}\right]$$

$R_3 \to R_3 + \frac{1}{3}R_2$ (Replacing R_3)

$$\left[\begin{array}{ccc|c} 3 & 2 & 0 & 5 \\ 0 & 4 & 3 & 8 \\ 0 & 0 & 0 & \frac{4}{3} \end{array}\right].$$

The last matrix is in row echelon form, and hence the row equivalent system is given by

$$\begin{aligned} 3x_1 + 2x_2 &= 5 \\ 4x_2 + 3x_3 &= 8 \\ 0 &= \tfrac{4}{3} \end{aligned}$$

We observe that the last equation in the linear system above is a contradiction to the fact that $0 \neq \frac{4}{3}$. Consequently, the given linear system has no solution.

Theorem 3.14. *Consider the system of linear equations in* (3.2)*. If A and b are the matrices of coefficients and the column vector of numbers, respectively. Then the following statements are true.*

(i) If $rank(A) = rank([A|b]) =$ number of unknowns, then the linear system has only one solution.

(ii) If $rank(A) = rank([A|b]) <$ number of unknowns, then the linear system has infinitely many solutions.

(iii) If $rank(A) < rank([A|b])$, then the linear system has no solution.

Remark.

(a) From Theorem 3.14, we observe that the linear system (3.2) has no solution if an echelon form of the augmented matrix has a row of the form $[0, 0, ..., 0 \quad b]$ with b nonzero.

(b) A linear system has unique solution when there are no free variable, and it has infinitely many solutions when there is at least one free variable.

Example 3.51. Use matrix rank to determine the number of solutions for the system.

$$(a) \quad \begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_2 + 4x_3 &= 2 \\ 2x_1 + 7x_3 &= 5 \end{aligned} \quad , \quad (b) \quad \begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ 2x_2 + 2x_3 &= 4 \\ x_2 + x_3 &= 2 \end{aligned} \quad (c) \quad \begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_2 + 2x_3 &= -2 \\ -2x_2 - 2x_3 &= 3 \end{aligned}$$

Solution:

(a) We have a linear system

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_2 + 4x_3 &= 2 \\ 2x_1 + 7x_3 &= 5 \end{aligned}$$

and the augmented matrix given by

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 7 & 5 \end{array}\right].$$

After a sequence of elementary row operations, we obtain its row echelon form

$$\widetilde{[A|b]} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

From the transformed matrix, we can see that the matrix A in its row echelon form is

$$\widetilde{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, we have $rank(A) = rank([A|b]) =$ *number of unknowns*. Hence, the given linear system has only one solution.

(b) We have a linear system

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ 2x_2 + 2x_3 &= 4 \\ x_2 + x_3 &= 2 \end{aligned}$$

In this case, the augmented matrix and its row echelon form, respectively, are given by

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 2 & 2 & 4 \\ 0 & 1 & 1 & 2 \end{array}\right] \text{ and } \widetilde{[A|b]} = \left[\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The matrix A in its row echelon form is

$$\widetilde{A} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Here, the matrices $\widetilde{A}$, and $\widetilde{[A|b]}$ have only two nonzero rows. Thus,
$rank(A) = rank([A|b]) <$ *number of unknowns*. Therefore, by Theorem 3.14, the given system has infinitely many solutions.

(c) Here, we have a linear system

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_2 + 2x_3 &= -2 \\ -2x_2 - 2x_3 &= 3. \end{aligned}$$

The augmented matrix $[A|b]$ and its row echelon form $\widetilde{[A|b]}$, respectively, are given by

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 2 & 2 & -2 \\ 0 & -2 & -2 & 3 \end{array}\right] \text{ and } \widetilde{[A|b]} = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Here, the number of nonzero rows of the row echelon form of A and that of $[A|b]$ are 2 and 3, respectively. Therefore, the given linear system has no solution.

Exercise 3.9. Solve the following linear systems using the method of Gaussian elimination.

$$(a) \quad \begin{aligned} -x_1 + x_2 &= 4 \\ -2x_1 + x_2 &= 0 \end{aligned}$$

$$(b) \quad \begin{aligned} x_1 + x_2 &= -1 \\ x_1 - x_2 &= 0 \\ -2x_1 + x_2 &= 3 \end{aligned}$$

$$(c) \quad \begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ 4x_1 + 5x_2 + 6x_3 &= 3 \\ 7x_1 + 8x_2 + 9x_3 &= 6. \end{aligned}$$

$$(d) \quad \begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ 2x_2 + 3x_2 - 2x_3 &= 0 \end{aligned}$$

3.9.2 Cramer's rule

Cramer's Rule is a method for solving linear systems where the number of equations and the number of unknowns are equal. Cramer's rule relies on determinants. Consider the following linear system of n-**equations** in n-**unknowns** $x_1, x_2, x_3, ..., x_n$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ &\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n &= b_n \end{aligned} \tag{3.4}$$

which has a matrix notation

$$Ax = b.$$

Let us define the determinants

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix}, \quad D_j = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1(j-1)} & b_1 & a_{1(j+1)} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2(j-1)} & b_2 & a_{2(j+1)} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n(j-1)} & b_n & a_{n(j+1)} & \dots & a_{nn} \end{vmatrix} \tag{3.5}$$

for $j = 1, 2, 3, ..., n$. Here, D is the determinant of the coefficient matrix A, and for each j D_j represents the determinant of a matrix which is obtained from A after replacing the j-th column by the column vector b.

Theorem 3.15 (Cramer's rule)**.**

(a) If a linear system (3.4) *of n-equations in the same number of unknowns $x_1, x_2, x_3, ..., x_n$, has a nonzero coefficient determinant $D = |A|$, then the system has precisely one solution. This solution is given by*

$$x_1 = \frac{D_1}{D}, x_2 = \frac{D_2}{D}, ..., x_n = \frac{D_n}{D}$$

where D and D_j for $j = 1, 2, 3, ..., n$ are defined in (3.5).

(b) If the system (3.4) *is homogeneous and $D \neq 0$, then it has only the trivial solution $x_1 = 0, x_2 = 0, x_3 = 0, ..., x_n = 0$. If $D = 0$ the homogeneous system also has nontrivial solutions.*

Example 3.52. Use Cramer's rule to solve the system of linear equations.

$$\begin{aligned} 4x_1 - 2x_2 &= 10 \\ 3x_1 - 5x_2 &= 11 \end{aligned}$$

Solution: Here, the coefficient matrix A and the column vector b, respectively, are

$$\begin{bmatrix} 4 & -2 \\ 3 & -5 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 10 \\ 11 \end{bmatrix}.$$

And the determinants D, D_1, D_2 are

$$D = \begin{vmatrix} 4 & -2 \\ 3 & -5 \end{vmatrix} = (-20) - (-6) = -14, \quad D_1 = \begin{vmatrix} 10 & -2 \\ 11 & -5 \end{vmatrix} = (-50) - (-22) = -28,$$

$$D_2 = \begin{vmatrix} 4 & 10 \\ 3 & 11 \end{vmatrix} = (44) - (30) = 14.$$

Therefore, by Theorem 3.15, the unique solution of the given linear system is

$$(x_1, x_2) = \left(\frac{D_1}{D}, \frac{D_2}{D} \right) = (2, -1).$$

Example 3.53. Solve the following system of linear equations using Cramer's rule

$$\begin{aligned} 2x_1 - x_2 &= 0 \\ -x_1 + 2x_2 - x_3 &= 0 \\ -x_2 + x_3 &= 1 \end{aligned}$$

Solution: With the coefficient matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}, \text{ and column vector } b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

the determinants D, D_1, D_2 and D_3 are computed as follows;

$$D = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{vmatrix} = 1, \quad D_1 = \begin{vmatrix} 0 & -1 & 0 \\ 0 & 2 & -1 \\ 1 & -1 & 1 \end{vmatrix} = 1, \quad D_2 = \begin{vmatrix} 2 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 1 \end{vmatrix} = 2$$

and

$$D_3 = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & -1 & 1 \end{vmatrix} = 3.$$

Thus, by Theorem 3.15, the only solution of the given linear system is

$$(x_1, x_2, x_2) = \left(\frac{D_1}{D}, \frac{D_2}{D}, \frac{D_3}{D}\right) = (1, 2, 3).$$

Remark. Cramer's rule doesn't work if the determinant of the coefficient matrix is zero or the coefficient matrix is not square.

Exercise 3.10. Solve the following linear systems using Cramer's rule (if possible).

$$(a) \quad \begin{aligned} 4x_1 - 2x_2 &= 10 \\ 3x_1 - 5x_2 &= 11 \end{aligned}$$

$$(b) \quad \begin{aligned} -x_1 + 2x_2 - 3x_3 &= 1 \\ 2x_1 + x_3 &= 0 \\ 3x_1 - 4x_2 + 4x_3 &= 2. \end{aligned}$$

$$(c) \quad \begin{aligned} x_1 \qquad\quad &= 7 \\ 2x_2 \quad &= 8 \\ 3x_3 &= 24. \end{aligned}$$

3.9.3 Inverse method

The Inverse method is one of the important methods to solve a linear system with n equations in n unknowns.

Example 3.54. Consider a linear system

$$\begin{aligned} x - y &= 1 \\ x + y &= 3. \end{aligned}$$

Using matrix notation, it can be rewritten as

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}. \tag{3.6}$$

And if we denote the coefficient matrix by A, then we have

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \text{ and } A^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Now, multiplying (from the left) both sides of equation (3.6) by A^{-1}, we have

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

And using the fact $A^{-1}A = I_2$, we have

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \text{ This implies } \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Thus, $(x, y) = (2, 1)$ is the only solution of the given system of linear equations. This shows the usefulness of the matrix inverse to solve linear systems.

Consider the following linear system with n**-equations** in n**-unknowns** $x_1, x_2, x_3, ..., x_n$;

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ &\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n &= b_n. \end{aligned} \tag{3.7}$$

The matrix notation of the linear system (3.7) is

$$Ax = b,$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Theorem 3.16 (Inverse Method). *If A is an invertible matrix, then for each $b \in R^n$, the linear system $Ax = b$ has a unique solution, which is given by*

$$x = A^{-1}b.$$

Example 3.55. Solve the following system of linear equations using matrix inverse method.

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ 3x_1 + 2x_2 &= 12 \end{aligned}$$

Solution: The matrix of coefficients A, the inverse A^{-1}, and the column vector b, respectively, are given by

$$A = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} \\ -\frac{3}{7} & \frac{2}{7} \end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 12 \end{bmatrix}.$$

Thus, by Theorem 3.16, the only solution of the given linear system is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1}b = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} \\ -\frac{3}{7} & \frac{2}{7} \end{bmatrix} \begin{bmatrix} 1 \\ 12 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

.

Example 3.56. Use matrix inversion to solve the following linear system.

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 1 \\ x_1 + 2x_2 &= -2 \\ x_3 &= 3 \end{aligned}$$

Solution: The coefficient matrix A, the column vector b and the inverse A^{-1}, respectively, are given by

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} 2 & -3 & -2 \\ -1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, by Theorem 3.16, the unique solution of the given linear system is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A^{-1}b = \begin{bmatrix} 2 & -3 & -2 \\ -1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}.$$

Exercise 3.11. Solve the following linear systems using the method of matrix inversion (if possible).

$$(a) \quad \begin{aligned} 3x_1 + 4x_2 &= -4 \\ 5x_1 + 3x_2 &= 4 \end{aligned}$$

$$(b)\quad \begin{aligned} 4x_1 - x_2 - x_3 &= 1 \\ 2x_1 + 2x_2 + 3x_3 &= 10 \\ 5x_1 - 2x_2 - 2x_3 &= -1. \end{aligned}$$

$$(c)\quad \begin{aligned} 3x_1 \qquad\qquad &= 12 \\ 4x_2 \qquad &= 16 \\ 5x_3 &= 20. \end{aligned}$$

Review exercises

1. For every square matrix A, show that $A + A^t$ is symmetric.

2. Given matrices

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix}$$

   (i) Compute the products $A(BC)$, $(AB)C$, and verify that $A(BC) = (AB)C$.

   (ii) Compute the products $\alpha(AB)$, $(\alpha A)B$, $A(\alpha B)$, and verify that

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

3. A fruit grower raises two crops, apples and peaches. The grower ships each of these crops to three different outlets. In the matrix

$$A = \begin{bmatrix} 125 & 100 & 75 \\ 100 & 175 & 125 \end{bmatrix}$$

   a_{ij} represents the number of units of crop i that the grower ships to outlet j. The matrix $B = \begin{bmatrix} \$3.5 & \$6.00 \end{bmatrix}$ represents the profit per unit. Find the product BA and state what each entry of the matrix represents.

4. A corporation has three factories, each of which manufactures acoustic guitars and electric guitars. In the matrix

$$A = \begin{bmatrix} 70 & 50 & 25 \\ 35 & 100 & 70 \end{bmatrix}$$

   a_{ij} represents the number of guitars of type i produced at factory j in one day. Find the production levels when production increases by 20%.

5. Find the value of x for which the matrix is equal to its own inverse

$$(a)\ \begin{bmatrix} 3 & x \\ -2 & -3 \end{bmatrix} \quad (b)\ \begin{bmatrix} 2 & x \\ -1 & -2 \end{bmatrix} \quad (c)\ \begin{bmatrix} x & 2 \\ -3 & 4 \end{bmatrix}$$

6. If $A = \begin{bmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{bmatrix}$, then

   i. show that $A = A^{-1}$

   ii. show that $A^n = \begin{bmatrix} cos(n\theta) & sin(n\theta) \\ -sin(n\theta) & cos(n\theta) \end{bmatrix}$.

7. If $A = \begin{bmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{bmatrix}$, and $B = \begin{bmatrix} cos(\phi) & sin(\phi) \\ -sin(\phi) & cos(\phi) \end{bmatrix}$, then show that

$$AB = \begin{bmatrix} cos(\theta + \phi) & sin(\theta + \phi) \\ -sin(\theta + \phi) & cos(\theta + \phi) \end{bmatrix}.$$

8. Determine the values of α for which the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & \alpha \end{bmatrix}$ is invertible and find A^{-1}.

9. Show that if A is invertible, then so is A^m for every positive integer m; moreover, $(A^m)^{-1} = (A^{-1})^m$.

10. If A and B are $n \times n$ matrices with A is invertible, then show that

$$(A + B)A^{-1}(A - B) = (A - B)A^{-1}(A + B).$$

11. Solve the following systems of linear equations using Gaussian elimination

$$(a) \quad \begin{array}{l} x_1 - x_2 + 2x_3 = 4 \\ x_1 + x_3 = 6 \\ 2x_1 - 3x_2 + 5x_3 = 4 \\ 3x_1 + 2x_2 - x_3 = 1 \end{array}$$

$$(b) \quad \begin{array}{l} x_1 - 2x_2 + 3x_3 = 9 \\ -x_1 + 3x_2 = -4 \\ 2x_1 - 5x_2 + 5x_3 = 17 \end{array}$$

$$(c) \quad \begin{array}{l} 2x_1 + x_2 - x_3 + 2x_4 = -6 \\ 3x_1 + 4x_2 + x_4 = 1 = 2 \\ x_1 + 5x_2 + 2x_3 + 6x_4 = -3 \\ 5x_1 + 2x_2 - x_3 - x_4 = 1 \end{array}$$

12. Use Cramer's rule (if possible) to solve the following linear systems.

$$(a) \quad \begin{array}{l} x_1 + 2x_2 = 5 \\ -x_1 + x_2 = 1 \end{array}$$

$$
(b) \quad \begin{aligned} 4x_1 - x_2 - x_3 &= 1 \\ 2x_1 + 2x_2 + 3x_3 &= 10 \\ 5x_1 - 2x_2 - 2x_3 &= -1 \end{aligned}
$$

$$
(c) \quad \begin{aligned} 4x_1 - 2x_2 + 3x_3 &= -2 \\ 2x_1 + 2x_2 + 5x_3 &= 16 \\ 8x_1 - 5x_2 - 2x_3 &= 4 \end{aligned}
$$

13. Use matrix inversion method (if possible) to solve the following linear systems.

$$
(a) \quad \begin{aligned} 2x_1 + 3x_2 + x_3 &= -1 \\ 3x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 + 4x_2 + x_3 &= -2 \end{aligned}
$$

$$
(b) \quad \begin{aligned} 2x_1 + 3x_2 + x_3 &= 4 \\ 3x_1 + 3x_2 + x_3 &= 8 \\ 2x_1 + 4x_2 + x_3 &= 5 \end{aligned}
$$

$$
(c) \quad \begin{aligned} 4x_1 - 2x_2 + 3x_3 &= 0 \\ 2x_1 + 2x_2 + 5x_3 &= 0 \\ 8x_1 - 5x_2 - 2x_3 &= 0 \end{aligned}
$$

Chapter Four
Introduction to calculus

Chapter Objectives

At the end of this chapter you should be able to:

- become familiar with the concept of limits.
- explain the intuitive meaning of limit of a function.
- evaluate limits of a function at given points.
- identify and evaluate one-sided limits.
- have an understanding of the basic limit theorems.
- acquire basic knowledge on infinite limits and limits at infinity to find asymptotes.
- get acquainted with the concept of continuity of a function.
- apply the intermediate value theorem to locate roots of equations.
- become familiar with the derivative of a function.
- find the slope and equation of a tangent line to a curve.
- get basic knowledge on the techniques of differentiation.
- evaluate the derivative of polynomial, rational and composite functions.
- find the derivatives of the exponential and logarithmic functions.
- develop an appreciation of higher derivatives of functions.
- apply the concepts of the derivative to find rates of change of variable quantities.
- evaluate maximum and minimum values of functions.
- use the concepts of the derivative to sketch the graph of a function.
- get acquainted with related rate problems.
- define an anti-derivative of a continuous function.
- find indefinite integrals of some elementary functions.
- evaluate the integrals of functions using the techniques of substitution, integration by parts and integration by partial fractions.
- solve integrals involving trigonometric functions.
- find the definite integral of continuous functions.
- apply the concepts of definite integrals to find areas of regions bounded by continuous functions.

4.1. Limits and continuity

At the end of this section you should be able to

- become familiar with the concept of limits.
- explain the intuitive meaning of limit of a function.
- evaluate limits of elementary functions at given points.
- identify right-hand limit from left-hand limit.
- evaluate one-sided limits.
- become aware of the relationship between one-sided limits and the existence of limit of a function.
- find limit of a function in terms of its one-sided limits.
- describe the basic limit theorems.
- find limits of functions given in terms of combinations of function.
- evaluate limit of powers of functions.
- evaluate the limit of composite functions.
- apply the squeeze theorem to evaluate limits.
- gain an understanding of the relationship between infinite limits and vertical asymptotes.
- describe horizontal asymptotes in terms of limits at infinity.
- see the relationship between infinite limits at infinity and oblique asymptotes.
- give the definition of continuous function.
- identify the difference between continuous and discontinuous functions.
- state the theorems on continuity.

In this section we study the concepts of limits and continuity of functions. The concept of limit is fundamental to our main subjects of the branch of mathematics called differential and integral calculus. When we ask about the limit of a function at a point c, we are to ask about tendencies of the values of $f(x)$ as x gets arbitrarily closer and closer to c.

Consider the function $f(x) = 2x$ and find values of f for values of x close to 3 (but not necessarily equal to 3).

- values of x to the left of 3

x	2	2.5	2.9	2.99	2.999	…
$f(x)$	4	5	5.8	5.98	5.998	…

- values of x to the right of 3

x	4	3.5	3.1	3.01	3.001	…
$f(x)$	8	7	6.2	6.02	6.002	…

As you can see from the above two tables, the values of *f(x)* = 2*x* tend to approach to 6 as *x* gets closer and closer to 3 from both sides of 3.
Intuitively, we say "6 is the limit of *f(x)* = 2*x* as x approaches 3" and we write

$$\lim_{x\to 3}(2x) = 6.$$

In general, if for a given real number c, the values of a function *f(x)* approaches a number L as *x* gets close to *c*, we write

$$\lim_{x\to c} f(x) = L$$

We may sometimes write this as $f(x) \to L$ as $x \to c$.

Suppose *f* is a function and *c* is a fixed real number. When one ask for the behavior (approximate value) of *f*(*x*) for *x* near *c*, normally one is not interested about the value *f*(*c*). Instead, one is asking about values of *f* at $x \in (c - \delta, c + \delta)$ for $x \neq c$, with $\delta > 0$ *(δ - delta).* We call the interval $(c - \delta, c + \delta)$ a ***neighborhood of c***. When we exclude *c* from the neighborhood, we obtain a union of two disjoint intervals.

i.e., $(c - \delta, c) \cup (c, c + \delta)$.

Such a set is called a ***deleted neighborhood of c***. For $\delta > 0$, the interval $(c - \delta, c)$ may be called a ***left neighborhood of c*** while $(c, c + \delta)$ a ***right neighborhood of c***. Thus when we talk of *f* ***near*** *c*, we are interested in the function values ***only*** in a deleted neighborhood of *c*.

Therefore, when our interest is to know limit of *f* at *c*, we are mainly curious to know about the tendencies of *f*(*x*) for *x* in a deleted neighborhood of *c*.

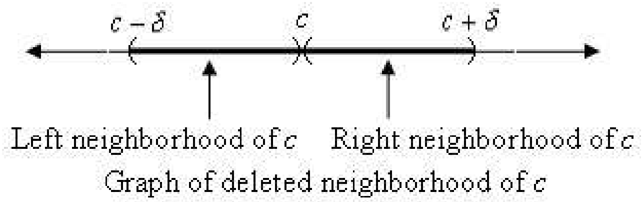


Figure 4.1: Deleted neighborhood of *c*

Similarly, if *x* gets close to 2, the function *f(x)* = *x* + 3 gets close to 5, so that $\lim_{x\to 2}(x+3) = 5$

and if *x* gets close to 1, $f(x) = x^2 - 3$ approaches –2, so that $\lim_{x\to 1}(x^2 - 3) = -2$.

You can also see that

$$\lim_{x\to 2}(x^3 + 1) = 9, \qquad \lim_{x\to 8}\sqrt{x+1} = 3, \quad \text{and} \quad \lim_{x\to 1}\frac{1}{x-5} = -\frac{1}{4}$$

In the above examples we were able to find the limits without much difficulty. However, finding certain limits are not so immediate. For example consider

$$\lim_{x\to 2}\frac{x^2 - 4}{x - 2}$$

Here both $x^2 - 4$ and $x - 2$ approach to 0 as x approaches to 2, and $0/0$ is not determined. But note that

$$\frac{x^2-4}{x-2} = \frac{(x-2)(x+2)}{x-2} = x+2, \quad \text{for } x \neq 2.$$

Thus, for x close to 2 (but not necessarily equal to 2), the behavior of $\frac{x^2-4}{x-2}$ is similar to that of $x + 2$ and it seems reasonable to conclude that

$$\lim_{x\to 2}\frac{x^2-4}{x-2} = \lim_{x\to 2}(x+2) = 4.$$

In the same manner, we have

$$\lim_{x\to 0}\frac{(x+2)^2-4}{x} = \lim_{x\to 0}\frac{x^2+4x+4-4}{x} = \lim_{x\to 0}\frac{x(x+4)}{x} = \lim_{x\to 0}(x+4) = 4, \text{ (for } x \neq 0)$$

$$\lim_{x\to 1}\frac{\sqrt{x}-1}{x-1} = \lim_{x\to 1}\frac{\sqrt{x}-1}{x-1}, \frac{\sqrt{x}+1}{\sqrt{x}+1} = \lim_{x\to 1}\frac{x-1}{(x-1)\left(\sqrt{x}+1\right)} = \lim_{x\to 1}\frac{1}{\sqrt{x}+1} = \frac{1}{\sqrt{x}+1} = \frac{1}{2}, \quad \text{(for } x \neq 1)$$

Even though standard textbooks of calculus give the formal (analytic) definition of limit of a function using the notion of neighborhoods, we shall give here a working definition in terms of what we call ***one-sided limits***.

Definition 4.1:
Suppose f is a function and c is a fixed real number.

1. A real number L is called the **left-hand limit** of f at c, written as $\lim_{x\to c^-} f(x) = \text{L}$ if and only if for all values of x sufficiently close to c from the **left side** of c, the corresponding values of f approach to L.
2. A real number R is called the **right-hand limit** of f at c, written as $\lim_{x\to c^+} f(x) = \text{R}$ if and only if for all values of x sufficiently close to c from the **right side** of c, the corresponding values of f approach to R.

Note that, if the set $(c - \delta, c) \cup (c, c+\delta)$ is a deleted neighborhood of c, then for left-hand limit we take $x \in (c - \delta, c)$, i.e. $x < c$, and for right-hand limit we take $x \in (c, c+\delta)$, i.e. $x > c$ (but not necessarily $x = c$).

Example 4.1: Let $f(x) = \begin{cases} x^2, & for\ x < 1 \\ 2x, & for\ x > 1 \end{cases}$ Then $\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} x^2 = 1$, while $\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} 2x = 2$

Example 4.2: Let $f(x) = \frac{|x|}{x} = \begin{cases} -1, & for\ x < 0 \\ 1, & for\ x > 0 \end{cases}$

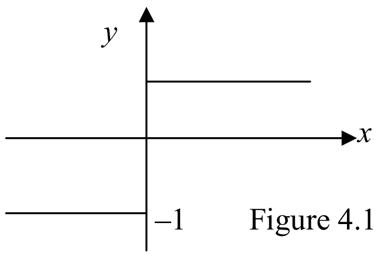


Figure 4.1

$\lim_{x\to 0^-} f(x) = \lim_{x\to 1^-}(-1) = -1$

and $\lim_{x\to 0^+} f(x) = 1$

Example 4.3 : Let f(x) = $\begin{cases} 2x+1, & for\ x < 0 \\ x^2+1. & for\ x > 0 \end{cases}$

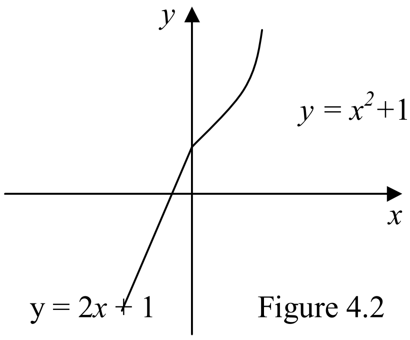


Figure 4.2

The $\lim_{x\to 0^-} f(x) = \lim_{x\to 0^+} (2x + 1) = 1$

and $\lim_{x\to 0^+}$ f(x) = $\lim_{x\to 0^+} (x^2 + 1) = 1$

Note that in this example $\lim_{x\to 0^-} f(x) = \lim_{x\to 0^+} f(x)$

Definition 4.2:
Suppose f is function and c is a fixed real number. A real number L is called the **limit of f at *c*** if and only if the left-and right-hand limits exist and are both equal to L;
i.e. $\lim_{x\to c} f(x)$ = L if and only if $\lim_{x\to c^-} f(x)$ = L = $\lim_{x\to c^+} f(x)$

Thus for $\lim_{x\to c} f(x)$ *to exist, the following conditions must be satisfied:*

i) $\lim_{x\to c^-} f(x)$ *must exist*

ii) $\lim_{x\to c^+} f(x)$ *must exist*

iii) $\lim_{x\to c^-} f(x) = \lim_{x\to c^+} f(x)$

Otherwise, we say $\lim_{x\to c} f(x)$ **does not exist.**

Thus in **Example 4.3**, where *f(x)* = $\begin{cases} 2x+1, & for\ x < 0 \\ x^2+1, & for\ x > 0 \end{cases}$

we have seen above that $\lim_{x\to 0^-} f(x) = 1 = \lim_{x\to 0^+} f(x)$ Thus, $\lim_{x\to 0} f(x) = 1$.

Example 4.4: Let $f(x) = 2^x$ for $x \in \Re$. Then

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} 2^x = 2^1 = 2 \quad \text{and} \quad \lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} 2^x = 2^1 = 2$$

Since $\lim_{x\to 0^-} 2^x = \lim_{x\to 1^+} 2^x = 2$, we have $\lim_{x\to 1} 2^x = 2$

In fact, if $a > 0$, $a \neq 1$, then $\lim_{x\to c} a^x = a^c$, for any $c \in \mathbf{R}$

Similarly, you can show that $\lim_{x\to c} \log_a^x = \log_a^x = \log_a^c$, for c > 0 and $\lim_{x\to c} \sin x = \sin c$, $\forall c$.

Example 4.5: Let f(x) = $\begin{cases} x^2, & for\ x \leq 1 \\ 3, & for\ x > 1 \end{cases}$

Then $\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} x^2 = 1$ while $\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} 3 = 3$. Since $1 \neq 3$, $\lim_{x\to 1} f(x)$ does not exist.

Example 4.6: Let $f(x) = \sqrt{x}$, for $x \geq 0$. Then $\lim_{x\to 0^+} f(x) = 0$. But since $f(x) = \sqrt{x}$ is not defined to the left of 0, $\lim_{x\to 0^-} \sqrt{x}$ does not exist. Hence $\lim_{x\to 0} \sqrt{x}$ does not exist.

Remark: *If a function f has a limit as x approaches a number c, then the limit is* ***unique****; i.e. if* $\lim_{x\to c} f(x) = L_1$ *and* $\lim_{x\to c} f(x) = L_2$*, then* $L_1 = L_2$.

- **Basic Limit Theorems**

> **Theorem 4.1:** Suppose $\lim_{x\to c}$ f(x) = L, $\lim_{x\to c}$ g(x) = M and k is a constant.
>
> Then i) $\lim_{x\to c}$ kf(x) = k $\lim_{x\to c}$ f(x) = KL … **Constant Rule**
>
> ii) $\lim_{x\to c}$ (f + g)(x) = $\lim_{x\to c}$ f(x) + $\lim_{x\to c}$ g(x) = L + M … **Addition Rule**
>
> iii) $\lim_{x\to c}$ (f – g)(x) = $\lim_{x\to c}$ f(x) - $\lim_{x\to c}$ g(x) = L – M … **Difference Rule**

Example 4.7: $\lim_{x\to \pi/2} 5\sin x = 5 \lim_{x\to \pi/2} \sin x = 5(1) = 5$

Example 4.8: Let $f(x) = 2x$ and $g(x) = 5x - 1$. Then

$\lim_{x\to 1} (f+g)(x) = \lim_{x\to 1} f(x) + \lim_{x\to 1} g(x) = \lim_{x\to 1} (2x) + \lim_{x\to 1} (5x-1) = 2(1) + 5(1) - 1 = 2 + 5 - 1 = 6$

$\lim_{x\to 3} (f-g)(x) = \lim_{x\to 3} f(x) - \lim_{x\to 3} g(x) = \lim_{x\to 3} (2x) - \lim_{x\to 3} (5x-1) = 2(3) - [5(3) - 1] = 6 - 14 = -8$

> **Theorem 4.2:** Assume that $\lim_{x\to c}$ f(x) = L and $\lim_{x\to c}$ g(x) = M.
>
> Then $\lim_{x\to c}$ (fg)(x) = $\left(\lim_{x\to c} f(x)\right)\left(\lim_{x\to c} g(x)\right)$ = L.M … **Product Rule**

Example 4.9: $\lim_{x\to \pi} x \cos x = \lim_{x\to \pi} x . \lim_{x\to \pi} \cos x = \pi . \cos \pi = \pi(-1) = -\pi$.

It follows from Theorem 4.2 that $\lim_{x\to c} x^2 = \lim_{x\to c} x.x = \lim_{x\to c} x . \lim_{x\to c} x = c.c = c^2$

In general, if n is a positive integer, $\lim_{x \to c} x^n = c^n$.

Thus, if $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$ is any polynomial function of degree n and c is any real number, then from Theorems 4.1 and 4.2, we get

$$\lim_{x \to c} P(x) = \lim_{x \to c} (a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0)$$
$$= \lim_{x \to c} a_n x^n + \lim_{x \to c} a_{n-1} x^{n-1} + \ldots + \lim_{x \to c} a_2 x^2 + \lim_{x \to c} a_1 x + \lim_{x \to c} a_0$$
$$= a_n \lim_{x \to c} x^n + a_{n-1} \lim_{x \to c} x^{n-1} + \ldots + a_2 \lim_{x \to c} x^2 + a_1 \lim_{x \to c} x + a_0$$
$$= a_n c^n + a_{n-1} c^{n-1} + \ldots + a_2 c^2 + a_1 c + a_0 = P(c)$$

Example 4.10: Let $P(x) = 2x^3 + 4x^2 - 3x + 1$. Then

$$\lim_{x \to -1} P(x) = \lim_{x \to -1} (2x^3 + 4x^2 - 3x + 1) = 2(-1)^3 + 4(-1)^2 - 3(-1) + 1 = -2 + 4 + 3 + 1 = 6$$

> **Theorem 4.3:** Assume that $\lim_{x \to c} f(x) = L$, $\lim_{x \to c} g(x) = M$ and suppose $M \neq 0$
>
> Then $\lim_{x \to c} \left(\frac{f}{g}\right)(x) = \frac{\lim_{x \to c} f(x)}{\lim_{x \to c} g(x)} = \frac{L}{M}$ … **Quotient Rule**

Example 4.11: $\lim_{x \to 10} \frac{\log x}{x} = \frac{\lim_{x \to 10} \log x}{\lim_{x \to 10} x} = \frac{1}{10}$.

If $f(x) = \frac{p(x)}{q(x)}$ *is a rational function, then* $\lim_{x \to c} f(x) = \lim_{x \to c} \frac{p(x)}{q(x)} = \frac{\lim_{x \to c} p(x)}{\lim_{x \to c} q(x)} = \frac{p(c)}{q(c)} = f(c)$ if $f(c) \neq 0$.

Example 4.12: $\lim_{x \to -2} \frac{x^3 - 4x + 1}{4x^2 + x - 6} = \frac{\lim_{x \to -2}(x^3 - 4x + 1)}{\lim_{x \to -2}(4x^2 + x - 6)} = \frac{(-2)^3 - 4(-2) + 1}{4(-2)^2 + (-2) - 6} = \frac{-8 + 8 + 1}{16 - 2 - 6} = \frac{1}{8}$.

> **Theorem 4.4:** Suppose $\lim_{x \to c} f(x) = L$, $L \neq 0$ and $a \in \mathbf{R}$ such that $L^a \in \mathbf{R}$
>
> Then, $\lim_{x \to c} (f(x))^a = L^a$ ………….... **Power Rule**

Example 4.13: $\lim_{x \to \frac{\pi}{2}} \sqrt{\sin x} = \lim_{x \to \frac{\pi}{2}} (\sin x)^{1/2} = \left[\lim_{x \to \frac{\pi}{2}} (\sin x)\right]^{1/2} = 1^{\frac{1}{2}} = \sqrt{1} = 1$.

Example 4.14: $\lim_{x \to 4} \sqrt[3]{(x^2 + 2x + 3)^2} = \lim_{x \to 4} (x^2 + 2x + 3)^{2/3} = (16 + 4 + 3)^{2/3} = (27)^{2/3} = 3^2 = 9$

Theorem 4.5 (The Squeezing Theorem). Suppose f, g and h are functions such that f(x) ≤ h(x) ≤ g(x) for all x in some deleted neighborhood of c. If $\lim_{x\to c} f(x) = L = \lim_{x\to c} g(x)$, then $\lim_{x\to c} h(x) = L$.

Example 4.14: Evaluate $\lim_{x\to 0} x^2 \sin\frac{1}{x}$

Solution: It may be tempting to consider $x^2 \sin\frac{1}{x}$ as the product of x^2 and $\sin\frac{1}{x}$ and then use the Product Rule. Unfortunately it can be shown that $\lim_{x\to 0} \sin\frac{1}{x}$ does not exist. Thus we cannot use the Product Rule to evaluate the given limit. However since the sine function has range [–1, 1], it follows that $-1 \le \sin\frac{1}{x} \le 1$, for $x \ne 0$. Multiplying both sides by x^2, we get

$$-x^2 \le x^2\sin\frac{1}{x} \le x^2 \text{ with } \lim_{x\to 0}(-x^2) = 0 = \lim_{x\to 0} x^2$$

Thus, by the Squeeze Theorem, we get $\lim_{x\to 0} x^2 \sin\frac{1}{x} = 0$.

Remark: One of the most important applications of the Squeezing Theorem is evaluating $\lim_{x\to 0}\frac{\sin x}{x}$. We cannot apply the Quotient Rule to evaluate this limit since the limit of the denominator is 0. But using some geometric constructions and *the Squeeze Theorem* it can be shown that

$$\lim_{x\to 0}\frac{\sin x}{x} = 1$$

Remark: *The above result has important consequences especially in the evaluation of some limits involving trigonometric functions.*

Example 4.15: Find $\lim_{x\to 0}\frac{\sin 5x}{x}$

Solution: $\lim_{x\to 0}\frac{\sin 5x}{x} = \lim_{x\to 0}\frac{5\sin 5x}{5x} = 5\lim_{x\to 0}\frac{\sin 5x}{5x}$

If we put y = 5x, we have as x → 0, 5x → 0 so that y → 0. Thus $\lim_{x\to 0}\frac{\sin 5x}{x} = 5\lim_{y\to 0}\frac{\sin y}{y} = 5$.

In general, for any $a \in \Re$, $\lim_{x\to 0}\frac{\sin ax}{x} = a$.

Example 4.16: Find $\lim_{x\to 0}\frac{\tan x}{x}$

Solution: $\lim_{x\to 0}\frac{\tan x}{x} = \lim_{x\to 0}\left(\frac{\sin x}{\cos x}.\frac{1}{x}\right) = \lim_{x\to 0}\left(\frac{\sin x}{x}\right)\cdot\lim_{x\to 0}\left(\frac{1}{\cos x}\right) = 1.\frac{1}{1} = 1.$

Example 4.17: Evaluate $\lim_{x\to 0}\frac{\cos x - 1}{x}$

Solution: By multiplying both numerator and denominator by cos x + 1 we get

$$\lim_{x\to 0}\frac{\cos x-1}{x} = \lim_{x\to 0}\left(\frac{\cos x-1}{x}\right)\left(\frac{\cos x+1}{\cos x+1}\right) = \lim_{x\to 0}\frac{\cos^2 x-1}{x(\cos x+1)}$$

$$= \lim_{x\to 0}\frac{-\sin^2 x}{x(\cos x+1)} \qquad \text{(since } \sin^2 x + \cos^2 x = 1\text{)}$$

$$= \lim_{x\to 0}\left(\frac{\sin x}{x}\right)\left(\frac{-\sin x}{\cos x+1}\right) = \lim_{x\to 0}\frac{\sin x}{x}.\lim_{x\to 0}\frac{-\sin x}{\cos x+1} = 1.\frac{0}{1} = 1.0 = 0.$$

- **Infinite Limits, Limits at Infinity and Asymptotes**

When $\lim_{x\to c^+}$ f(x) does not exist, it may happen that as x approaches *c* from right, the value of f(x) becomes indefinitely large or becomes negative and indefinitely large in absolute value. The value of f(x) may behave similarly when the left-hand limit at *c* does not exist. We shall use the symbols ∞ (infinity) and - ∞ to express these cases, respectively.

To explain these concepts consider the function f(x) = $\frac{1}{x}$, for x ≠ 0

As x gets close to 0 from right, the values of f(x) = $\frac{1}{x}$ become arbitrarily large positive.

In this case we write $\lim_{x\to 0^+}\frac{1}{x} = \infty$

and when x gets close to 0 from left, the values of f(x) = $\frac{1}{x}$ become arbitrarily small negative.

In this case we write $\lim_{x\to 0^-}\frac{1}{x} = -\infty$. See Figure 4.3.

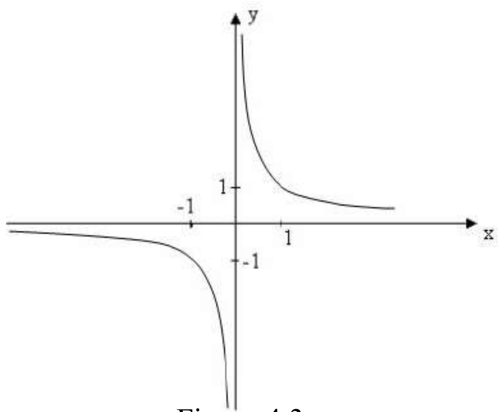


Figure 4.3

Definition 4.3:

Let f be a function defined in a deleted neighborhood of c.

i) We say that the **left-hand limit** of f(x) at c is **infinity**, and write $\lim_{x\to c^-}$ f(x) = ∞ if for every real number M, we have f(x) > M for every x close to c from the left side of

ii) We say that the **right-hand limit** of f(x) at c is **infinity**, and write $\lim_{x\to c^+}$ f(x) = ∞ if for every real number M, we have f(x) > M for every x close to c from the right side if c.

iii) We say that the **limit** of f(x) at c is **infinity** and write $\lim_{x \to c} f(x) = \infty$

if and only if $\lim_{x \to c^-} f(x) = \infty$ and $\lim_{x \to c^+} f(x) = \infty$

Definition 4.4:

Let f be a function defined in a deleted neighborhood of c.

i) We say that the **left-hand limit** of f(x) at c is **negative infinity**, and write $\lim_{x \to c^-} f(x) = -\infty$

if for every real number M, we have f(x) < M for every x close to c from the left side of c.

ii) We say that the **right-hand limit** of f(x) at c is **negative infinity**, and write $\lim_{x \to c^+} f(x) = -\infty$

if for every real number M, we have f(x) < M for every x close to c from the right side if c.

iii) We say that the **limit** of f(x) at c is **negative infinity** and write $\lim_{x \to c} f(x) = -\infty$

if and only if $\lim_{x \to c^-} f(x) = -\infty$ and $\lim_{x \to c^+} f(x) = -\infty$

Example 4.18: For f(x) = $\frac{1}{x}$, for x ≠ 0, $\lim_{x \to 0^+} \frac{1}{x} = \infty$ and $\lim_{x \to 0} \frac{1}{x} = -\infty$. Hence $\lim_{x \to 0} \frac{1}{x}$ does not exist. Whereas for f(x) = $\frac{1}{x^2}$, x ≠ 0, $\lim_{x \to 0^+} \frac{1}{x^2} = \infty = \lim_{x \to 0^-} \frac{1}{x^2}$. Hence $\lim_{x \to 0} \frac{1}{x^2} = \infty$.

In general, for any real number c and f(x) = $\frac{1}{x-c}$ we have $\lim_{x \to c^+} f(x) = \lim_{x \to c^+} \frac{1}{x-c} = \infty$ and

$\lim_{x \to c^-} f(x) = \lim_{x \to c^-} \frac{1}{x-c} = -\infty$.

Definition 4.5:

Suppose f is a function and c is a fixed real number. We say that the line x = c is a **vertical asymptote** of the graph of f if and only if either

$$\lim_{x \to c^-} f(x) = \pm\infty \quad \text{or} \quad \lim_{x \to c^+} f(x) = \pm\infty$$

Remark: *From the above examples, we can see that the line x = 0 (i.e. the y-axis) is a vertical asymptote of the graphs of the functions f(x) =* $\frac{1}{x}$ *and f(x) =* $\frac{1}{x^2}$*, while the line x = c is a vertical asymptote of the graph of f(x) =* $\frac{1}{x-c}$.

Example 4.19: Find all the vertical asymptotes of f(x) = $\frac{x+2}{x^2-1}$

Solution: If c is any number different from 1 or -1, then by the Quotient Rule,

$$\lim_{x \to c} f(x) = \lim_{x \to c} \frac{x+2}{x^2-1} = \frac{c+2}{c^2-1} \in \mathbf{R}$$

Thus any line x = c for c ≠ ± 1 cannot be a vertical asymptote.

For c = 1, $\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} \frac{x+2}{x^2-1} = \lim_{x\to 1^+} \left(\frac{x+2}{x+1}\right)\left(\frac{1}{x-1}\right)$

$$= \lim_{x\to 1^+} \left(\frac{x+2}{x+1}\right) . \lim_{x\to 1^+}\left(\frac{1}{x-1}\right) = \frac{3}{2}(\infty) = \infty$$

Similarly, for c = -1, $\lim_{x\to 1^+} f(x) = \lim_{x\to -1^+} \left(\frac{x+2}{x-1}\right) . \left(\frac{1}{x+1}\right)$

$$= \lim_{x\to -1^+} \left(\frac{x+2}{x-1}\right) . \lim_{x\to 1^+}\left(\frac{1}{x+1}\right) = \left(-\frac{1}{2}\right)(\infty) = -\infty$$

Hence the lines x = 1 and x = -1 are vertical asymptotes of the graph of the function f(x) = $\frac{x+2}{x^2-1}$.

Next, we try to investigate the behavior of a function f as x increases (or decreases) indefinitely, and try to see if we have $\lim_{x\to\infty}$ f(x) or $\lim_{x\to -\infty}$ f(x). Such limits, if they exist, are in general called limits at infinite.

Definition 4.6:

i) Suppose f is a function defined on an interval of the form (c, ∞), for some c∈**R**. We say that the limit of f(x) **as x approaches to infinity** is the number L, and write $\lim_{x\to\infty}$ f(x) = L if when x is assigned sufficiently large positive values, the corresponding values of f approach to L.

ii) Suppose f is a function defined on an interval of the form (-∞, c) for some c∈**R**. We say that the limit of f(x) as **x approaches to negative infinity** is the number L, and write $\lim_{x\to -\infty}$ f(x) = L if when x is assigned sufficiently small negative values, the corresponding values of f approach to L.

Example 4.20: Let f(x) = $\frac{1}{x}$, for x ≠ 0.

When x is assigned sufficiently large positive values, the values of f(x) = $\frac{1}{x}$ become close to 0. Similarly for values of x sufficiently small negative values, f(x) = $\frac{1}{x}$ becomes close to 0. Hence $\lim_{x\to\infty} \frac{1}{x} = 0$ and $\lim_{x\to -\infty} \frac{1}{x} = 0$. See Figure 4.3 above.

Similarly, $\lim_{x\to\infty} \frac{1}{x^2} = 0$, $\lim_{x\to -\infty} \frac{1}{x^2} = 0$ and in general, $\lim_{x\to\infty} \frac{1}{(x-c)^2} = 0$.

Definition 4.7:

If for a function f and a real number L, $\lim_{x\to\infty}$ f(x) = L or $\lim_{x\to -\infty}$ f(x) = L, then the line y = L is called a **horizontal asymptote** to the graph of f.

Thus the line y = 0 (i.e. the x-axis) is a horizontal asymptote for both the function f(x) = $\frac{1}{x}$ and f(x) = $\frac{1}{x^2}$. See Figure 4.3 above.

Example 4.21: Find a horizontal asymptote to the graph of f(x) = $\frac{3x^2 - x + 1}{2x^2 + 5}$

Solution: Since we are interested with the behavior of f for large values of |x|, we divide both numerator and denominator of f by the leading exponent (i.e.x^2) to get

$$f(x) = \frac{3x^2 - x + 1}{2x^2 + 5} = \frac{\frac{3x^2}{x^2} - \frac{x}{x^2} + \frac{1}{x^2}}{\frac{2x^2}{x^2} + \frac{5}{x^2}} = \frac{3 - \frac{1}{x} + \frac{1}{x^2}}{2 + \frac{5}{x^2}}$$

Then $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{3x^2 - x + 1}{2x^2 + 5} = \frac{\lim_{x\to\infty}\left(3 - \frac{1}{x} + \frac{1}{x^2}\right)}{\lim_{x\to\infty}\left(2 + \frac{5}{x^2}\right)} = \frac{3 - 0 - 0}{2 + 0} = \frac{3}{2}$

Thus $\lim_{x\to\infty} f(x) = \frac{3}{2}$ and the line y = $\frac{3}{2}$ is a horizontal asymptote to the graph of f.

Similarly, $\lim_{x\to-\infty} f(x) = \frac{3}{2}$.

Remark: *For a rational function f(x) =* $\frac{p(x)}{q(x)}$*, with deg(p)<deg(q), we find a horizontal asymptote by applying the above technique.*

As a combination of the above two subsections, it may happen that as the values of |x| increase without bound, the corresponding values of |f(x)| also increases without bound leading to what are generally called infinite limits at infinity.

Definition 4.8:
Let f be defined on an interval of the form (c, ∞), for c∈**R**. We say that the limit of f(x) as **x approaches to infinity is infinity**, written $\lim_{x\to\infty} f(x) = \infty$ whenever x is assigned sufficiently large positive values, the corresponding values of f(x) increase without bound.

Remark: *Analogous definitions can be given for*
$\lim_{x\to\infty} f(x) = -\infty$, $\lim_{x\to-\infty} f(x) = \infty$ *and* $\lim_{x\to-\infty} f(x) = -\infty$.

Example 4,22: For f(x) = x^3, we have $\lim_{x\to\infty} x^3 = \infty$ and $\lim_{x\to-\infty} x^3 = -\infty$

Example 4.23: $\lim_{x\to\infty} \frac{x^4 - 3x + 2}{2x^2 + 1} = \lim_{x\to\infty} \frac{1 - \frac{3}{x^3} + \frac{2}{x^4}}{\frac{2}{x^2} + \frac{1}{x^4}} = \infty$. (By dividing by x^4).

Definition 4.9:
If for a function f and for two real numbers a and b $\lim_{x\to\pm\infty}$ [f(x) – (ax + b)] = 0, then the line y = ax + b is called an **oblique** (or a **skew**) asymptote to the graph of f.

In general, for a rational function f(x) = $\frac{p(x)}{q(x)}$, *we have*

i) *When degree(p) < degree (q),* $\lim_{x\to\pm\infty} f(x) = 0$ *and the x-axis is a horizontal asymptote of f.*

ii) *When degree(p) = degree(q), then f has a horizontal asymptote given by the quotient of the leading coefficients of p and q.*

iii) *When degree(p) > degree(q), then* $\lim_{x\to\pm\infty} f(x) = \pm\infty$, *and in particularly if degree(p) = degree(q)+1, then f has an oblique asymptote obtained as a quotient when we divide p by q.*

Example 4.24: Let f(x) = $\frac{4x^2 + 5x - 1}{x + 3}$, find all asymptotes of f.

Solution: Since $\lim_{x\to -3^+} f(x) = \lim_{x\to -3^+} \frac{4x^2 + 5x - 1}{x + 3} = \infty$, the line x = -3 is a vertical asymptote.

By the long division method, we get f(x) = $\frac{4x^2 + 5x - 1}{x + 3} = (4x - 7) + \frac{20}{x + 3}$

$\Rightarrow \lim_{x\to\infty} [f(x) - (4x - 7)] = \lim_{x\to\infty} \frac{20}{x + 3} = 0$

Therefore, the line y = 4x – 7 is an oblique asymptote of f.

- **A special Limit in Exponential Function**

Consider the function f(x) = $\left(1 + \frac{1}{x}\right)^x$ with domain (-∞, -1) ∪ (0, ∞)

The following two tables indicate the behavior of the values of f(x) as x approaches to positive and negative infinity, respectively,

x	2	10	100	1000	10,000	100,000
$f(x)$	2.75	2.593743	2.704814	2.716924	2.718146	2.718268

x	-2	-10	-100	-1000	-10,000	-100,000
$f(x)$	4	2.867972	2.731999	2.719642	2.718418	2.718295

As is tried to be indicated from the above tables, the values of $\left(1+\frac{1}{x}\right)^x$ tend to approach to an irrational number whose value is 2.7182818…. This number, denoted by e, is called the **base of the natural logarithm**, and plays an important role in calculus.

Remark: *The natural logarithmic function (with base e) is given by f(x) =* $\log_e x$ *and is denoted by f(x) = ℓnx. Its inverse, the* ***natural exponential function*** *is given by f(x) = exp(x) = e^x.*

Thus from the above constructions, we have

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e = \lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x.$$

This limit has important consequences.

Example 4.25: $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^{x+3} = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x.\left(1+\frac{1}{x}\right)^3$ - Rule of exponents

$= \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x . \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^3$ - Product Rule

$= e.1^3 = e$

In general, for any real number *a*, $\lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^{x+a} = e$.

Example 4.26: Show that $\lim_{t\to 0}(1+t)^{1/t} = e$

Solution: We prove this by showing that $\lim_{t\to 0^+}(1+t)^{1/t} = e = \lim_{t\to 0^-}(1+t)^{1/t}$. First use the substitution $t = \frac{1}{x}$, so that $x = \frac{1}{t}$ and as $x \to \infty$, $t \to 0^+$. Hence, $\lim_{t\to 0^+}(1+t)^{1/t} = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$.

Similarly, as $x \to -\infty$, $t \to 0^-$. Hence, $\lim_{t\to 0^-}(1+t)^{1/t} = \lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = e$

Therefore, $\lim_{t\to 0^+}(1+t)^{1/t} = e = \lim_{t\to 0^-}(1+t)^{1/t} \Rightarrow \lim_{t\to 0}(1+t)^{1/t} = e$.

Example 4.27: Evaluate $\lim_{x\to\infty}\left(1-\frac{5}{x}\right)^x$

Solution: Let $t = \frac{-5}{x}$. Then $x = \frac{-5}{t}$ and

$$\lim_{x\to\infty}\left(1-\frac{5}{x}\right)^x = \lim_{t\to 0^-}(1+t)^{-5/t} = \lim_{t\to 0^-}\left[(1+t)^{1/t}\right]^{-5} = \left[\lim_{t\to 0}(1+t)^{1/t}\right]^{-5} = e^{-5} = \frac{1}{e^5}.$$

In general, for any real number *a*, $\lim_{x\to\pm\infty}\left(1+\frac{a}{x}\right)^x = e^a$.

- **Continuity of a Function**

In our everyday usage, the word continuity refers to something that happens without any interruption. In calculus, the term continuity is used to describe functions whose graphs can be traced without any break. We shall give its formal definition using the concept of limits.

Definition 4.10:

- Let f be a function and c be a number in the domain of f. f is said to be **continuous at c** if

$$\lim_{t\to c} f(x) = f(c)$$

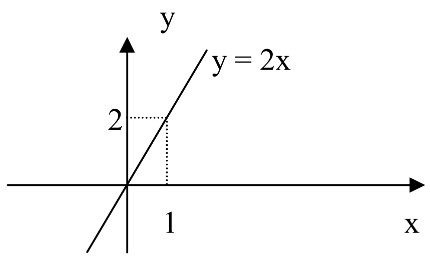
- If f fails to be continuous at c, then we say that f is **discontinuous** (or **not continuous) at c.**
- f is said to **continuous** if it is continuous at each point of its domain.

Example 4.28: Let f(x) = 2x and c = 1

Then $\lim_{x\to 1} f(x) = \lim_{x\to 1} 2x = 2$

and f(1) = 2(1) = 2.

Since $\lim_{x\to 1} 2x = 2 = f(1)$, f is continuous at 1.



In fact f is a continuous function. See Figure 4.4.

Figure 4.4

Remark: *For a function f to be continuous at c, the following conditions must be satisfied*

a. *f(c) must be defined*

b. $\lim_{x\to c}$ *f(x) must exist*

c. $\lim_{x\to c} f(x) = f(c)$

Otherwise if one of the above conditions is not satisfied, then f is discontinuous at c.

Example 4.29: Let $f(x) = \begin{cases} 3x, & for\ x < 0 \\ 2, & for\ x = 0 \\ x^2, & for\ x > 0 \end{cases}$

Then f(0) = 2 so that f(0) is defined . $\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} 3x = 0$

and $\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} x^2 = 0$. Thus $\lim_{x\to 0} f(x) = 0$

But since $\lim_{x\to 0} f(x) \neq f(0)$, f is not continuous at 0. .

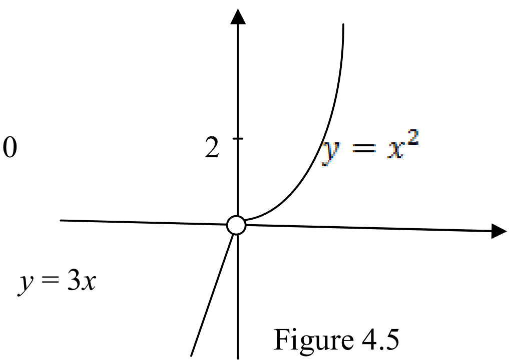


Figure 4.5

Example 4.30: Let *f(x) = sinx*. Then, $\lim_{x\to \pi/2} f(x) = \lim_{x\to \pi/2} \sin x = 1 = \sin\pi/2 = f(\pi/2)$

Hence *f(x) = sinx* is continuous at $\pi/2$.

In fact *f(x) = sinx* is a continuous function. Similarly, the functions *f(x) = cosx,* the **exponential function** with base a, f(x) = a^x, the **logarithmic function** with base a, f(x) = $\log_a x$, the **natural exponential** function f(x) = e^x and the **natural logarithmic** function f(x) = ℓnx are all continuous functions in their respective domains.

Theorem 4.6: Suppose f and g are functions with common domain such that both f and g are continuous at c. Then

1) f + g is continuous at c.
2) f – g is continuous at c.
3) if k is a scalar, kf is continuous at c.
4) fg is continuous at c.
5) if g(c) ≠ 0, $\frac{f}{g}$ is continuous at c.

Example 4.31: Let P(x) = $a_nx^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$ be any polynomial of degree n, and let c∈**R**, arbitrary. Then,

$$\lim_{x\to c} P(x) = \lim_{x\to c} (a_nx^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0) = a_nc^n + a_{n-1}c^{n-1} + \ldots + a_2c^2 + a_1c + a_0 = P(c)$$

Hence, P(x) is continuous at c, and since c was taken arbitrarily, every polynomial function is continuous.

Example 4.32: Let f(x) = $\frac{p(x)}{q(x)}$ be any rational function. Then if c is any real number such that q(c) ≠ 0, then $\lim_{x\to c} f(x) = \lim_{x\to c} \frac{p(x)}{q(x)} = \frac{\lim_{x\to c} p(x)}{\lim_{x\to c} q(x)} = \frac{p(c)}{q(c)} = f(c)$

Thus any rational function is continuous in its domain.

From the above theorem we can see that f(x) = $5x^2 - 4x + 7$ is continuous in **R,** g(x) = $\frac{x^3+x-1}{x^2-4}$ is continuous in **R**\{-2, 2}, h(x) = |x| cos x - $\frac{3}{x^2}$ is continuous for x ≠ 0 and f(x) = $\frac{5}{x-1}$ + ℓnx is continuous for x∈(0, 1) ∪ (1, ∞).

As a generalization of the Power Rule for limits, we have the following theorem

Theorem 4.7 (**Substitution Rule**): Suppose f and g are real valued functions such that $\lim_{x\to c} f(x)$ = L and g is continuous at L. Then $\lim_{x\to c} g(f(x)) = g\left(\lim_{x\to c} f(x)\right) = g(L)$

Example 4.33: For f(x) = sinx, g(x) = $\sqrt{x}$, and c = $\pi/2$, we have $\lim_{x\to \pi/2} f(x)$ = 1 and g is continuous at 1. Thus $\lim_{x\to c} g(f(x)) = \lim_{x\to c} \sqrt{\sin x} = \sqrt{1} = 1$.

Using Substitution Rule we have continuity of the composite of two functions as given by the following theorem.

Theorem 4.8: Suppose f and g are functions such that f is continuous at c and g is continuous at f(c). Then, gof is continuous at c.

Proof: Since f is continuous at c, $\lim_{x\to c}$ f(x) = f(c). Now $\lim_{x\to c}$ (gof)(x) = $\lim_{x\to c}$ g(f(x)) = g$\left(\lim_{x\to c} f(x)\right)$ = g(f(c)) = (gof)(c). Therefore, gof is continuous at c.

Example 4.34: For f(x) = $x^2 + 5$, g(x) = e^x and c = 1, we have (gof)(x) = g(f(x)) = e^{x^2+5} and $\lim_{x\to 1}$ (gof)(x) = $\lim_{x\to 1} e^{x^2+5}$ = e^6 = e^{1^2+5} = e^6. Thus, (gof)(x) = e^{x^2+5} is continuous at 1.

- **Intermediate Value Theorem**

Recall that for a function f continuous on a closed interval [a, b] its graph can be traced between the points (a, f(a)) and (b, f(b)) without any break or interruption. In this section we shall see an important application of continuous functions: namely, the Intermediate Value Theorem, and some of its consequences.

For a function continuous on [a, b], the intermediate value property asserts that if L is any number between (intermediate to) f(a) and f(b), then there is at least one number c between a and b whose image under f is L. See Figure 4.6.

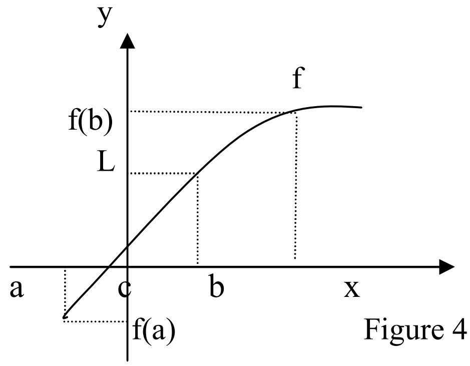


Figure 4.6

Theorem 4.9: (Intermediate Value Theorem) Suppose f is continuous on a closed interval [a, b]. Let L be any number between f(a) and f(b), (either f(a) ≤ L ≤ f(b), or f(b) ≤ L ≤ f(a). Then there exists a number c in [a, b] such that f(c) = L.

Example 4.35: Let f(x) = x^2. Then f is continuous on [0, 3] with f(0) = 0 and f(3) = 9. By the Intermediate Value Theorem f assumes (takes on) every value between 0 and 9. For instance for L = 4, we have 2 ∈[0, 3] with f(2) = 4, and for L = 7, we have $\sqrt{7}$ ∈[0, 3] with f$\left(\sqrt{7}\right)$ = 7.

Example 4.36: Let f(x) = $x^3 + 2x^2 + x = 4$ on [-2, 1]. Show that there exists some c ∈[-2, 1] such that f(c) = 4.

Solution: f is continuous on [-2, 1] with f(-2) = 2 and f(10 = 8. Since $2 \le 4 \le 8$, it follows, by the Intermediate Value Theorem that there exists c $\in$[-2, 1] such that f(c) = 4. i.e. f(c) = $c^3 + 2c^2 + c + 4 = 4$. In this case we can find such c by solving

$c^3 + 2c^2 + c + 4 = 4$

$\Leftrightarrow \quad c^3 + 2c^2 + c = 0$

$\Leftrightarrow \quad c(c^2 + 2c + 1) = 0$

$\Leftrightarrow \quad c\,(c + 1)^2 = 0$ which gives either c = 0 or c = -1

Since both of these values are in [-2, 1], for this particular case we have two values in [-2, 1] with image under f equal to 4.

One of the most important applications of the Intermediate Value Theorem is given in the following theorem.

Theorem 4.10: Suppose f is continuous on a closed interval [a, b] and assume that f(a) and f(b) have opposite signs. Then there is at least one c $\in$(a, b) such that f(c) = 0.

Proof: Without loss of generality, assume that f(a) < 0 and f(b) > 0. Then choose L = 0, between f(a) and f(b). By the Intermediate Value Theorem, there is at least one c between a and b such that f(c) = L = 0.

Remark: *This means that the equation f(x) = 0 has at least one root in the interval (a, b).*

Example 4.37: The function $f(x) = x^3 - x - 2$ is continuous on [1, 2]. f(1) = -2 < 0 and f(2) = 4 > 0. Thus there is a number c in (1, 2) such that f(c) = 0 or $c^3 - c - 2 = 0$.

Example 4.38: Show that the graphs of $y = e^x$ and y = 3x intersect in the interval [0, 1]

Solution: Define the function $f(x) = e^x - 3x$. Then f is continuous on [0, 1] with $f(0) = e^0 - 3(0) = 1 - 0 = 1 > 0$ and $f(1) = e^1 - 3(1) = e - 3 < 0$. Thus there is a number c $\in$(0, 1) such that $f(c) = e^c - 3c = 0$ and the graphs of $y = e^x$ and y = 3x intersect at c$\in$(0, 1).

Exercise 4.1

1. Evaluate the following limits, if they exist.

   a. $\lim_{x\to 4}(7-2x)$ b. $\lim_{x\to 2}\frac{2x-1}{3x+1}$ c. $\lim_{x\to 3}\frac{\sqrt{x+1}}{2x-1}$

   d. $\lim_{x\to -1}\frac{x^2-1}{x+1}$ e. $\lim_{x\to 2}\frac{x-2}{\sqrt{x+2}-2}$ f. $\lim_{x\to 1}\frac{\frac{1}{x}-1}{x-1}$

2. Find $\lim_{x\to c^-} f(x)$, $\lim_{x\to c^+} f(x)$ and $\lim_{x\to c} f(x)$, if it exists, for

   a. f(x) = cos x , at c = $\pi/6$

   b. f(x) = $\sqrt{x+3}$, at c = -3

c. $f(x) = \begin{cases} x, & \text{for } x < 2 \\ 2x-2, & \text{for } x > 2 \end{cases}$, at c = 2

d. $f(x) = \begin{cases} |x+1|, & \text{for } x \geq -2 \\ 1 \quad , & \text{for } x < 2 \end{cases}$, at c = -2

3. Evaluate each of the following limits, if it exists.

a. $\lim_{x\to 3} (2x^2 - 3x + 5)$ b. $\lim_{x\to 0} 2^x \sin x$ c. $\lim_{x\to \pi/4} (\cos x)^4$

d. $\lim_{x\to 263} \sqrt[3]{\dfrac{x^2-4x+3}{x-3}}$ e. $\lim_{x\to 4} \left(x^{1/2} - x^{3/2} \right)$ f. $\lim_{x\to 1} \sqrt[4]{7x^3+9}$

g. $\lim_{x\to 0} \dfrac{\sin x}{\tan x}$ h. $\lim_{x\to 0} \dfrac{\sin(2x)\sin(3x)}{5x^2}$ i. $\lim_{x\to 0} \dfrac{\sin 3x}{\sin 4x}$

4. Evaluate the following limits, if they exist

a. $\lim_{x\to 3^-} \dfrac{2}{x-3}$ b. $\lim_{x\to 1} \dfrac{x}{(x-1)^2}$ c. $\lim_{x\to \infty} \dfrac{4}{x-1}$

d. $\lim_{x\to -\infty} \cos x$ e. $\lim_{x\to \infty} \dfrac{x^3-1}{2x^2}$ f. $\lim_{x\to \infty} \dfrac{2x^3+3x-5}{5x^3+1}$

5. Find all the asymptotes, if any, for the following functions

a. f(x) = tanx b. $f(x) = \dfrac{x^2-9}{x+4}$ c. $f(x) = \dfrac{3x^3+2x+1}{(x+1)^2}$

6. Evaluate the following limits, if they exist.

a. $\lim_{t\to \infty} \left(1+\dfrac{1}{x}\right)^{x+4}$ b. $\lim_{t\to -\infty} \left(1+\dfrac{4}{x+1}\right)^{x}$ c. $\lim_{t\to \infty} \left(\dfrac{2x+3}{2x+1}\right)^{x-1}$

7. Check whether or not the following functions are continuous at the indicated points.

a. $f(x) = x^2 + 1$, at c = 2 b. $f(x) = |x^2 - 1|$, at c = -1, 0, 1

c. $f(x) = \dfrac{3}{x-2}$, at c = 2 d. $f(x) = \begin{cases} x^2, & \text{for } x \neq 1 \\ 3, & \text{for } x = 1 \end{cases}$, at c = 1

8. Show that the following equations have roots in the indicated intervals.

a) logx = 0, in $\left[\dfrac{1}{2}, 2\right]$ b) $2^x - 2 = 0$, in [0, 2] c) cos x – x = 0, in $\left[0, \pi/2\right]$

9. Using the Intermediate Value Theorem show that the graphs of f and g intersect in the given interval.

a. $f(x) = x^3 + 4x + 2$ and g(x) = -1, in [-1, 0]
b. f(x) = 2sinx and g(x) = 1 – x, in [0, 2]
c. $f(x) = x \ln x$ and g(x) = sinx, in $\left[\dfrac{1}{e}, e\right]$

4.2. Derivatives

Objectives

At the end of this section you should be able to

- get acquainted with the concept of the derivatives of a function.
- evaluate the derivative of elementary functions using the definition.
- find the slope and equation of a tangent line to a curve at a given point.
- evaluate the derivatives of combinations of functions.
- find the derivatives of polynomial and rational functions.
- have a good understanding of the Chain Rule.
- apply the Chain Rule to evaluate derivatives of composite functions and algebraic functions.
- find the derivative of the logarithmic function.
- find the derivative of the exponential function.
- apply the above derivatives to the natural logarithmic and natural exponential functions as special cases.
- evaluate derivatives of composite functions with the logarithmic and exponential functions.
- have an understanding of the derivative of a derivative.

Using the concepts discussed in section 4.1, we are now ready to study one of the central concepts of calculus: the derivative of a function. Even though the derivative is connected with finding the tangent lines to curves at a point, its main applications are in finding rates of change of variable quantities relative to the change in another quantity.

Consider a function f continuous at a point c in its domain.

Then, by definition of continuity $\lim_{x \to c} f(x)$ = f(c)

This means for x close to c, f(x) is close to f(c). If we denote the **increment** (or change) x – c in the x-direction by h = x – c (so that x = c + h) as is seen in Figure 4.7, then the corresponding change in the y-direction is given by

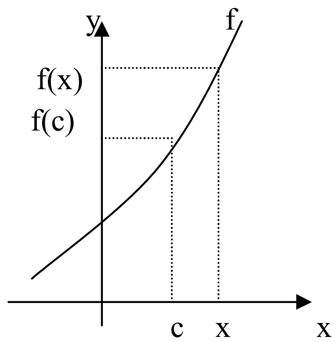


Figure 4.7

$$f(x) - f(c) = f(c + h) - f(c).$$

The ratio of these two increments is given by

$$\frac{f(x) - f(c)}{x - c}$$

and is called the **difference quotient of f at c.**
For instance, if f(x) = x^2 + 2 and c = 3, then

$$\frac{f(x)-f(3)}{x-3} = \frac{(x^2+2)-(3^2+2)}{x-3} = \frac{x^2-9}{x-3}$$

We shall define the derivative of a function of f at c as the limit of the above difference quotient, if the limit exists.

Definition 4.11
Let c be a number in the domain of a function f. If

$$\lim_{x\to c}\frac{f(x)-f(c)}{x-c}$$

exists, we call this limit the **derivative of f at** c, and denote it by f'(c), so that

$$f'(c) = \lim_{x\to c}\frac{f(x)-f(c)}{x-c}$$

If this limit exists we say that f **has a derivative at c**, or f is **differentiable at c** or f'**(c) exists.**

Remarks: 1. *Observe that we can alternatively write*

$$f'(c) = \lim_{h\to 0}\frac{f(c+h)-f(c)}{h}$$

since for h = x – c, we have x = c + h and as x → c, h → 0.

2. *The notation f′(c) is read as "the derivative of f at c" or for short "f prime at c".*
Other notations are given by $\frac{df}{dx}(c)$ *or Df(c)*

3. *The quantity f′(c) describes the rate of change of the function f around the point (c, f(c)).*

Example 4.39: Let f(x) = 2x + 3. Then, for any c ∈**R**, the point (c, f(c)), we have

$$f'(c) = \lim_{x\to c}\frac{f(x)-f(c)}{x-c} = \lim_{x\to c}\frac{(2x+3)-(2c+3)}{x-c} = \lim_{x\to c}\frac{2x-2c}{x-c} = 2\lim_{x\to c}\frac{x-c}{x-c} = 2\lim_{x\to c}(1) = 2.$$

Since c∈**R** is arbitrarily taken, we have for f(x) = 2x + 3, f'(x) =2 for all x∈**R.**
In fact for any linear function f(x) = ax + b, we have

$$f'(c) = \lim_{x\to c}\frac{f(x)-f(c)}{x-c} = \lim_{x\to c}\frac{(ax+b)-(ac+b)}{x-c} = a\lim_{x\to c}\frac{x-c}{x-c} = a$$

for any c ∈**R**. Thus f'(x) = a
Note that the graph of a linear function is a straight line and the rate of change (a constant) is measured by the slope of the line.

Example 4.40: Let f(x) = $3x^2$ + 5. Then for any x∈**R**

f(x + h) = $3(x + h)^2$ + 5 = $3x^2$ + 6xh + $3h^2$ + 5 and

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}\frac{(3x^2+6xh+3h^2+5)-(3x^2+5)}{h}$$

$$= \lim_{h\to 0}\frac{6xh+3h^2}{h} = \lim_{h\to 0}\frac{3h(2x+h)}{h} = 3\lim_{h\to 0}(2x+h) = 6x.$$

Thus, for f(x) = $3x^2$ + 5, f'(x) = 6x for any x∈**R.**

In particular, when c = 1, $f'(1) = 6(1) =$. 6 is the slope of the tangent line to the graph of f at (1,8)

Example 4.41: Let f(x) = c, where c is a constant.

Then for any $x \in \mathbf{R}$,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{c-c}{h} = \lim_{h \to 0} 0 = 0.$$

Thus, for f(x) = c, a constant, $f'(x) = 0$ for all $x \in \mathbf{R}$.

Hence, for f(x) = 15, $f'(x) = 0$, for $f(x) = -\sqrt{2}$, $f'(x) = 0$, and so on.

Applying the above definition, we can get the following derivatives.

$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
$f(x) = \frac{1}{x}$	$f'(x) = \frac{-1}{x^2}$, for all $x \neq 0$	f(x) = sinx	$f'(x) = \cos x$, for all $x \in \mathbf{R}$
$f(x) = \sqrt{x}$	$f'(x) = \frac{1}{2\sqrt{x}}$, for x > 0	$f(x) = \cos x$	$f'(x) = -\sin x$, for all $x \in \mathbf{R}$,

Using the definition to evaluate the derivative of more complicated combinations and compositions of functions becomes cumbersome. At this stage the student must be able how to find the derivatives of various types of functions quickly and efficiently without always resorting to the definition. In the table below we list some techniques of differentiation which can be proved using the definition.

Theorem 4.11: Suppose f and g are differentiable at c, and k is a constant, then

a)	$(kf)'(c) = k\,f'(c)$ …	**Constant Rule**
b)	$(f+g)'(c) = f'(c) + g'(c)$ …	**Addition Rule**
c)	$(f-g)'(c) = f'(c) - g'(c)$ …	**Difference Rule**
d)	$f(x) = x^n$, n an integer, $f'(x) = n\,x^{n-1}$ …	**Power Rule**
d)	$(fg)'(c) = f'(c)\,g(c) + f(c)\,g'(c)$ …	**Product Rule**
e)	$\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}$ provided $g(c) \neq 0$ …	**Quotient Rule**
f)	$(gof)'(c) = g'(f(c)).\,f'(c)$ …..	**The Chain Rule**

Thus, if $f(x) = x^4$, then $f'(x) = 4x^3$ and if $g(x) = x^{12}$, then $f'(x) = 12x^{11}$, and so on.

Example 4.42: Let $f(x) = x^2 + 3$ and g(x) = sinx. Then

$$(f+g)'(x) = f'(x) + g'(x) = \frac{d}{dx}(x^2+3) + \frac{d}{dx}(\sin x) = 2x + 0 + \cos x \quad = 2x + \cos x$$

$$\frac{d}{dx}(g(x) - 4f(x)) = \frac{d}{dx}(\sin x) - 4\frac{d}{dx}(x^2+3) = \cos x - 4(2x+0) = \cos x - 8x.$$

Since polynomials are sums or differences of constant multiples of powers of x, the first four rules help us to evaluate their derivatives.

Remark: *Given a polynomial of degree n,* $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$

$$P'(x) = \frac{d}{dx}(a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0)$$

$$= a_n\frac{d}{dx}(x^n) + a_{n-1}\frac{d}{dx}(x^{n-1}) + \ldots + a_2\frac{d}{dx}(x^2) + a_1\frac{d}{dx}(x) + \frac{d}{dx}(a_0).$$

$$= na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \ldots + 2a_2x + a_1$$

Example 4.43: For $p(x) = 5x^4 - 2x^3 + x^2 + 7x - 1$, we have $p'(x) = 20x^3 - 6x^2 + 2x + 7$.

For $q(x) = 6x^3 + \sqrt{2}\,x^2 - 3x + \pi$, we have $q'(x) = 18x^2 + 2\sqrt{2}\,x - 3$.

As an application of the product rule, we have the following examples.

Example 4.44: Let $k(x) = 2x \sin x$. Find $k'(x)$.

Solution: If we put $f(x) = 2x$ and $g(x) = \sin x$, then $f'(x) = 2$ and $g'(x) = \cos x$.

Thus, $k'(x) = f'(x)g(x) + f(x)g'(x) = 2\sin x + 2x\cos x$.

Remark: *In practice, to evaluate the derivative of a product of two functions, we do not need to identify which one is f and which one is g.*

Example 4.45: Let $h(x) = x^3 \cos x$. Then

$$h'(x) = (x^3)' \cos x + x^3(\cos x)' = 3x^2\cos x + x^3(-\sin x) = 3x^2\cos x - x^3 \sin x.$$

For the derivative of the product of three functions f, g and h, we have

$$(fgh)'(x) = f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x).$$

Example 4.46: Let $k(x) = x^3 \sin x \cos x$. Find $k'(x)$.

Solution: Put $f(x) = x^3$, $g(x) = \sin x$ and $h(x) = \cos x$ in the above statement with $f'(x) = 3x^2$, $g'(x) = \cos x$ and $h'(x) = -\sin x$. Then $k'(x) = 3x^2 \sin x \cos x + x^3 \cos x.\cos x + x^3\sin x(-\sin x) = 3x^2 \sin x \cos x + x^2 \cos^2 x - x^3 \sin^2 x$.

The Quotient Rule is used to find the derivative of any rational function. If $f(x) = \frac{p(x)}{q(x)}$, *for p, q polynomials, we then have* $f'(x) = \left(\frac{p(x)}{q(x)}\right)^1 = \frac{p'(x)q(x) - p(x)q'(x)}{(q(x))^2}$, *for* $q(x) \neq 0$.

Example 4.47: Let $f(x) = \frac{3x^2 - 5}{2x + 1}$. *Find* $f'(x)$

Solution: Putting $p(x) = 3x^2 - 5$ and $q(x) = 2x + 1$, we get

$$f'(x) = \frac{6x(2x+1) - (3x^2 - 5)(2)}{(2x+1)^2} = \frac{12x^2 + 6x - 6x^2 + 10}{(2x+1)^2} = \frac{6x^2 + 6x + 10}{(2x+1)^2}$$

As an important consequence of the Quotient Rule, we can now find the derivatives of the remaining four trigonometric functions.

Example 4.48: Let $f(x) = \tan x$. Show that $f'(x) = \sec^2 x$

Solution: f(x) = tanx = $\frac{\sin x}{\cos x}$. Then

$$f'(x) = \frac{d}{dx}(\tan x) = \left(\frac{\sin x}{\cos x}\right)^1 = \frac{(\sin x)^1 \cos x - \sin x(\cos x)^1}{(\cos x)^2}$$

$$= \frac{\cos x.\cos x - \sin x(-\sin x)}{\cos^2 x}$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = sec^2x.$$

In the same manner, we can show that

$$\frac{d}{dx}(cotx) = -csc^2x,\ \frac{d}{dx}(secx) = secxtanx \text{ and } \frac{d}{dx}(\csc x) = -cscx.cotx.$$

The Chain Rule states that (gof)′(x) = g′(f(x)) f'(x), for all x such that f is differentiable at x and g is differentiable at f(x).

Example 4.49: Find the derivative of h(x) = cos(x^2 + 1)

Solution: Let f(x) = x^2 + 1 and g(x) = cosx. Then, h(x) = (gof)(x) = g(f(x)) = g(x^2 + 1) = cos(x^2 + 1) and h′(x) = g′(f(x)). f(x) = -sin(x^2 + 1) . (x^2 + 1)′ = 2xsin(x^2 + 1).

If a and b are any real numbers, we can easily show that

$$\frac{d}{dx}(\sin ax) = a \cos ax \quad \text{and} \quad \frac{d}{dx}(\cos bx) = -b\sin bx$$

Thus, $\frac{d}{dx}$(sin4x) = 4 cos4x and $\frac{d}{dx}$(cos5x) = -5 sin5x

Example 4.50: Find the derivative of h(x) = $(1+3x -5x)^{12}$

Solution: Let f(x) = 1 + 3x – x^5 and g(x) = x^{12}. Then h = gof and

$$h'(x) = \frac{d}{dx}(1+3x - x^5)^{12} = 12(1 + 3x - x^5)^{11}\ (1 + 3x - x^5)' = 12(3 - 5x^4)\ (1 + bx - x^5)^{11}.$$

Example 4.51: Find the equations of tangent and normal lines to the semicircle

$$y = f(x) = \sqrt{1-x^2} \text{ at } \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$$

Solution: The slope of the tangent line T is given by the derivative of y = f(x) = $\sqrt{1-x^2}$ at x = $\frac{1}{2}$. Thus, by Chain Rule,

$$f'(x) = \frac{dy}{dx} = \frac{d}{dx}\sqrt{1-x^2} = \frac{1}{2\sqrt{1-x^2}}(1-x^2)' = \frac{-2x}{2\sqrt{1-x^2}} = \frac{-x}{\sqrt{1-x^2}}$$

so that the slope of T is

$$m = f'\left(\frac{1}{2}\right) = \frac{-1/2}{\sqrt{1-1/4}} = -\frac{1}{2}.\frac{2}{\sqrt{3}} = \frac{-1}{\sqrt{3}}$$

and since the tangent line passes through the point $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, its equation in slope-point form is

$$y - \frac{\sqrt{3}}{2} = -\frac{1}{\sqrt{3}}\left(x - \frac{1}{2}\right) \quad \text{or} \quad x + \sqrt{3}\,y - 2 = 0$$

The slope of the normal line at $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ is $\sqrt{3}$ and its equation is

$$y - \frac{\sqrt{3}}{2} = \sqrt{3}\left(x - \frac{1}{2}\right) \quad \text{or} \quad y - \sqrt{3}\,x = 0.$$

Remark: *The Chain Rule can be extended to more than two functions.*
Suppose k(x) = (hogof)(x) = h(g(f(x))) and let f be differentiable at x, g be differentiable at f(x) and h be differentiable at g(f(x)). Then k′(x) = (hogof)′(x) = h′(g(f(x)).g′f(x)). f′(x)
Similarly, if ℓ(x) = (kohogof)(x) = k(h(g(f(x)))), then

ℓ′(x) = *(kogohof)′(x) = k′(h(g(f(x)))). h′(g(f(x)).g′(f(x)).f′(x).*

You can now see why this method is called the Chain Rule!

Example 4.52: Find the derivative of the function

$$k(x) = \cos\sqrt{2x^2 - 3}$$

Solution: Let $k(x) = \cos x\sqrt{2x^2 - 3} = h(g(f(x)))$ with

$f(x) = 2x^2 - 3$, $g(x) = \sqrt{x}$ and $h(x) = \cos x$. Then

$$k'(x) = \frac{d}{dx}\left(\cos\sqrt{2x^2 - 3}\right) = h'(g(f(x)).\ g'(f(x))\ f'(x)$$

$$= -\sin\sqrt{2x^2 - 3}\ .\ \frac{1}{2\sqrt{2x^2 - 3}}.4x = \frac{-2x\sin\sqrt{2x^2 - 3}}{\sqrt{2x^2 - 3}}$$

Example 4.53: Let $f(x) = \sin(\tan x^2)$. Find $f'(x)$

Solution: $f'(x) = \frac{d}{dx}(\sin(\tan x^2))$

$$= \cos(\tan x^2)\sec^2 x^2(2x) = 2x.\cos(\tan x^2)\sec^2 x^2.$$

- **Derivatives of Logarithmic and Exponential Function**

Recall that for $a > 0$, and $a \neq 1$, the logarithmic function with base a is given by

$$f(x) = \log_a x \qquad \text{for } x > 0.$$

In particular, when a = e, we get the natural logarithmic function

$$f(x) = \log_a x = \ell n x \quad , \qquad \text{for } x > 0.$$

Theorem 4.12: Let a > 0, a ≠ 1 and let f(x) = $\log_a x$. Then

$$f'(x) = \frac{1}{x}\log_a x$$

From Theorem 4.12, when the base a = e , it follows that

$$(\log_e x)' = (\ell n x)' = \frac{1}{x}\log_e e = \frac{1}{x}.1 = \frac{1}{x} \text{ : i.e. } \boxed{(\ln x)' = \frac{1}{x}}$$

Also, by applying change of base of logarithms, we get

$$(\log_a x)' = \frac{1}{x}\log_a e = \frac{1}{x}\frac{\log_e e}{\log_e a} = \frac{1}{x}.\frac{1}{\ln a} \text{: i.e. } \boxed{(\log_a x)' = \frac{1}{x\ln a}}$$

Example 4.54: For $f(x) = \log_5 x$, we have $f'(x) = \frac{1}{x}\log_5 e = \frac{1}{x\ln 5}$

Example 4.55: Find the derivative of the following

a) f(x) = $\log_3(x^2 + x - 1)$ b) g(x) = $\frac{x}{\ln x}$

Solution: a) $f'(x) = \frac{1}{x^2+x-1}.\log_3 e\ (x^2 + x - 1)' = \frac{2x+1}{x^2+x-1}.\log_3 e$

b) $g'(x) = \left(\frac{x}{\ln x}\right)' = \frac{1.\ln x - x.\frac{1}{x}}{(\ln x)^2} = \frac{\ln x - 1}{\ln^2 x}$

Theorem 4.13: Let a > 0, a ≠ 1 and let f(x) = a^x. Then, $f'(x) = (a^x)' = \frac{a^x}{\log_a e}$.

By applying change of base we also have

$$\boxed{(a^x)' = \frac{a^x}{\log_a e} = a^x \ell n\, a}$$

When the base a = e, we get $(e^x)' = \frac{e^x}{\log_e e} = e^x \ell n e = x^x.1 = e^x$*: i.e.* $\boxed{(e^x)' = e^x}$

Example 4.56: For f(x) = 3^x, we have $f'(x) = \frac{3^x}{\log_3 e} = 3^x\, \ell n3$.

Example 4.57: Find the derivative of the following

a) f(x) = $e^{\sqrt{x+1}}$ b) g(x) = $3^{\sin x}$

c) f(x) = $\sqrt{x + e^{4x}}$ d) g(x) = $e^{x^2}\, \ell n x$

Solution: a) By using the Chain Rule, we get

$$f'(x) = \left(e^{\sqrt{x+1}}\right)' = e^{\sqrt{x+1}}.\left(\sqrt{x+1}\right)' = e^{\sqrt{x+1}}, \frac{1}{\sqrt{x+1}} = \frac{e^{\sqrt{x+1}}}{\sqrt{x+1}}$$

b) $(3^{\sin x})' = 3^{\sin x}, \ell n3. (\sin x)' = \ell n3.\cos x.3^{\sin x}$

c) $\frac{d}{dx}\sqrt{x+e^{4x}} = \frac{1}{2\sqrt{x+e^{4x}}}(x+e^{4x})' = \frac{1+4e^{4x}}{2\sqrt{x+e^{4x}}}$

d) By the Product Rule and Chain Rule we get:

$$g'(x) = \left(e^{x^2}\ln x\right)' = 2xe^{x^2}\ell nx + e^{x^2}.\frac{1}{x} = e^{x^2}\left(2x\ln x + \frac{1}{x}\right)$$

- **Higher Derivatives**

If a function f is differentiable at a point x in its domain, we denote its derivative by $f'(x)$, where

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$, provided the limit exists.

This derivative is usually called the **first derivative of f at x.**

If the new function f' is differentiable at a point x, then we can repeat the process and find its derivative as

$$(f'(x))' = f''(x) = \lim_{h\to 0}\frac{f'(x+h)-f'(x)}{h},$$ provided the limit exists.

we call $f''(x)$ the **second derivative of f at x**, and it is often read as "f double prime of x".

Observe that $f''(x)$ is simply the derivative of the function f' at x and is no more difficult than finding the first derivative.

Example 4.58: If $f(x) = \ell nx$, then $f'(x) = \frac{1}{x}$ and hence $f''(x) = \frac{d}{dx}\left(\frac{1}{x}\right) = \frac{-1}{x^2}$.

We can similarly find the derivative of $f''(x)$ to get

$$(f''(x))' = f'''(x) = \lim_{h\to 0}\frac{f''(x+h)-f''(x)}{h}$$, and so on,

and call this the **third derivative of f at x**.

Thus, for $f(x) = \ell nx$, $f'(x) = \frac{1}{x}$, $f''(x) = \frac{-1}{x^2}$ and $f'''(x) = \frac{2}{x^3}$

These derivatives when they exist are called ***higher derivatives (or derivatives of derivatives)***

The n^{th} derivative $f^{[n]}(x)$ can also be denoted by $\frac{d^n f}{dx^n}(x)$.

Thus the second derivative is $f''(x)$ or $f^{[2]}(x)$ or $\frac{d^2 f}{dx^2}(x)$ and the third derivative is $f'''(x)$ or $f^{[3]}(x)$ or $\frac{d^3 f}{dx^3}(x)$.

Example 4.59: Find the higher derivatives of the following

a) $f(x) = 4x^3 + x^2 - 3x + 7$ b) $g(x) = e^x$

c) $f(x) = \sin x$ d) $g(x) = \ell n(3x)$

Solution: a) For $f(x) = 4x^3 + x^2 - 3x + 7$, we have

$f'(x) = 12x^2 + 2x - 3$

$f''(x) = 24x + 2$

$f'''(x) = 24$

$f^{[4]}(x) = 0$ and for $n \geq 4$, $f^{[n]}(x) = 0$

b) For $g(x) = e^x$, $g'(x) = e^x$, $g''(x) = e^x$, and in general for $n > 1$, $g^{[n]}(x) = e^x$

c) $f(x) = \sin x$, $f'(x) = \cos x$

$f''(x) = -\sin x$, $f'''(x) = -\cos x$

$f^{[4]}(x) = \sin x$ and so on

d) $g(x) = \ell n(3x)$, $g'(x) = \frac{3}{3x} = \frac{1}{x}$ by Chain Rule

$g''(x) = -\frac{1}{x^2}$, $g'''(x) = \frac{2}{x^3}$, $g^{[4]}(x) = -\frac{3.2}{x^4}, \ldots$

Exercise 4.2

1. For each of the following functions, find $f'(c)$ using the definition
   a. $f(x) = 2x - 4$, at $c = 1$ b. $f(x) = x^2 + 3$, at $c = -1$
   c. $f(x) = x^3 - 2$ at $c = 0$ d. $f(x) = |x + 2|$, at $c = 2$
2. Find the equations of the tangent and normal lines to the graph of f at the given point.
   a. $f(x) = x^2 + x - 1$, at (2, 5) b. $f(x) = \sqrt{x}$, at (4, 2)
   c. $f(x) = 2\cos x$, at $(\pi/2, 0)$ d. $f(x) = \frac{1}{x}$, at $(2, \frac{1}{2})$
3. Find the derivative of the following functions
   a. $f(x) = (x^2 - 5)\cos x$ b. $g(x) = \sqrt{x}\sec x$
   c. $\frac{2x^3 - 5x}{x^2 + 3}$ d. $g(x) = \frac{2x}{\tan x}$
4. Find the equations of the tangent and normal lines to the functions at the indicated point.
   a. $f(x) = \sin x \cos x$, at $(\pi/4, 1/2)$
   b. $f(x)\ \frac{x}{x^2 + 1}$, at $(1, 1/2)$

5. Find the derivative of the following functions.

a. $f(x) = \tan^3 x$ b. $g(x) = x\sqrt{1-x^2}$ c. f(x) = **Error! Objects cannot be created from editing field codes.** sin x^2

d. $g(x) = \dfrac{1}{(x^2-x)^4}$ e. $f(x) = \sqrt{x^2+\sqrt{x^2+1}}$ f. $g(x) = x\cos x + \sqrt[3]{5x-4}$

g. $f(x) = \sin\sqrt{2x+1}$ h. $g(x) = \dfrac{\cot x}{\sqrt{x+1/x}}$ i. $f(x) = e^{\tan x}$

j. $g(x) = \ell n(\ell n x)$ k. $f(x) = (\ell n x + e^{\sqrt{x}})^3$ l. $f(x) = \ell n^2 x + \ell n x^2$

6. Find the first, second and third derivatives of the following function

a. $f(x) = e^{x^2}$ b. $g(x) = \sec x$

c. $f(x) = \sin(2x) + \cos(3x)$ d. $g(x) = \ell n(\sin x)$

4.3. Applications of the derivative

At the end of this section you should be able to:

- define maximum and minimum values of a function on a given interval.
- explain the fundamental theorem of local extrema values.
- identify the regions where a function is increasing and decreasing.
- apply the first and second derivative tests to find local extrema values of a function.
- solve practical problems related to extrema.
- state the important points that are necessary to sketch the graph of a function.
- sketch the graph of a function applying the above concepts.
- solve related rates problems.

At the beginning of this unit we have mentioned that the derivative of a function at a point c in its domain measures the rate of change of the function around that point. In this section we shall see how the derivative can be applied to solve a variety of problem in the areas of engineering, the natural sciences, business and the social sciences. We see how it can be used to solve maximum and minimum values of a function (i.e., where it has "peaks" and where it has "valleys"), where it curves upward and where it curves downward, and in general, to sketch the graph of the function. At the end we shall introduce related rates problems and see how to solve them using the derivative.

a) Extrema of a Function

Definition. Let f be a function defined on an interval I. If there is a number d in I such that $f(x) \le f(d)$ for all x in I, then f(d) is called the **maximum value of f**

on I. Similarly, if there is a number c in I such that f(x) ≥ f(c) for all x in I, then f(c) is called the **minimum value of f on I**. (See Figure 4.8) A value of f that is either a maximum value or a minimum value of f on I is called an **extreme value of f on I**.

Remark: *If the set I is the domain of the function f and if f has a maximum value on I, then this maximum value is called the **(absolute or global) maximum** of f.*
Similar for minimum value of f.

Example 4.60: Let $f(x) = x^2$ on I = [-2, 4]. Then f has the maximum value of 16 = f(4) and the minimum value of 0 = f(0). Both 0 and 16 are extreme values of f.
- On the interval [-2, 4), the minimum value of f is 0 but f has no maximum.
- On the interval (0, 4) f has neither a maximum nor a minimum. See Figure 4.8.

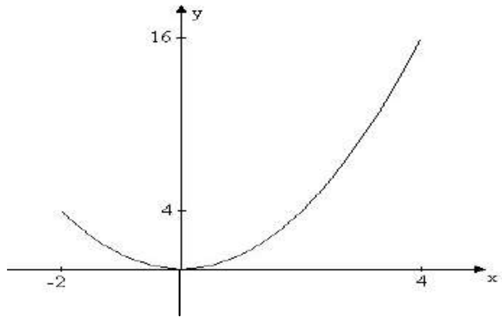


Figure 4.8

Example 4.61: Let $f(x) = \frac{1}{x}$ for x ≠ 0.

The domain of f is I = (-∞, 0) ∪ (0, ∞) and f has neither a maximum nor a minimum value on I. See Figure 4,10
- On the interval [-1, 0] f has the maximum value -1 = f(-1), but no minimum.
- On the interval (0, 2] f has the minimum value $\frac{1}{2} = f(2)$, but has no maximum.
- On the interval [-1, 2], f has no extrema.

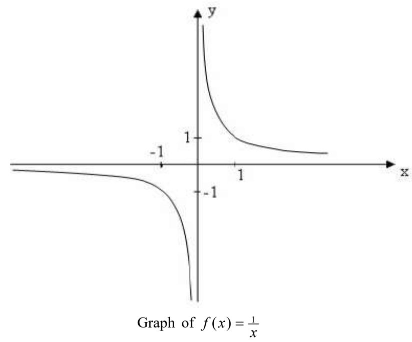


Figure 4.9

Note that in the first example when the interval is open we have no extrema, while in the second example, when the function is not continuous, we had no extrema. Continuity of a function on a closed interval gives us a sufficient condition for the existence of both extreme values.

Theorem 4.14: (Maximum-Minimum Theorem). Let f be continuous on a closed bounded interval [a, b]. Then f has a maximum and a minimum value on [a, b].

Hence the function $f(x) = x^2$ for $-2 \le x \le 4$ has both extreme values on [-2, 4].
Similarly, the function $f(x) = x^3 - 4x + 5$ for $0 \le x \le 2$ which is continuous on [0, 2] has a maximum and a minimum value on [0, 2], by Theorem 4.14. Even though the above theorem tells us about the existence of extreme values on [a, b], it does not tell us where they occur or how to find them. The following theorem will help us in determining such values.

Theorem 4.15: Let f be defined and continuous on [a, b]. If f has an extreme value at c in (a, b) and f is differentiable at c, then $f'(c) = 0$.

Example 4.62: Let $f(x) = x^3 - 3x + 1$.
Then f is differentiable and the critical points of f are the values of x for which $f'(x) = 0$,

But $f'(x) = 3x^2 - 3 = 0$

$\Leftrightarrow 3(x - 1)(x + 1) = 0$

$\Rightarrow$ x = 1 and x = -1 are critical points of f.

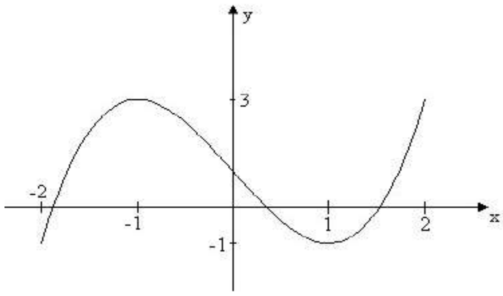


Figure 4.10

If we want to find extreme values of f on, say, the interval [-3, 3] we compute and compare the values of f at -3, -1, 1 and 3 to get f(-3) = 17 , f(-1) = 3, f(1) = -1 and f(3) = 19.
Thus the minimum value of f on [-3, 3] is -17 which occurs at -3 and the maximum value of f is 19 which occurs at 3.

- **Monotonic Functions**

One of the important points needed to sketch the graph of a function is to find the regions in which the graph slopes upward to the right (increases) or it slopes downward to the right (decreases) as seen in Figure 4.11 (a) and (b), respectively.

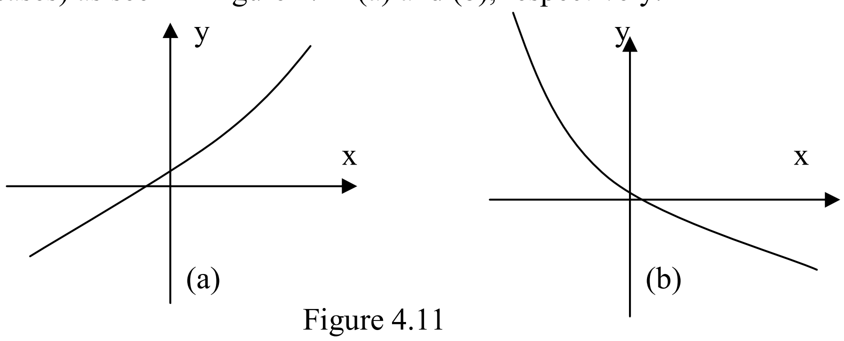


Figure 4.11

Definition 4.12: Suppose f is a function defined on an interval I.

i) f is said to be **increasing** on I if $f(x_1) \le f(x_2)$ whenever $x_1 < x_2$

ii) f is said to be **decreasing** on I if $f(x_1) \ge f(x_2)$ whenever $x_1 < x_2$

iii) f is said to be **monotonic** on I if f is either increasing or decreasing on I.

Remark: *we can similarly define the terms **strictly increasing, strictly decreasing** and **strictly monotonic** by replacing $\le$ by $<$ and $\ge$ by $>$.*

Example 4.63: Let $f(x) = x^2 - 1$.
Find the intervals of monotonicity of f.

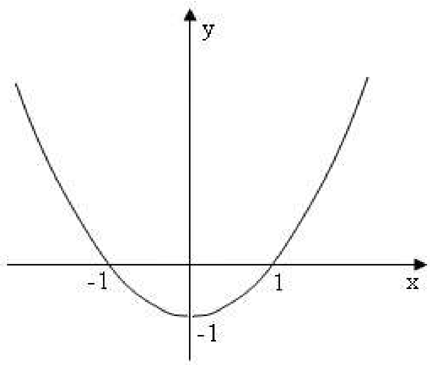
Solution: For $x_1, x_2 \in (-\infty, 0)$ with $x_1 < x_2$, we have

$f(x_1) = x_1^2 - 1 > x_2^2 - 1 = f(x_2)$

$\Rightarrow$ f is strictly decreasing on $(-\infty, 0)$.

For $x_1, x_2 \in (0, \infty)$ with $x_1 < x_2$, we have

$f(x_1) = x_1^2 - 1 < x_2^2 - 1 = f(x_2)$



$\Rightarrow$ f is strictly increasing on $(0, \infty)$. Figure 4. 12

The derivative of a function gives us a test for monotonicity as is indicated in the following theorem.

> **Theorem: 4.16** Suppose f is continuous and differentiable on an interval I.
>
> i) If $f'(x) > 0$, for every $x \in I$, then f is strictly increasing on I.
>
> ii) If $f'(x) < 0$, for every $x \in I$, then f is strictly decreasing on I.

Example 4.64: Find the intervals over which the following function $f(x) = x^3 - 3x + 1$is monotonic.

Solution: For $f(x) = x^3 - 3x + 1$, $f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1)$

To find the intervals over which f is increasing and decreasing we find the sign of $f'(x)$ using the critical points x = 1 and x = -1 and the Sign Chart Method.

	-1		1	
x – 1	- - - - - - -	- - - - - - - -	0	+ + + + + + ++ + +
x + 1	- - - - - - -	0 + ++ + + + + + + + +		+ + + + + +
$f'(x)$	+ + + + +	0 - - - - - -	0	+ + + + + + + + +

From the above "sign chart" we can see that

$f'(x) > 0$ for $x \in (-\infty, -1) \cup (1, \infty)$ and $f'(x) < 0$ for $x \in (-1, 1)$.

Thus f is strictly increasing on $(-\infty, -1) \cup (1, \infty)$ and strictly decreasing on $[-1, 1]$.See Figure 4.10.

- **The First and Second Derivative Tests for Relative Extrema**

If f is a differentiable function, we have seen that at relative extreme values $f'(c) = 0$. Thus in order to locate relative extreme values of f we find the values of x for which $f'(x) = 0$ or $f'(x)$ does not exist. But this method does not help us to determine which of these values of x give relative extreme values (or which value is a maximum or which is a minimum). The next two theorems will provide us with conditions that guarantee that f has relative extreme values. These conditions will also help in sketching the graphs of functions and in solving applied problems.

> **Theorem 4.17: (The First Derivative Test)**
>
> Let f be continuous on an interval I, and let $c \in I$.
>
> a) If $f'(x)$ changes its sign from positive to negative at c
> i.e. if $f'(x) > 0$ to the left of c and $f'(x) < 0$ to the right of c, then f has a relative maximum value at c.
>
> b) If $f'(x)$ changes its sign from negative to positive at c, then f has a relative minimum value at c.

Example 4.65: Consider again the function f(x) = $x^3 - 3x + 1$.
$f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1) = 0$ gives the critical points x = 1 and x = -1
For the critical point x = -1 check the sign of f' at -2 and 0 with $f'(-2) = 9 > 0$ and $f'(0) = -3 < 0$.
Thus $f'(-1) = 3$ is a relative maximum value of f.
Similarly taking the critical point x = 1 between 0 and 2, we get $f'(0) = -3 < 0$ and $f'(2) = 9 > 0$.
Thus f(1) = -1 is a relative minimum value of f. (See Figure 4.11 above)

The above theorem needs to check the signs of two distinct points to the left and to the right of each critical point. The next theorem makes use of the sign of the second derivative directly at the critical points.

Theorem: 4.18 (The Second Derivative Test)
Let f be differentiable in an interval I and let c $\in$I with $f'(c) = 0$.

a) If $f''(c) < 0$, then $f(c)$ is a relative maximum value of f.
b) If $f''(c) > 0$, then $f(c)$ is a relative minimum value of f.

If $f''(c) = 0$, then we can not draw any conclusion about f(c).

Example 4.66: Consider again the function f(x) = $x^3 - 3x + 1$ with $f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1)$. We have $f'(1) = f'(-1) = 0$ and $f''(x) = 6x$. Since $f''(-1) = -6 < 0$, $f(-1) = 3$ is a local maximum value of f. Since $f''(1) = 6 > 0$, $f(1) = -1$ is a local minimum values of f.

Example 4.67: Let f(x) = $\dfrac{4x}{x^2+4}$, Find the local extreme values of f.

Solution: $f'(x) = \dfrac{4(x^2+4) - 4x(2x)}{(x^2+4)^2}$, Quotient Rule.

$$= \frac{4x^2+16-8x^2}{(x^2+4)^2} = \frac{16-4x^2}{(x^2+4)^2}$$

$f'(x) = 0 \Rightarrow 16 - 4x^2 = 0 \Rightarrow x = 2$ or $x = -2$

$f''(x) = \dfrac{-8x(x^2+4)^2 - (16-4x^2)2(x^2+4)(2x)}{(x^2+4)^4}$ - Quotient Rule and Chain Rule

$= \dfrac{8x(x^2-12)}{(x^2+4)^3}$ - Simplification.

Thus $f''(2) = \dfrac{16(-8)}{8^3} = \dfrac{-1}{4} < 0$ $\Rightarrow$ f(2) = 1 is a local maximum value of f and

$f''(-2) = \dfrac{-16(-8)}{8^3} = \dfrac{1}{4} > 0$ $\Rightarrow$ f(-2) = -1 a local minimum value of f.

- **Practical Applications of the Extrema**

A lot of practical problems can be expressed as a continuous function on a closed and bounded interval we may be interested to find points where f attains its maximum or its minimum values. For instance we may be interested in finding the maximum area of a region to be enclosed by a fixed perimeter; the minimum distance from a fixed point to a curve. In economics a function may represent a profit or cost function and we may want to find the value of x to find maximum profit and minimum cost, and so on. The Maximum – Minimum Theorem and the first and second derivative test will be crucial in finding such points as are illustrated in the following examples.

Example 4.68: A landowner wishes to use 2000 meters of fencing to enclose a rectangular region. Suppose one side of the land lies along a river and does not need fencing. What should be the sides of the region in order to maximize the area?

Solution: Suppose the rectangle is to have length x and width y meters as seen in Figure 4.13.

Since the length of the fencing is 2000 meters, we have

$$x + 2y = 2000$$

$\Rightarrow \quad 2y = 2000 - x \;\Rightarrow\; y = 1000 - x/2$

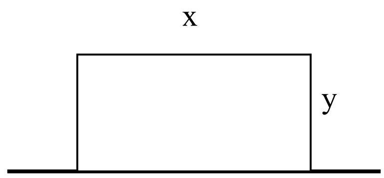


Figure 4.13

The area of the rectangle is A = xy which can be written as a function of x alone as

$$A(x) = xy = x(1000 - x/2) = 1000x - \frac{x^2}{2} \qquad \text{for } 0 \le x \le 2000$$

Thus we find the maximum value of A on [0, 2000].

$$A'(x) = 1000 - x = 0 \;\Rightarrow\; x = 1000 \text{ is a critical point.}$$

Comparing the value of A at the critical point and at the endpoints 0 and 2000, we get

$$A(0) = 0, \quad A(1000) = 500{,}000 \text{ and } A(2000) = 0 \quad \text{(check!)}$$

Thus the maximum value of A occurs when x = 1000 so that

$$y = 1000 - x/2 = 1000 - 500 = 500.$$

Consequently, to enclose maximum area, the fence should have a length of 1000 mts and a width of 500 mts.

Example 4.69: Ethiopian Airlines offers a round trip discount on group flight from Addis Ababa to Lalibela. If x people sign up for the flight, the cost of each ticket is to be 1000 – 2x Birr. Find the number of people the airline gets maximum revenue from the sales of tickets for the flight,

Solution: Since individual cost of a ticket is 1000 – 2x, the total cost of the group will be

$$C(x) = (1000 - 2x)x = 1000x - 2x^2.$$

To find a critical point, we solve $C'(x) = 1000 - 4x = 0$, which gives the only critical point x = 250 of C(x).

You can easily check that for x < 250, C′(x) > 0 and for x > 250, C′(x) < 0. Thus by the First Derivative Test C has an (absolute) maximum value at x = 250.

The maximum revenue the airline gets from the sales of 250 tickets is then

$$C(250) = 1000(250) - 2(250)^2 = 125{,}000 \text{ Birr}.$$

Example 4.70: A manufacturer wishes to produce rectangular containers with square bottom and top each of which is to have a capacity of 1000 cubic inches. If the cost of production of each container is proportional to its surface area, what should be the dimensions so as to minimize the cost of production?

Solution: Let x be the side of the base and h be the height of the container as seen in Figure 4.15.

Then the volume is

$$V = x^2h = 1000$$

$$\Rightarrow \quad h = \frac{V}{x^2} = \frac{1000}{x^2} \quad \text{for } x > 0$$

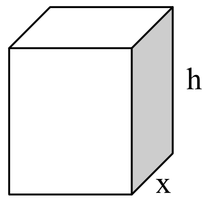


Figure 4.15

To find the surface area, we have the area of the top and bottom as $2x^2$ and the area of the four sides as

$$4xh = 4x\left(\frac{1000}{x^2}\right) = \frac{4000}{x}$$

Hence the total surface area is given by

$$s(x) = 2x^2 + \frac{4000}{x} \qquad \text{for } x > 0.$$

Since the cost of production is proportional to the surface area, to minimize cost, we find the minimum value of s.

$$s'(x) = 4x - \frac{4000}{x^2} = \frac{4x^3 - 4000}{x^2} = 0$$

$$\Rightarrow 4x^3 - 4000 = 0 \qquad \Rightarrow \qquad x^3 = 1000$$

$\Rightarrow$ x = 10 is the only critical point.

By the Second Derivative Test, we have

$$s''(x) = 4 + \frac{8000}{x^3} \quad \text{with} \quad s''(10) = 4 + 8 = 12 > 0$$

Thus x = 10 gives the minimum value s(10) = 600 sq. in.

The height is $h = \frac{1000}{x^2} = \frac{1000}{100} = 10$ in.

Hence the manufacturer would minimize the cost of production by manufacturing cubes of side 10 inches.

Curve Sketching

As a second application of the derivative we shall see here sketching the graphs of functions. You have been sketching the graphs of polynomial and rational functions starting from your high school mathematics. Here we systematically apply the notions of differential calculus to give precise meaning to the asymptotes, intervals of increase and decrease, the turning points and find the range of the functions.

First we shall list the important items that will help us in sketching the graph of a function y = f(x).

1) Determine the domain of the function f.
2) Find the intercepts of the function f.
    - x-intercepts are points of the form (x, 0)
    - y-intercepts are points of the form (0, y)
3) Determine the asymptotes, if any, of the function f.
    - A line x = c is a vertical asymptote of the graph of f iff
      $\lim_{x\to c^+} f(x) = \pm\infty$ or $\lim_{x\to c^-} f(x) = \pm\infty$.
    - A line y = L is a horizontal asymptote of the graph of f iff
      $\lim_{x\to\infty} f(x) = L$ or $\lim_{x\to-\infty} f(x) = L$.
    - A line y = ax + b is an oblique (or skew) asymptote of the graph of f iff
      $\lim_{x\to\pm\infty} [f(x) - (ax + b)] = L$
4) Determine the intervals of monotonicity of the function f.
    - f is increasing for all x at which $f'(x) > 0$
    - f is decreasing for all x at which $f'(x) < 0$
5) Find extreme values of f, if any.
   Find the critical points of f and apply the first or second derivative tests to determine whether they are relative extreme points or not.
6) If necessary plot some additional points to help you see the behavior of the function.

Example 4.71: Sketch the graph of $f(x) = \frac{x-2}{x+2}$.

Solution. 2.1 The domain of f is **R**\{-2} and the x-intercept is the value of x for which

$$f(x) = \frac{x-2}{x+2} = 0 \Rightarrow x = 2. \text{ Hence x-intercept at } (2, 0)$$

The y-intercept is the value of y when x = 0, i.e. $f(0) = \frac{0-2}{0+2} = -1$. Hence y-intercept at (0, -1).

Since $\lim_{x\to-2^+} f(x) = \lim_{x\to-2^+} \frac{x-2}{x+2} = -\infty$, the line x = -2 is a vertical asymptote to the graph of f.

Also you can check that $\lim_{x\to-2^-} f(x) = \infty$

Since $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{x-2}{x+2} = \lim_{x\to\infty} \frac{x-2/x}{x+2/x} = 1$, the line y = 1 is a horizontal asymptote for the graph of f.

To find the intervals of monotonicity, let us first find $f'(x)$.

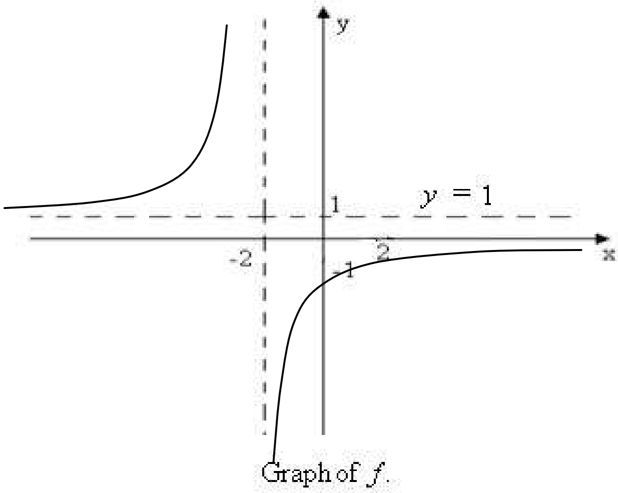
By the Quotient Rule for Differentiation,

$$f'(x) = \left(\frac{x-2}{x+2}\right)' = \frac{(x+2)(x-2)' - (x-2)(x+2)'}{(x+2)^2} = \frac{x+2-(x-2)}{(x+2)^2} = \frac{4}{(x+2)^2}.$$

Hence $f'(x) > 0$ for every element x in the domain of f. It follows that f is strictly increasing on $(-\infty, -2)$ and on $(-2, \infty)$.

f has no critical number and hence no local extrema.

Additional points: f(-1) = -3, f(1) = -1/3



The graph of f is given in Figure 4.16. Figure 4.16

Example 4.72: Sketch the graph of $f(x) = x + \frac{1}{x}$, for $x \neq 0$

Solution: Since $f(x) = x + \frac{1}{x} = \frac{x^2+1}{x} \neq 0$ and since $x \neq 0$ f has no intercepts.

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+}\left(x + \frac{1}{x}\right) = \infty \text{ and } \lim_{x\to 0^-} f(x) = \lim_{x\to 0^-}\left(x + \frac{1}{x}\right) = -\infty$$

The line x = 0 (the y-axis) is a vertical asymptote of f.

$$\lim_{x\to\infty}[f(x) - x] = \lim_{x\to\infty}\frac{1}{x} = 0$$

Hence the line y = x is an oblique asymptote of the graph of f.

$f'(x) = 1 - \frac{1}{x^2} = \frac{x^2-1}{x^2} = \frac{(x-1)(x+1)}{x^2} = 0$ gives two critical points x = 1 and x = -1.

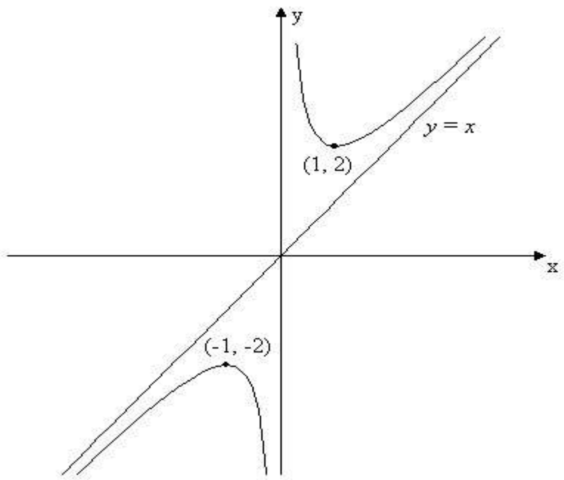
Using a sign chart to find the intervals of monotonicity:

	-1	1	
x – 1	- - - - - - - -	- - - - - - - - 0	+ + + + + + ++ + +
x + 1	- - - - - - - 0	+ ++ + + + + + + +	+ + + + + +
$f'(x)$	+ + + + + 0	- - - - - - - 0	+ + + + + + + + +

$f'(x) > 0$ in the interval $(-\infty, -1) \cup (1, \infty)$ so that it is strictly increasing in $(-\infty, -1) \cup (1, \infty)$.

$f'(x) < 0$ in the interval (-1, 1) \ {0} so that f is strictly decreasing in (-1, 1) \{0}.
Using the first derivative test, you can see that f(-1) = -2 is a local maximum and f(1) = 2 is a local minimum. You can also apply the second derivative test to see this.

Additional points: f(-2) = $-5/2$, $f(-1/2) = \frac{-5}{2}$, $f(1/2) = 5/2$, f(2) = $\frac{5}{2}$.



Graph of $f(x) = x + \frac{1}{x}$

The graph is given in Figure 4.17. Figure 4.17

Related Rates

One of the most important applications of the derivative is to solve problems involving rates of change. As was mentioned at the beginning of this section the derivative measures the rate of change of a variable quantity (which is the independent variable x) with respect to another variable (which is the dependent variable y = f(x)). Here we shall apply this to solve some practical related rates problems.

Example 4.73: Suppose a particle P starts from a point 0 and moves along a straight line in the positive direction as See in Figure 4.18

Figure 4.18

Let s(t) devote the distance traveled from 0 in t seconds. If we assume that the speed is constant, then we can compute the speed as

$$\text{speed} = \frac{\textit{distance traveled}}{\textit{time elapsed}}$$

If we are interested to find the average speed of the particle between two times t_1 and t_2 (with $t_1 < t_2$), we get

$$\text{Average speed} = \frac{\textit{change in distance}}{\textit{change in time}} = \frac{s(t_2) - s(t_1)}{t_2 - t_1}$$

In particular if t_1 is any time t and t_2 is a short time later say $t_2 = t + h$ for $h > 0$, then we have

Approximate speed (at t = t_1) = $\frac{s(t+h)-s(t)}{t+h-t} = \frac{s(t+h)-s(t)}{h}$

If the speed is not even constant, by taking h smaller and smaller we can approximate the speed of the particle at time t, to get what is called the (**instantaneous**) **velocity** of the particle as

$$v(t) = \lim_{h\to 0} \frac{s(t+h)-s(t)}{h}$$

Thus, if s(t) denotes the position function of the particle its velocity is given by

$$v(t) = s'(t) = \frac{ds}{dt} \text{ - rate of change of position.}$$

Similarly, the **acceleration** of the particle can be obtained by

$$a(t) = v'(t) = \frac{dv}{dt} \text{ - rate of change of velocity}$$

$$= \lim_{h\to 0} \frac{v(t+h)-v(t)}{h} = s''(t) = \frac{d^2s}{dt^2}$$

For instance if $s(t) = t^3 - 6t^2 + 20$ for $0 \le t \le 6$,

then $v(t) = s'(t) = 3t^2 - 12t$

and $a(t) = v'(t) = s''(t) = 6t - 12$

In general, if any quantity q is a function of time t, then the rate of change of the quantity with respect to time is given by the derivative q′(t).

Example 4.74 : Water is flowing into a vertical cylindrical tank of radius 2 feet at the rate of 8 ft^3/min. How fast is the water level rising after t minutes?

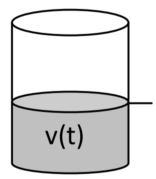


Figure 4.19

Solution: Let v(t) denote the volume of water in the tank after t minutes and let h(t) denote the height of water in the tank after t minutes. See Figure 4.19.

Since the rate at which water is flowing into the tank is 8 ft^3/min. the volume of water in the tank after t minutes is

$$v(t) = 8t$$

On the other hand since the base of the cylinder is 2 feet and height in h minutes is h(t), we have the volume

$$v(t) = \pi r^2 h(t) = 4\pi h(t)$$

Thus $4\pi h(t) = 8t \quad \Rightarrow h(t) = \frac{2}{\pi}t$

The rate at which the water level is rising in then

$$h'(t) = \frac{2}{\pi} \text{ ft|min,} \qquad \text{a constant!} \qquad \text{(why?)}$$

Example 4.75: Two automobiles start from a point A at the same time. One travels west at 60 km/hr and the other travels north at 35 km/hr. How fast is the distance between them increasing 3 hrs later?

Solution: Let s(t) denote the distance between the two cars after t hrs. In t hrs the car due north travels 35t kms and the car due west travels 60t kms as seen in Figure 4.20

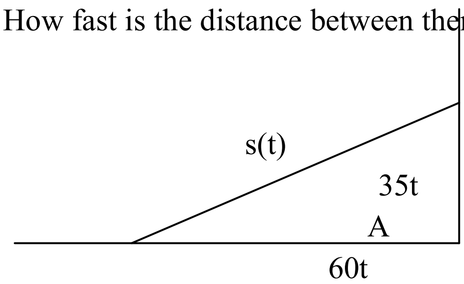


Figure 4.20.

Hence the distance s(t) between the two cars in t hrs is

$$s(t) = \sqrt{(35t)^2 + (60t)^2}$$

The rate of change of the distance between the cars is

$$s'(t) = \frac{2(35)^2 t + 2(60)^2 t}{2\sqrt{(35t)^2 + (60t)^2}} \qquad \ldots \text{ How ?}$$

Hence after 3 hrs the distance between the two cars is increasing at the rate of

$$s'(3) = \frac{3(35)^2 + 3(60)^2}{3\sqrt{(35)^2 + (60)^2}} = 5\sqrt{193} \text{ km/hr}$$

Exercise 4.3

1. Find relative extrema and the intervals in which the given function is increasing or decreasing

   a) $f(x) = 5 - 4x - x^2$ b) $g(x) = x^3 + x^2 - x - 4$

   c) $f(x) = \frac{x}{x^2+1}$ d) $g(x) = x^2 + \frac{1}{x^2}$

2. Use the First or Second Derivative Test to determine relative extreme values of the function

   a) $f(x) = 5x^2 - 2x + 1$ b) $g(x) = \frac{x^2}{4} + \frac{4}{x}$

   c) $f(x) = x^4 + \frac{1}{2}x$ d) $g(x) = \frac{1}{x^2+1}$

   e) $f(x) = \frac{\cos x}{1+\sin x}$ f) $g(x) = (x^2 + 2)^6$

3. Sketch the graph of the following functions

   a) $f(x) = (x^2 - 1)^2$ b) $g(x) = \frac{e^x}{x}$ c) $g(x) = \frac{x^3 - 3x^2 + 4}{x^2 - 1}$

4. A menu of total area of 100 sq. in. is printed with 2 in. margins at the top and bottom and 1in. margins at the sides. For what dimensions of the menu is the printed area largest?
5. A rectangle of perimeter p is rotated about one of its sides so as to from a cylinder. Of all such possible rectangles, which generated a cylinder of maximum volume?
6. The volume of a spherical balloon is increasing at a constant rate of 8 cubic feet per minute. How fast is the radius of the sphere increasing when the radius is exactly 10 feet?
7. At midnight ship B was 90 miles due south of ship A. Ship A sailed east at 15 m/hr and ship B sailed north at 20 m/hr. At what time were they closest to each other?

4.4. Integrals and their applications

In this section we shall introduce the second major part of calculus known as integral calculus. Just like subtraction is the inverse process of addition, integration is the inverse process of taking the derivative of a function. Historically, integral calculus was developed in solving problems connected with finding areas of regions with curved boundaries.

Section Objectives

At the end of this section you should be able to:

- define an anti-derivative of a continuous function.
- state properties of anti-derivatives.
- find indefinite integrals of some elementary functions.
- evaluate the integrals of functions using the techniques of integration.
- solve integrals involving trigonometric functions.
- find the definite integral of continuous functions.
- apply the concepts of definite integrals to find areas of regions bounded by continuous functions.

The Indefinite Integral

As is mentioned above the process of integration is the inverse process of differentiation and hence is sometimes called taking anti-derivatives.

Definition 4.13: A function F(x) is called an **anti-derivative** of a continuous function f(x) if and only if $F'(x) = f(x)$ for every x in the domain of f.

Example 4.76: Let $f(x) = 3x^2 + 4x$. Then the function $F_1(x) = x^3 + 2x^2$ is an anti-derivative of f(x), since $F'(x) = \frac{d}{dx}(x^3 + 2x^2) = 3x^2 + 4x = f(x)$.

Note that F_1 is not the only anti-derivative of f(x). You can also check that $F_2(x) = x^3 + 2x^2 + 5$ and $F_3(x) = x^3 + 2x^2 - 7$ are also anti-derivatives of f.

In fact, if c is any real number, then $F(x) = x^3 + 2x^2 + c$ is an anti-derivative of $f(x) = 3x^2 + 4x$ since $F'(x) = \frac{d}{dx}(x^3 + 2x^2 + c) = 3x^2 + 4x = f(x)$

Theorem 4.19: If F(x) is an anti-derivative of f(x), then F(x) + c, where c is an arbitrary constant, is also an anti-derivative of f(x).

Notation and terminologies: Given a function f, the symbol $\int f(x)dx$ stands for any (and hence all) anti-derivatives of f. i.e. if F(x) is an anti-derivative of f(x), we write $\int f(x)dx = F(x) + c$, for any constant c. The symbol $\int$ is called the **integral sign.** The function f(x) is called the **integrand**, x is called the **variable of integration**, and c is called a **constant of integration.** $\int f(x)dx$ is also called the **indefinite integral** of f with respect to x.

Examples 4.77: We have

a) $\int 3x^2dx = x^3 + c$ d) $\int \sin x dx = -\cos x + c$; g) $\int \cos x\, dx = \sin x + c$

b) $\int e^x dx = e^x + c$ e) $\int \frac{1}{x} dx = \ell n|x| + c$

c) $\int \sec^2 x dx = \tan x + c$ f) $\int \csc x \cot x dx = -\csc x + c$

- **Properties of the Indefinite Integral**

Suppose F and G are antiderivatives of f and g, respectively, and k is a constant. Then

1) $\int kf(x)dx = k\int f(x)dx = kF(x) + c.$

2) $\int (f(x) + g(x))dx = \int f(x)dx + \int g(x)dx = F(x) + G(x) + c.$

3) $\int (f(x) - g(x))dx = \int f(x)dx - \int g(x)dx = F(x) - G(x) + c.$

Examples 4.78:

1) $\int 4\cos x dx = 4\int \cos x dx = 4\sin x + c$

2) $\int \left(e^x - \frac{1}{x}\right) dx = \int e^x dx - \int \frac{1}{x} dx = e^x - \ell n|x| + c$

3) If $f(x) = x^r$, for any rational $r \neq -1$, then

$$\int f(x)dx = \int x^r dx = \frac{x^{r+1}}{r+1} + c \qquad \text{(verify!)}$$

Thus, $\int x^5 dx = \frac{x^6}{6} + c$ and $\int x^{-3/2}\, dx = \frac{x^{-1/2}}{-1/2} + c = \frac{-2}{\sqrt{x}} + c$.

4) $\int\left(\frac{1}{x^3} + 2\sec^2 x\right)dx = \int x^{-3} dx + 2\int \sec^2 x dx$

$= -\frac{1}{2x^2} + 2\tan x + c.$

5) If $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2x^2 + a_1x + a_0$ is a polynomial, then its anti-derivative is given by

$$P(x) = \int P(x)dx = \frac{a_n x^{n+1}}{n+1} + \frac{a_{n-1}x^n}{n} + \ldots + \frac{a_2x^3}{3} + \frac{a_1x^2}{2} + a_0x$$

Thus, $\int (3x^4 + \sqrt{2}\, x^3 - 5x + 2)dx = 3\int x^4 dx + \sqrt{2}\int x^3 dx - 5\int x dx + \int 2dx$

$= \frac{3}{5}x^5 + \frac{\sqrt{2}}{4}x^4 - \frac{5}{2}x^2 + 2x + c$

- **Some Techniques of Integration**

In the previous section we were trying to find anti-derivatives of some functions whose derivatives can easily be found from the previous unit on differentiation. But there are various functions such as

$f(x) = (x + 3)^5$, $g(x) = xe^{-x}$ and $h(x) = \frac{2x}{x(x^2-4)}$

whose anti-derivatives are not readily found. In this section we shall see some techniques to find the integrals of such functions.

a) Integration by Substitution

This technique is basically developed by reversing the Chain Rule. It is very helpful in finding the integrals of functions that appear as the composite of two functions.
Suppose we want to find the indefinite integral

$$\int (x+3)^5 dx$$

we may expand $(x + 3)^5$ and then integrate term by term using the formula

$\int x^r dx = \frac{1}{r+1}x^{r+1} + c$.

But this would obviously be very tedious and cumbersome. On the other hand if we replace or substitute u for x + 3, we get

$(x + 3)^5 = u^5$ and $\frac{du}{dx} = \frac{d}{dx}(x + 3) = 1 \Rightarrow dx = du.$

Thus, $\int (x+3)^5 dx = \int u^5 du = \frac{1}{6}u^6 + c$. Hence, $\int (x+3)^5 dx = \frac{1}{6}(x+3)^6 + c$, for some constant c.

Theorem 4.20: If $g'(x)$ is continuous $\forall x \in [a, b]$ and f is continuous at g(x), then

$\int f(g(x))g'(x)dx = \int f(u)du$ - **Integration by Substitution**

Example 4.79: Evaluate $\int 2x(x^2 - 5)^6 dx$

Solution: Let $u = x^2 - 5$. Then, $\frac{du}{dx} = 2x$ which implies that $du = 2xdx$. Thus,

$\int 2x(x^2-5)^6 dx = \int u^6 du = \frac{1}{7}u^7 + c$ and hence $\int 2x(x^2-5)^6 dx = \frac{1}{7}(x^2-5)^7 + c$.

Example 4.80: Integrate $\int \frac{x}{\sqrt{1+x^2}}dx$

Solution: Let $u = 1 + x^2$. Then, $\frac{du}{dx} = 2x$ which implies that $xdx = \frac{1}{2}du$. Therefore,

$$\int \frac{x}{\sqrt{1+x^2}}dx = \int \frac{1}{\sqrt{u}}\frac{du}{2} = \frac{1}{2}\int u^{-1/2}du = \frac{1}{2}.2.u^{1/2} + c = \sqrt{u} + c = \sqrt{1+x^2} + c.$$

Example 4.81: Integrate $\int \sin 4x dx$

Solution: Let $u = 4x$. Then $\frac{du}{dx} = 4 \Rightarrow dx = \frac{1}{4}du$. Thus $\int \sin 4x dx = \frac{1}{4}\int \sin u du = -\frac{1}{4}\cos u + c$

$= -\frac{1}{4}\cos 4x + c$.

In general $\int \sin ax dx = -\frac{1}{a}\cos ax + c$, and $\int \cos bx dx = \frac{1}{b}\sin bx + c$

These two formulas can be used to find integrals involving trigonometric functions together with trigonometric identities.

Example 4.82: Integrate $\int \sin^2 x \cos^2 x dx$

Solution: From trigonometric identities we have

$\sin^2 x = \dfrac{1-\cos 2x}{2}$ and $\cos^2 x = \dfrac{1+\cos 2x}{2}$

Thus $\int \sin^2 x\cos^2 x dx = \int \frac{1}{2}(1-\cos 2x)\frac{1}{2}(1+\cos 2x)dx$

$$= \frac{1}{4}\int (1-\cos^2 2x)dx = \frac{1}{4}\int\left[1-\frac{1}{2}(1+\cos 4x)\right]dx$$

$$= \frac{1}{8}\int dx - \frac{1}{8}\int \cos 4x dx = \frac{1}{8}x - \frac{1}{32}\sin 4x + c.$$

Example 4.83: Find $\int \tan x dx$

Solution: $\int \tan x dx = \int \frac{\sin x}{\cos x}dx$. Let u = cos x. Then $\frac{du}{dx}$ = -sinx ⇒ sinx dx = -du.

Hence, $\int \tan x dx = \int \frac{-du}{u}$ = $-\int \frac{1}{u}du$ = -ℓn|u| + c = -ℓn|cosx| + c.

You can similarly find $\int \cot x dx$.

Example 4.84: Integrate $\int e^{-2x}dx$

Solution: Let u = - 2x. Then $\frac{du}{dx}$ = -2 ⇒ dx = $-\frac{1}{2}$ du.

So that $\int e^{-2x}dx = -\frac{1}{2}\int e^u du = \frac{1}{2}e^u + c = \frac{1}{2}e^{-2x} + c$

In general since $\frac{d}{dx}e^{f(x)} = f'(x)\, e^{f(x)}$, *we have* $\int f'(x)e^{f(x)}dx = e^{f(x)} + c$.

Thus $\int xe^{x^2}dx = \frac{1}{2}e^{x^2} + c$ and $\int 3x^2 e^{x^3 dx} = e^{x^2} + c$.

b) Integration by Parts

The method of integration by parts is basically developed from the Product Rule for differentiation. If f and g are differentiable functions, we have

$$(f(x)g(x))' = f'(x)\, g(x) + g'(x)\, f(x)$$

Integrating on both sides with respect to x, we get

$$f(x)\, g(x) = \int f'(x)g(x)dx + \int g'(x)f(x)dx$$

If one of the integrals on the right can be easily evaluated, we can find the other integral using the following theorem

Theorem 4.21: If f and g are differentiable functions, then

$$\int f(x)g'(x)dx = f(x)g(x) - \int g(x)f'(x)dx$$ **- Integration by parts**

Example 4.85: Find $\int xe^x dx$

Solution: Let f(x) = x and g′(x) = e^x. Then $f'(x) = 1$ and g(x) = e^x. Therefore,

$$\int xe^x dx = xe^x - \int e^x.1.dx = xe^x - e^x + c, \text{ for some constant c.}$$

Integration by parts can be easily remembered using the following substitutions.

Let u = f(x) and v = g(x)

Then du = $f'(x)$dx and dv = g′(x)dx

So that $\int f(x)g'(x)dx = \int udv$ = f(x)g(x) - $\int g(x)f'(x)dx$ = uv- $\int vdu$

Thus, $\int udv = uv - \int vdu$ **- Integration by Parts**.

Example 4.86: Find $\int x\ell nxdx$

Solution: Let u = ℓnx, dv = xdx. Then, du = $\frac{1}{x}$dx, v = $\frac{x^2}{2}$.

Thus $\int x\ell nxdx = \frac{x^2}{2}\ell nx - \int \frac{x^2}{2}\frac{1}{x}dx = \frac{x^2}{2}\ell nx - \int \frac{x}{2}dx = \frac{x^2}{2}\ell nx - \frac{x^2}{4} + c.$

Example 4.87: Find $\int \ell nxdx$

Solution: Let u = ℓnx and dv = dx. Then, du = $\frac{1}{x}$dx and v = x

Hence $\int \ell nxdx = x\ell nx - \int x.\frac{1}{x}dx = x\ \ell nx - \int dx = x\ell nx - x + c = x(\ell nx - 1) + c.$

Example 4.88: Find the integral $\int x^2 e^x dx$

Solution: Let u = x^2 and dv = e^x dx. Then, du = 2xdx, v = e^x and $\int x^2 e^x dx = x^2e^x - \int 2xe^x dx$.

But we have seen above that $\int xe^x dx = xe^2 - e^x + c$. Hence,

$\int x^2 e^x dx = x^2e^x - 2(xe^x - e^x + c) = e^x(x^2 - 2x + 1) + c_1$ where c_1 = -2c is a constant.

In some cases we may have to apply integration by parts more than once to arrive at the required result as in the following example.

Example 4.89: Find $\int e^x \cos xdx$

Solution: Let u = e^x and dv = cosxdx. Then, du = e^xdx and v = sinx. Thus,

$$\int e^x \cos xdx = e^x\sin x - \int e^x \sin xdx.$$

To evaluate the integral on the right, we again use integration by parts.

Let u = e^x and dv = sinx dx. Then, du = e^xdx and v = -cosx.

Thus, $\int e^x \sin x dx$ = -e^xcosx + $\int e^x \cos x dx$ which implies 2$\int e^x \cos x dx$ = e^xsinx + e^xcosx + c.

Therefore, $\int e^x \cos x dx = \frac{1}{2} e^x$ (sinx + cosx) + c.

c) Integration by the Method of Partial Fractions

The method of Partial Fractions is used for rational functions

$$f(x) = \frac{p(x)}{q(x)}$$

where degree of p(x) is less than degree of q(x). (If not we can apply long division to write f(x) as a sum of a polynomial and a rational function with the desired property.) The first step in this method is to factorize the denominator q(x) into linear factors, if possible. (The case where we have irreducible quadratic factors of q(x) will not be treated here.) Now with each linear factor $(ax + b)^m$ (of multiplicity m) we associate constants A_1, A_2, …,A_m and write $\frac{A_1}{ax+b} + \frac{A_2}{(ax+b)^2}$+…with the A_i's to be determined. Then, the rational function f(x) is then expressed as a sum of simple rational functions and can be easily integrated.

Example 4.90: Find $\int \frac{1}{x^2-4} dx$

Solution: By factorizing $x^2 - 4$ as (x – 2) (x + 2), we have

$$\frac{1}{x^2-4} = \frac{1}{(x-2)(x+2)} = \frac{A}{x-2} + \frac{B}{x+2} = \frac{A(x+2)+B(x-2)}{(x-2)(x+2)}$$

Since the denominators are equal, we equate the numerators as A(x+2) + B(x - 2) = 1.
From equality of polynomials, we get

$$\left.\begin{array}{r} A+B=0 \\ 2A-2B=1 \end{array}\right\} \Rightarrow A = 1/4 \text{ and } B = \frac{-1}{4}$$

Hence, $\int \frac{1}{x^2-4} dx = \int \left(\frac{1/4}{x-2} + \frac{-1/4}{x+2} \right) dx = \frac{1}{4} \int \frac{dx}{x-2} - \frac{1}{4} \int \frac{dx}{x+2} = \frac{1}{4} \ell n|x-2| - \frac{1}{4} \ell n|x+2| + c.$

Example 4.91: Find $\int \frac{3x^2+x-1}{x^3-x^2-2x} dx$

Solution: The denominator $x^3 - x^2 - 2x = x(x^2 - x - 2) = x(x + 1)(x - 2)$ has three roots 0, -1 and 2.

$$\frac{3x^2+x-1}{x^3-x^2-2x} = \frac{3x^2+x-1}{x(x+1)(x-2)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-2}$$

$$= \frac{A(x-1)(x-2)+Bx(x-2)+Cx(x-1)}{x(x+1)(x-2)}$$

$\Rightarrow$ A(x+1) (x - 2) + B x(x - 2) + Cx(x+1) = $3x^2 + x - 1$

This equation is true for all x∈**R**. In particular,

when x = 0, A(1)(-2) = -1 $\Rightarrow$ A = 1/2

when x = -1, B(-1)(-3) = 1 $\Rightarrow$ B = 1/3

when x = 2, C(2)(3) = 13 $\Rightarrow$ C = 13/6

Hence, $\int \frac{3x^2 + x - 1}{x^3 - x^2 - 2x}dx = \frac{1}{2}\int \frac{dx}{x} + \frac{1}{3}\int \frac{dx}{x+1} + \frac{13}{6}\int \frac{dx}{x-2} = \frac{1}{2}\ell n|x| + \frac{1}{3}\ell n|x+1| + \frac{13}{6}\ell n|x-2| + c.$

- **The Definite Integral**

For a very long time, mathematicians have struggled with the problem of finding areas of plane regions. Until the invention of the integral calculus, however, the regions considered were mostly those regions bounded by straight lines, called polygons, with a few exceptions such as the circle and the ellipse. The Greek mathematicians found the area of a polygon by first finding the area of a rectangle, then finding the area a parallelogram, and then finding the area of a triangle. The area of a polygon can be used to approximate the area of a region bounded by curved boundaries. For instance, the area of a circle can be found by drawing a sequence of inscribed polygons
$P_4, P_8, P_{16}, \ldots, P_n$, and then taking limit as $n \to \infty$.

To develop the idea for more general regions, consider the region bounded by the graphs of

$$y = 2x^2 + 1, \quad x = 0, \quad x = 6 \quad \text{and} \quad \text{x-axis}.$$

To find the area of the region, let us identify the region S by drawing its boundaries, namely the graphs of $y = f(x) = 2x^2 + 1, x = 0, x = 6$ and the x-axis as shown in Figure 4.21.

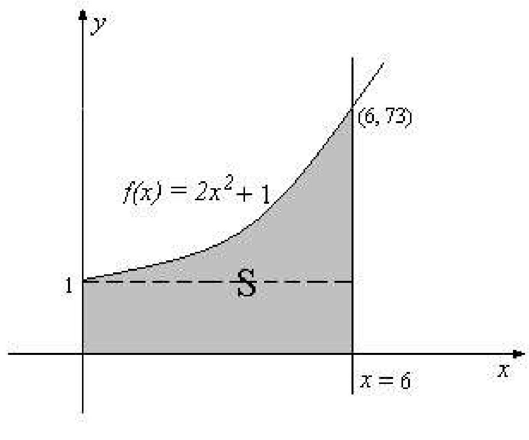


Figure 4.21

Unfortunately, since $f(x) = 2x^2 + 1$ is a curve that is not a line segment, we cannot find the area of the region by the elementary methods. So, it is necessary to develop a stronger technique that also generalizes the elementary method and enables us to find the area of such regions.

Let A(S) denote the area of the region S. It is not difficult to give lower and upper bounds of A(S). For instance, we consider the rectangle r that is enclosed by the boundaries of S and the rectangle R that encloses S, as shown in Figure 4.22.

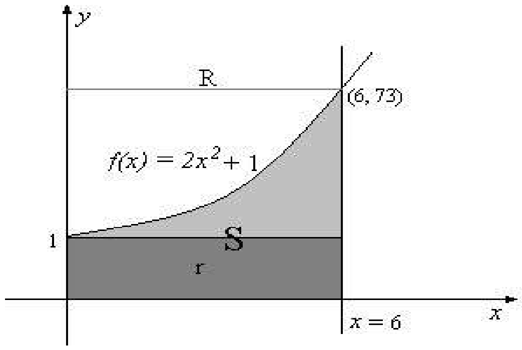


Figure 4.22

Then A(r) = 6 and A(R) = 6×73 = 438. Hence $6 \leq$ A(S) ≤ 438, which gives a wide range of bounds of A(S).

Better bounds of A(S) can be obtained if we consider the finer rectangles r_1, r_2, r_3, r_4, r_5 and r_6 that are enclosed by the boundaries of S and R_1, R_2, R_3, R_4, R_5 and R_6 that enclose S as shown in Figure 4.23.

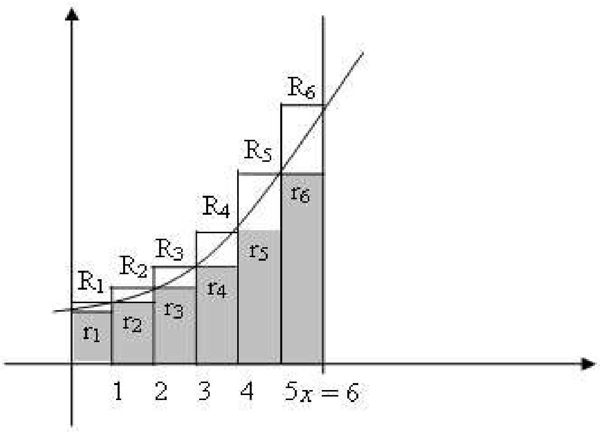


Figure 4.23

Evidently, each of the rectangles has base 1 unit but varying heights. It follows that

$$A(r_1)+A(r_2) + \ldots + A(r_6) \leq A(S) \leq A(R_1) + A(R_2) + \ldots + A(R_6)$$

i.e., $\sum_{i=1}^{6} A(r_i) \leq A(S) \leq \sum_{i=1}^{6} A(R_i)$ which gives $116 \leq A(S) \leq 188$.

To give a formal definition of the subdivisions, for any positive integer n, divide [a, b] into subintervals by introducing points of subdivision $x_0, x_1, \ldots x_n$

Definition 4.14: A **partition** of [a, b] is a finite set P of points $x_0, x_1, \ldots, x_n$ such that $a = x_0 < x_1 < x_2 < \ldots < x_n = b$. We describe P by writing $P = \{ x_0, x_1, \ldots x_n\}$

By definition, any partition of [a, b] must contain a and b.

The length of any subinterval $[x_{i-1}, x_i]$ of a partition P is defined and given by

$$\Delta x_i = x_i - x_{i-1}$$

In particular, when the lengths of each subintervals are equal, it is called a **regular** partition.

In this section we shall consider only regular partitions, so that the length of each subinterval is

$$\Delta x_i = x_i - x_{i-1} = \frac{b-a}{n}$$

Having chosen a partition P of [a, b], we inscribe and circumscribe rectangles on the region R using the division points of P as seen in Figure 7.3(a) and (b). Since f is continuous on [a, b], by the Maximum-Minimum Theorem, for each i between 1 and n, there is a minimum value m_i and a maximum value M_i of f on the subinterval $[x_{i-1}, xi]$. If r_i and R_i denote the inscribed and circumscribed rectangles on $[x_{i-1,}x_i]$, respectively, then the area of r_i is $A(r_i) = m_i\Delta x_i$ and the area of R_i is $A(R_i) = M_i\Delta x_i$, since the base of both r_i and R_i is $\Delta x_i = x_i - x_{i-1}$. From our observation in Figure 7.3 (a) and (b) we see that the area of the region R is between the sum of the inscribed rectangles and the sum of the circumscribed rectangles.

Definition 4.15: Let f be continuous on [a, b] and P be any partition of [a, b].

The sum $L_f(P) = m_1\Delta x_1 + m_2\Delta x_2 + \ldots + m_n\Delta x_n$

is called the **lower sum** of f associated with P and the sum

$$U_f(P) = M_1\Delta x_1 + M_2\Delta x_2 + \ldots + M_n\Delta x_n$$

is called the **upper sum** of f associated with P.

From our construction we see that if P is any partition of [a, b], then the area of R should be between $L_f(P)$ and $U_f(P)$ i.e.

$$\boxed{L_f(P) \le Area\ (R) \le U_f(P)}$$

Example 4.92: Let $f(x) = x^2$ for $o \le x \le 2$ and let $P = \left\{0, \frac{1}{2}, 1, \frac{3}{2}, 2\right\}$ be a partition of [0, 2].

Then the subdivision of [0, 2] associated with P are $\left[0, \frac{1}{2}\right]$, $\left[\frac{1}{2}, 1\right]$, $\left[1, \frac{3}{2}\right]$, $\left[\frac{3}{2}, 2\right]$. Since x^2 is an increasing function on [0, 2], the minimum value of f on each subinterval is at the left end point and the maximum value of f at the right end point. Thus

$$m_1 = f(0) = 0, \quad m_2 = f\left(\frac{1}{2}\right) = \frac{1}{4}, \quad m_3 = f(1) = 1, \quad m_4 = f\left(\frac{3}{2}\right) = \frac{9}{4}$$

and $M_1 = f\left(\frac{1}{2}\right) = \frac{1}{4}$, $M_2 = f(1) = 1$, $M_3 = f\left(\frac{3}{2}\right) = \frac{9}{4}$, $M_4 = f(2) = 4$

The base of each subinterval is $\Delta x_i = \frac{2-0}{4} = \frac{1}{2}$. Thus the lower sum of f associated with P is

$$L_f(P) = 0.\frac{1}{2} + \frac{1}{4}.\frac{1}{2} + 1.\frac{1}{2} + \frac{9}{4}.\frac{1}{2} = \frac{7}{4}$$

and the upper sum of f associated with P is

$$U_f(P) = \frac{1}{4}.\frac{1}{2} + 1.\frac{1}{2} + \frac{9}{4}.\frac{1}{2} + 4.\frac{1}{2} = \frac{15}{4}$$

Therefore the area of the region R below the graph of $f(x) = x^2$ above the x-axis on [0, 2] is between $\frac{7}{4}$ and $\frac{15}{4}$, i.e. $\frac{7}{4} \le \text{Area}(R) \le \frac{15}{4}$.

Definition 4.16: Let f be continuous on [a, b]. The **definite integral of f from a to b** is the unique number I satisfying $L_f(P) \le I \le U_f(P)$ for every partition P of [a, b].
This integral is denoted by

$$I = \int_a^b f(x)dx$$

The numbers a and b are called the **lower** and **upper limits of integration**, respectively.

Note that as the number of subdivisions of an interval [a, b] increases, the minimum and the maximum values of f on $[x_{i-1}, x_i]$ are close to each other. For each i from 1 to n if we take an arbitrary number t_i in $[x_{i-1}, x_i]$, then we get the sum

$$\sum_{i=1}^{n} f(t_i)\Delta x_i = f(t_1)\Delta x_1 + f(t_2)\Delta x_2 + \ldots + f(t_n)\Delta x_n$$

This sum is called a **Riemann sum** or an **Integral sum.**

Even though it is sometimes possible to calculate $\int_a^b f(x)dx$ by finding formulas for lower and upper sum we are to evaluate it here by the use of the Fundamental Theorem of Calculus.

For the moment we can conclude that if f is continuous and nonnegative on [a, b], then the area of the region R between the graph of f and the x-axis on [a, b] is given by

$$\text{Area}(R) = \int_a^b f(x)dx .$$

Remark: *The definite integral has the following properties.*

If f and g are integrable over [a, b] and k is a constant, then

a) $\int_a^b kf(x)dx = k\int_a^b f(x)dx$

b) $\int_a^b (f(x) \pm g(x))dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx$

c) *If* $f(x) \ge 0$, *for* $a \le x \le b$, *then* $\int_a^b f(x)dx \ge 0$ *and* *if f(x) ≤ 0, for a ≤ x ≤ b, then* $\int_a^b f(x)dx \le 0$.

d) *If m ≤ f(x) ≤ M for all x ∈ [a, b], then*

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a)$$

e) *If c is any number in (a, b), then*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \qquad \text{- Additive Property}$$

f) $\int_a^a f(x)dx = 0$ *for any number a.*

g) $\int_b^a f(x)dx = -\int_a^b f(x)dx$.

To develop a general method for evaluating $\int_a^b f(x)dx$ without computing lower and upper sums we shall state the most important theorem in calculus: **The Fundamental Theorem of Calculus**. To this end let f(t) be continuous on [a, b]. Then f is integrable on [a, b] and for any x∈[a, b] the definite integral $\int_a^x f(x)dt$ exists. Define a function F on [a, b] as F(x) = $\int_a^x f(t)dt$

In effect the Fundamental Theorem of Calculus states that the function F(x) is differentiable with derivative f(x) thereby eliminating the integral by the derivative. It also shows us how to evaluate the definite integral.

Theorem 4.22: (Fundamental Theorem of Calculus)
Let f(t) be continuous on [a, b] and for each x ∈[a, b] let

$$F(x) = \int_a^x f(t)dt$$

Then (i) F(x) is a differentiable function with F′(x) = f(x)

(ii) If F is any anti-derivative of f on [a, b], then $\int_a^b f(t)dt$ = F(b) – F(a).

Remarks: *a) From (ii) to evaluate* $\int_a^b f(x)dx$ *all we have to do is to find an anti-derivative of F of f and find the difference of its values at a and at b. This is usually denoted by* $[F(x)]_a^b$ *or* $F(x)\Big|_a^b$ *to mean F(b) – F(a).*

b) *If F is an anti-derivative of f, then F(x) + c, for any constant c is also an anti-derivative of f. But since* $\left[F(x)+c\right]_a^b$ *= (F(b) + c) – (F(a) + c) = F(b) – F(a) =* $F(x)\Big|_a^b$

the constant c does not play any role in evaluating the definite integral. Thus we can always take c = 0.

Example 4.93: Let f(x) = x^2 for $0 \le x \le 2$. Then F(x) = $\frac{1}{3}x^3$ is an anti-derivative of f, so that by the Fundamental Theorem of Calculus.

$$\int_0^2 x^2 dx = \frac{1}{3}x^3 \Big|_0^2 = F(2) - F(0) = \frac{1}{3}2^3 - \frac{1}{3}.0^3 = \frac{8}{3} - 0 = \frac{8}{3}$$

From our previous discussion, the area of the region R under the graph of f(x) = x^2 on [0, 2] above the x-axis is thus $\int_0^2 x^2 dx = 8/3$ sq. units.

Example 4.94: Evaluate each of the following definite integrals

a) $\int_1^4 3\sqrt{x}dx$ b) $\int_0^\pi \sin x dx$

c) $\int_0^{\pi/2} (x+\cos x)dx$ d) $\int_0^1 (5x^3+2x-e^x)dx$

Solution: a) Since F(x) = $3.\frac{2}{3}.x^{3/2} = 2x\sqrt{x}$ is an anti-derivative of f(x) = $3\sqrt{x}$, we have

$$\int_0^4 3\sqrt{x}dx = F(4) - F(1) = 2(4)\sqrt{4} - 2(1)\sqrt{1} = 16-2 = 14$$

b) An anti-derivative of sinx is –cosx. Thus

$$\int_0^\pi \sin x dx = -\cos x \Big|_0^\pi = -\cos\pi - (-\cos 0) = -(-1) + 1 = 2.$$

c) $$\int_0^{\pi/2} (x+\cos x)dx = \left[\frac{x^2}{2}+\sin x\right]_0^{\pi/2} = \left(\frac{\pi^2}{8}+1\right) - (0+0) = \frac{\pi^2}{8}+1.$$

d) $$\int_0^1 (5x^3+2x-e^x)dx = \left[\frac{5}{4}x^4+x^2-e^x\right]_0^1 = \left(\frac{5}{4}+1-e\right) - (0+0-1) = \frac{13}{4}-e$$

Remark: *For functions that are given by more than one formula we evaluate the definite integral using the additive property.*

Example 4.95: Evaluate $\int_{-2}^1 |x+1|dx$

Solution: By definition |x + 1| = $\begin{cases} x+1, & for\ x \geq -1 \\ -(x+1), & for\ x < -1 \end{cases}$

Then by Additive Property, we have

$$\int_{-2}^{1} |x+1| dx = \int_{-2}^{-1} -(x+1)dx + \int_{-1}^{1} (x+1)dx$$

$$= -\left[\frac{x^2}{2} + x\right]_{-2}^{-1} + \left[\frac{x^2}{2} + x\right]_{-1}^{1}$$

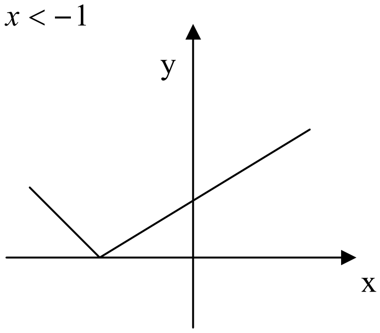


Figure 4.23

$$= -\left[\left(\frac{1}{2} - 1\right) - (2-2)\right] + \left[\left(\frac{1}{2} + 1\right) - \left(\frac{1}{2} - 1\right)\right] = \frac{1}{2} + 2 = 5/2$$

From the method of Integration by Substitution we have

$$\int f(g(x))g'(x)dx = \int f(u)du \qquad \text{where } u = g(x)$$

If we are to evaluate this integral between a and b, we have, when x = a, u = g(a) and when x = b, u = g(b). Thus it follows

$$\boxed{\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du}$$

- Change of Variable.

Example 4.96: Evaluate $\int_2^3 x\sqrt{x^2 - 4}dx$

Solution: We have two possibilities to evaluate such a definite integral. One way is to find an anti-derivative of $x\sqrt{x^2 - 4}$ and evaluate it between 2 and 3 by the Fundamental Theorem of Calculus. The other is to use the change of variable formula and change the limits of integration before integrating.

To this end, let u = g(x) = $x^2 - 4$. Then du = 2xdx.

When x = 2, u = g(2) = 0 and when x = 3, u = g(3) = 5

$$\text{Thus } \int_2^3 x\sqrt{x^2 - 4}dx = \int_0^3 \sqrt{u}\frac{du}{2} = \frac{1}{2}\int_0^5 u^{1/2}du$$

$$= \frac{1}{2}.\frac{2}{3}u\sqrt{u}\Big|_0^5 = \frac{1}{3}.5,\sqrt{5} \text{ - } 0 = \frac{5\sqrt{5}}{3}.$$

- ## Application of the Definite Integral: Area

The definite integral has several applications such as finding areas of regions, arc length of curves, surface areas and volumes of solids of revolution. In this section we shall see how to find areas of plane regions with curved boundaries using the definite integrals.

In the previous section we have seen that if f(x) ≥ 0 for all x ∈[a, b] and if f is continuous on [a, b], then $\int_a^b f(x)dx$ gives the area of the region R below the graph of f, above the x-axis, between the lines x = a and x = b. For instance, if f(x) = x^2 for 0 ≤ x ≤ 2, then the area of R as given in is given by A(R) = $\int_0^2 x^2 dx = \left.\frac{x^3}{3}\right|_0^2 = 8/3$ sq. units.

- If f(x) ≤ 0 on [a, b], then taking g(x) = -f(x) ≥ 0 for a ≤ x ≤ b, the area of the region R below the x-axis, above the graph of f on [a, b] is given by A(R) = $\int_a^b g(x)dx = \int_a^b -f(x)dx = -\int_a^b f(x)dx$.

For instance, if f(x) = 2x for -2 ≤ x ≤ 0, then the area of the region R below the x-axis, above the graph of f on [-2, 0] is given by

$$A(R) = -\int_{-2}^{0} f(x)dx = -\int_{-2}^{0} 2x dx = -x^2\Big|_{-2}^{0} = -[0 - 4] = 4 \text{ sq. units.}$$

Now let f and g be continuous on [a, b], and assume that f(x) ≥ g(x) for a ≤ x ≤ b. Then the area of the region R below the graph of f, above the graph of g, and between the lines x = a and x = b is given by

$$A(R) = \int_a^b [f(x) - g(x)]dx$$

Example 4.97: Find the area of the region bounded by $f(x) = 2\sqrt{x}$, $g(x) = -x$ and line $x = 9$.

Solution. Sketching the graphs of $y = f(x)$, $y = g(x)$ and $x = 9$, the region R can be identified as shown in Figure 4.24.

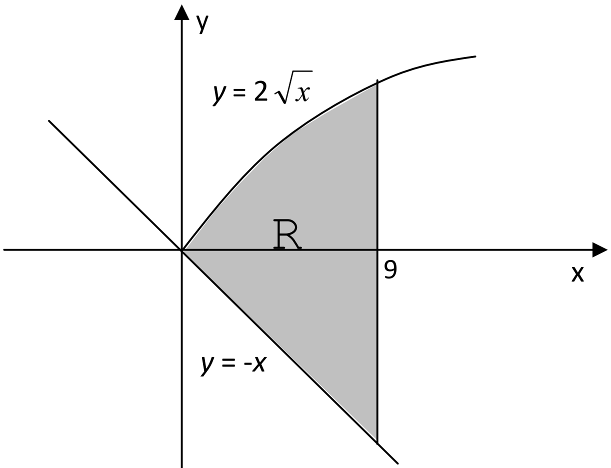


Figure 4.24

It follows that

$$A(R) = \int_0^9 (f(x) - g(x))dx = \int_0^9 \left(2\sqrt{x} + x\right)dx = \frac{4}{3}x^{\frac{3}{2}} + \frac{x^2}{2}\Big|_0^9 = 76.5.$$

Exercise 4.4

1. Evaluate the following indefinite integrals
   a) $\int (x^3 + 5)dx$ d) $\int (4 - x + 3x^2 - 2x^5)dx$
   b) $\int 2x^{-8}dx$ e) $\int (\cos x - 4e^x)dx$
   c) $\int 3 \sin x dx$ f) $\int \frac{x^3 - 4\sqrt{x}}{x^2}dx$
2. Find the following integrals by substitution
   a) $\int \frac{dx}{x-3}$ c) $\int \sin^2 x \cos x dx$
   b) $\int e^{x/3}dx$ d) $\int \frac{\ell n^4 x}{x}dx$
3. Find the following integrals by the method of Integration by Parts.
   a) $\int x \cos x dx$ c) $\int \frac{\ell n x}{x^2}dx$
   b) $\int (x+1)3^x dx$ d) $\int x^2 \sin x dx$
4. Integrate the following by the method of Partial Fractions
   a) $\int \frac{dx}{(x+2)(3x+4)}$ c) $\int \frac{dx}{(x-1)(x-2)(x-3)}$
   b) $\int \frac{x}{x^2 - x - 6}$ d) $\int \frac{2x}{(x-2)^2}dx$
5. Find the area of the region R between the graph of f and the x-axis on the given interval
   a) $f(x) = x^2 + 1$, on [1, 3]
   b) $f(x) = 2 + \cos x$, on $[0, 3\pi/2]$
   c) $f(x) = \frac{1}{x}$, on [1, 4]
   d) $f(x) = |x| - 1$, on [-1, 2]
6. Find the area of the region between the graphs of the following functions.
   a) $f(x) = x^2$ and $g(x) = 2 - x$
   b) $f(x) = e^x$, x = -1, x = 3 and the x-axis
   c) $f(x) = x^2 - 4$ and $g(x) = 4 - x^2$

References

1. Alemayehu Haile and Yismaw Alemu, *Mathematics an Introductory Course*, Department of Mathematics, Addis Ababa University
2. Demisu Gemeda and Seid Mohammed, Fundamental Cocepts of Algebra, AAU
3. Edwin J. Purcell, Dale Varberg, Calculus with Analytic Geometry
4. G. Chartrand, A. D. Polimeni and P. Zihang, Mathematical proofs: a transition to advanced mathematics 3rd edition, Pearson Education. Inc.
5. Goodman Hirsch, Precalculus-Understanding functions, 2000