MATHEMATICS-I

BCA 101

SELF LEARNING MATERIAL

DIRECTORATE
OF DISTANCE EDUCATION

SWAMI VIVEKANAND SUBHARTI UNIVERSITY

MEERUT – 250 005,

UTTAR PRADESH (INDIA)

SLM Module Developed By :

Author:

Reviewed by :

Assessed by:

Study Material Assessment Committee, as per the SVSU ordinance No. VI (2)

DISCLAIMER

Published by: Gayatri Sales

Typeset at: Micron Computers **Printed at:** Gayatri Sales, Meerut.

MATHEMATICS – I

Unit-I Set Theory: Sets and subsets, Finit and infinite sets, Algebra of sets: Union and Intersection, Complementation, Demorgan's law, Common application of algebra of sets.

Elementary Properties of Numbers: Mathematical Induction**,** Division Algorithm, The Greatest Common Divisor, The Euclidean Algorithm, The Diophantine Equation.

Unit-II Matrix: Matrix, Submatix, Types of matrices such as symmetric, skew symmetric, Hermitian, Skew Hermitian, Nilpotent, Involutary, Orthogonal etc., Singular and Non singular matrices, Addition and subtraction of matrices, Rank of matrices, Matrix Equation, Solution by Cramer's rule and Gauss Elimination method.

Unit-III Vectors: Vectors, Vector algebra, Addition and Subtraction of Vectors, Scalar and vector product of two vectors, Simple application of vectors.

Unit-IV Differentiation: Differentiation of Functions as polynomials, rationales, exponential, logarithmic and trigonometric function.

Unit-V Integration: Integration as inverse of differentiation, integration of simple Functions, integrationby parts, integration by substitution, definite integrals.

MATHEMATICS – I

Unit-I Set Theory:

Sets and subsets:-

Set - A collection of objects. The specific objects within the set are called the elements or members of the set. Capital letters are commonly used to name sets.

Examples: $Set\ A = \{a, b, c, d\}\ or\ Set\ B = \{1, 2, 3, 4\}$

Set Notation - Braces { } can be used to list the members of a set, with each member separated by a comma. This is called the “Roster Method.” A description can also be used in the braces. This is called “Set-builder” notation.

Example: Set A: The natural numbers from 1 to 10.
Members of A: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Set Notation: A = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
Set Builder Not.: $\{x|x\ is\ a\ natural\ number\ from\ 1\ to\ 10\}$

Ellipsis - Three dots (…) used within the braces to indicate that the list continues in the established pattern. This is helpful notation to use for long lists or infinite lists. If the dots come at the end of the list, they indicate that the list goes on indefinitely (i.e. an infinite set).

Examples: Set A: Lowercase letters of the English alphabet
Set Notation: $\{a, b, c, \dots, z\}$

Cardinality of a Set – The number of distinct elements in a set.

Example: Set A: The days of the week
Members of Set A: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday
Cardinality of Set A = $(\boldsymbol{A})$ = 7

Equal Sets– Two sets that contain exactly the same elements, regardless of the order listed or possible repetition of elements.

Example: A = {1, 1, 2, 3, 4} , B = {4, 3, 2, 1, 2, 3, 4,} .

Sets A and B are equal because they contain exactly the same elements (i.e. 1, 2, 3, & 4). This can be written as $\boldsymbol{A} = \boldsymbol{B}$.

Equivalent Sets – Two sets that contain the same number of distinct elements.

Example: A = $\{Football, Basketball, Baseball, Soccer\}$
B = $\{penny, nickel, dime, quarter\}$

(A) = 4 and (B) = 4
A and B $are\ Equivalent\ Sets$, $meaning$ (A) = (B).

Note: If two sets are equal, they are also Equivalent

Example: $Set\ A = \{a, b, c, d\}$ $Set\ B = \{d, d, c, c, b, b, a, a\}$

Are Sets A and B Equal?

Set A and B have exact same elements.

Are Sets A and B Equivalent?

Set A and B have the exact same number of distinct elements.. n(A)= n(B)= 4

The Empty Set (or Null Set)– The set that contains no elements.
It can be represented by either { } or ∅.
Note: Writing the empty set as {∅} is not correct!

Symbols commonly used with Sets –

$\in \rightarrow$ $indicates\ an\ object\ is\ an$ $\boldsymbol{element}$ $of\ a\ set.$
$\notin \rightarrow$ $indicates\ an\ object\ is$ $\boldsymbol{not}$ $an\ element\ of\ a\ set.$
$\subseteq \rightarrow$ $indicates\ a\ set\ is\ a$ $\boldsymbol{subset}$ $of\ another\ set.$
$\subset \rightarrow$ $indicates\ a\ set\ is\ a$ $\boldsymbol{proper\ subset}$ $of\ another\ set.$
$\cap \rightarrow$ $indicates\ the$ $\boldsymbol{intersection}$ $of\ sets.$
$\cup \rightarrow$ $indicates\ the$ $\boldsymbol{union}$ $of\ sets.$

Both Sets have 4 elements Note: If two sets are Equal, they are also Equivalent! Sets A and B have the exact same elements! $\{a, b, c, d\}$ Sets A and B have the exact same number of distinct elements! $(A) = (B) = 4$ 3

Subsets - For Sets A and B, Set A is a Subset of Set B if every element in Set A is also in Set B. It is written as $A \subseteq$ B.

Proper Subsets - For Sets A and B, Set A is a Proper Subset of Set B if every element in Set A is also in Set B, but Set A does not equal Set B. ($\boldsymbol{A \neq B}$) It is written as $\boldsymbol{A \subset B}$.

Example: $Set\ A = \{2, 4, 6\}$ $Set\ B = \{0, 2, 4, 6, 8\}$

$\{2, 4, 6\} \subseteq \{0, 2, 4, 6, 8\}$ and $\{2, 4, 6\} \subset \{0, 2, 4, 6, 8\}$

Set A is a Subset of Set B because every element in A is also in B. $A \subseteq B$

Set A is a Proper Subset of Set B because every element in A is also in B. $A \subset B$

Note: The Empty Set is a Subset of every Set.
The Empty Set is also a Proper Subset of every Set except the Empty Set.

Number of Subsets – The number of distinct subsets of a set containing n elements is given by .

Number of Proper Subsets – The number of distinct proper subsets of a set containing n elements is given by $\mathbf{2\ n - 1}$.

Example: How many Subsets and Proper Subsets does Set A have?

$Set\ A = \{bananas, oranges, strawberries\}$
$n = 3$

Subsets = $2\ n = 2_3 = 8$ Proper Subsets = $2\ n - 1 = 7$ Example

Example: List the Proper Subsets for the Example above.

1. $\{bananas\}$
2. $\{oranges\}$
3. $\{strawberries\}$
4. $\{bananas, oranges\}$
5. $\{bananas, strawberries\}$
6. $\{oranges, strawberries\}$
7. $\emptyset$

Intersection of Sets – The Intersection of Sets A and B is the set of elements that are in both A and B, i.e. what they have in common. It can be written as $\boldsymbol{A} \cap \boldsymbol{B}$.

Union of Sets – The Union of Sets A and B is the set of elements that are members of Set A, Set B, or both Sets. It can be written as $\boldsymbol{A} \cup \boldsymbol{B}$.

Example: Find the Intersection and the Union for the Sets A and B.

$Set\ A = \{Red,, Green\}$
$Set\ B = \{Yellow, Orange, Red, Purple, Green\}$
Intersection: $\boldsymbol{A} \cap \boldsymbol{B} = \{Red, Green\}$
Union: $\boldsymbol{A} \cup \boldsymbol{B} = \{Red, Blue, Green, Yellow, Orange, Purple\}$

Complement of a Set - The Complement of Set A, written as A’ , is the set of all elements in the given Universal Set (U), that are not in Set A.

Example: Let U = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} and A = {1,3, 5, 7, 9}

Find A'

U = {$\cancel{1}$, 2, $\cancel{3}$, 4, $\cancel{5}$, 6, $\cancel{7}$, 8, $\cancel{9}$, 10}

So, A = {2, 4, 6, 8, 10}

Try these on your own!

Given the set descriptions below, answer the following questions

$U = All\ Integers\ from$ 1 to 10. $A = Odd\ Integers\ from$ 1 to 10,

$B = Even\ Integers\ from$ 1 to 10, $C = Multiples\ of$ 2 $from$ 1 to 10.

1. Write each of the sets in roster notation. U = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} , A = {1, 3, 5, 7, 9} , B = {2, 4, 6, 8, 10} , C = {2, 4, 6, 8, 10}

2. What is the cardinality of Sets U and A? Cardinality: U-10, A-5

3. Are Set B and Set C Equal? Yes, they are Equal

4. Are Set A and Set C Equivalent? Yes, they are Equivalent

5. How many Proper Subsets of Set $U$are there? 2_{10}☐ − 1 = 1023

6. Find $\boldsymbol{B}$’and $\boldsymbol{C'}$ B’ = C’ = {1, 3, 5, 7, 9}

7. Find $\boldsymbol{A} \cup \boldsymbol{C'}$ $A \cup C$’= {1, 3, 5, 7, 9}

8. Find $\boldsymbol{B}$’$\cap$ $\boldsymbol{C}$ B’ $\cap$ C = { } or ∅

Finit and infinite sets

Definition of Finite Set

As the name represents, the finite set is a set having finite or countable number of elements.

Example

It is a set of all English alphabets.

As we can count the number of elements here, so it a finite set.

Cardinality of a Finite Set

The cardinality of a finite set is n(A)= a, where, a represents the number of elements of set A.

As in the above picture, the cardinality of this set is 26, as the number of elements are 26.

So, n (A) = 26.

This shows that if you can list all the elements of a set and write them in the curly braces or you can say in Roster form are called the finite sets.

Sometimes it may possible that the number of elements is very big but somewhere it is countable or it has starting and end point then it is a non empty finite set. Here we denote the number of elements with n(A) and if n(A) is a natural number then you can say that it is a finite set.

Properties of Finite Sets

- The subset of a finite set is always finite.
- The union of two finite sets is finite.
- The power set of a finite set is finite.

Let's see with example

A = {1, 2, 3, 4}

B = {2, 4, 6, 8}

C = {2, 3}

Here, all A,B and C are the finite sets as there number of elements are limited and countable.

- $C \subset A$, i.e., C is the subset of A, as all the elements of set C are present in set A. So the subset of a finite set is always finite.
- A B is {1, 2, 3, 4, 6, 8}, so the union of two finite sets is also finite.
- The number of elements of a power set of a set is 2^n, so the number of elements of the power set of set A is $2^5 = 32$, as the number of elements of set A is 5.This shows that the power set of a finite set is finite.

<u>Example</u>

Z = { a set of number of people live in Europe}

In this example, it is difficult to count the number of people live in Europe, but it is somewhere a natural number. So it is a non empty finite set.

The finite set can be represented in sequence,

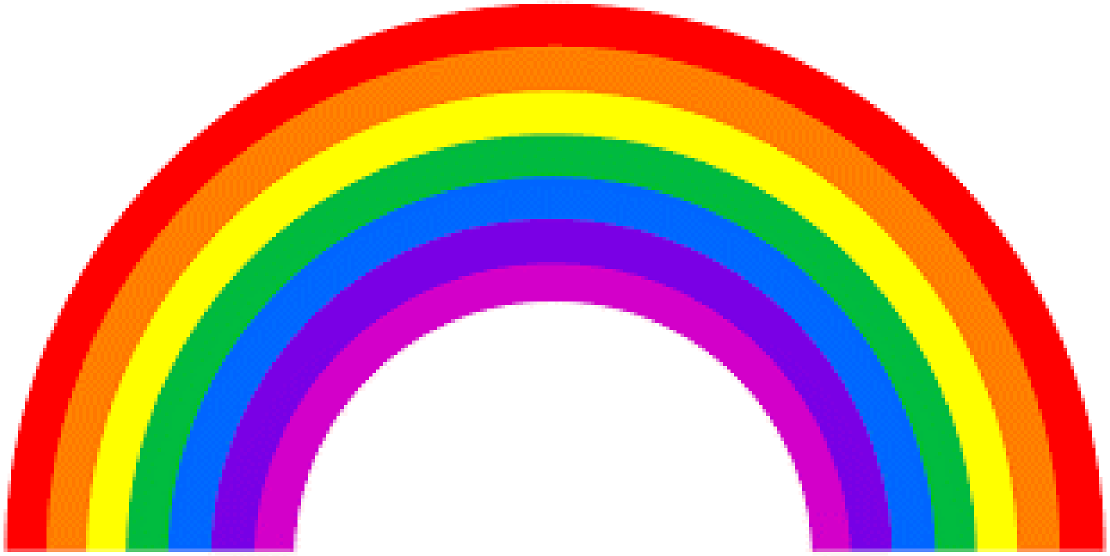
N is a set of natural numbers less then n.So the cardinality of set N is n.

N = {1, 2, 3,…, n}

Y = y_1, y2,…, n

Y = $\{y:y_1 \in N, y_1 \leq i \leq n\}$, where i is the integers between 1 and n.

Examples of Finite Sets



- Set of all colors of rainbow.

  R = {Violet, Indigo, Blue, green, yellow, orange, red}

  n(R) = 7

- Set of all natural numbers between 25 and 100.

  N= {25, 26, 27,…, 100}

  n(N) =76

- Set of all days in a week.

  D = {Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday}

  n (D) =7

All the above are the examples of finite sets because the number of elements is countable in them. And their cardinality is a natural number.

Is empty Set a Finite Set?

To understand the answer of this question, first we need to understand the meaning of Empty set.

Empty Set – Empty set is a set having no element in it. It can be represented as { }, which shows that there is no element in a set.

The cardinality of an empty set is 0 (zero), i.e., the number of elements is zero.

A= { } or Ø (phi)

n (A) = 0

The finite set is a set with a countable number of elements and as the empty set has zero element in it, so it is definite number of element.

The empty set is a finite set with a cardinality of zero.

Definition of Infinite Sets

A set which is not a finite set is called an Infinite Set. Or if you cannot count the number of elements of a particular set then it is said to be an infinite set.

As we represent a finite set in roster form, we cannot represent an infinite set in roster form easily as its elements are not limited so we use the three dots (ellipses) to represent the infinity of a set.

Example

Natural Numbers (N)
{1,2,3,4,….}
Integers (Z)
{…-2,-1,0,1,2,…}

As in the above example,

N is a set of all natural numbers starting from zero. Its number of elements is not countable so we use the three dots to represent its infinity.

Z is a set of all integers, as its elements are also uncountable so we use the three dots both the sides for the infinity of negative and positive integers.

It is important to note that the sets must have some well defined structure or pattern then only we can write it in roster form, so it is not possible to write all the infinite sets in roster form.

As we cannot write the set of real numbers in roster form as there is no proper structure of these numbers.

Cardinality of Infinite Sets

Cardinality of a set is n (A) = x, where x is the number of elements of a set A.

As the number of elements in an infinite set is unlimited, so the cardinality of an infinite set is n (A) = ∞, i.e., infinite.

Properties of Infinite Sets

- The union of two infinite sets is infinite
- The power set of an infinite set is infinite
- The super set of an infinite set is also infinite

As the number of elements of an infinite set is unlimited so its power set and supersets also need to be infinite.

Examples of Infinite Sets

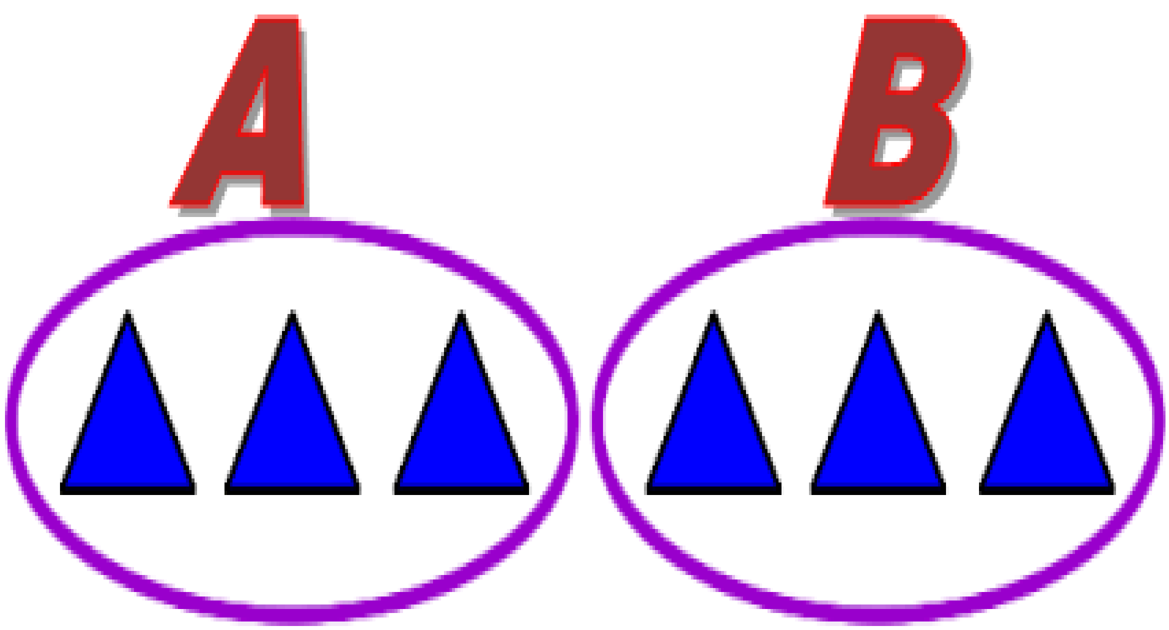
- A set of all whole numbers.

  W= {1, 2, 3, 4,…}

- A set of all points on a line.
- A set of all triangles.

What is the meaning of Equal Sets in Math?

In mathematics, we said a number equal to other if they are exactly same. Similarly in sets, we said two sets to be equal if there all the elements are same. The order of elements and the repetition of elements do not have any relevance.



Here, Set A and Set B are equal sets as there elements are exactly same.and there number of elements is also same.

<u>Example</u>

A = {5, 6, 7, 8}

B = {6, 8, 5, 7}

C = {5, 5, 6, 6, 7, 7, 8, 8}

Here all the three sets, set A, set B and set C are equal, as there elements are same irrelevance of order and the repetition.

What is the difference between Finite and Infinite Sets?

The difference between finite and infinite sets is as follows:

The sets could be equal only if there elements are same, so a set could be equal only if it is a finite set. And if a set is infinite, we cannot compare the elements of the sets.

No.	Points	Finite Sets	Infinite Sets
1	Definition	A set is a finite set if it is empty or a limited number of elements.	A set which is not a finite set is an infinite set.
2	Number of elements	Countable number of elements.	Uncountable number of elements.
3	Continuity	It starts and also stops.	It has no end either in the beginning or in last or could have both sides' continuous.
4	Cardinality	n (A) = n, n is the number of elements.	n (A) =∞, infinite as the number of elements are uncountable.
5	Union	Union of two finite sets is finite.	Union of two infinite sets is infinite.
6	Power set	Power set of a finite set is finite.	Power set of an infinite set is infinite.
7	Roster form	Can be easily represented in roster form.	All sets cannot be shown in roster form so we use three dots to represent the infinity.
8	Example	A = {2, 4, 6, 8} A set of even numbers less than 9.	X = {2, 4, 6, 8,…} A set of all even numbers.

How to determine if a Set is Finite or Infinite?

As we know that a set is finite if it has a starting point and an ending point both, but a set is said to be infinite if it has no end from any side or both sides.

Points to determine a set as finite or infinite are:

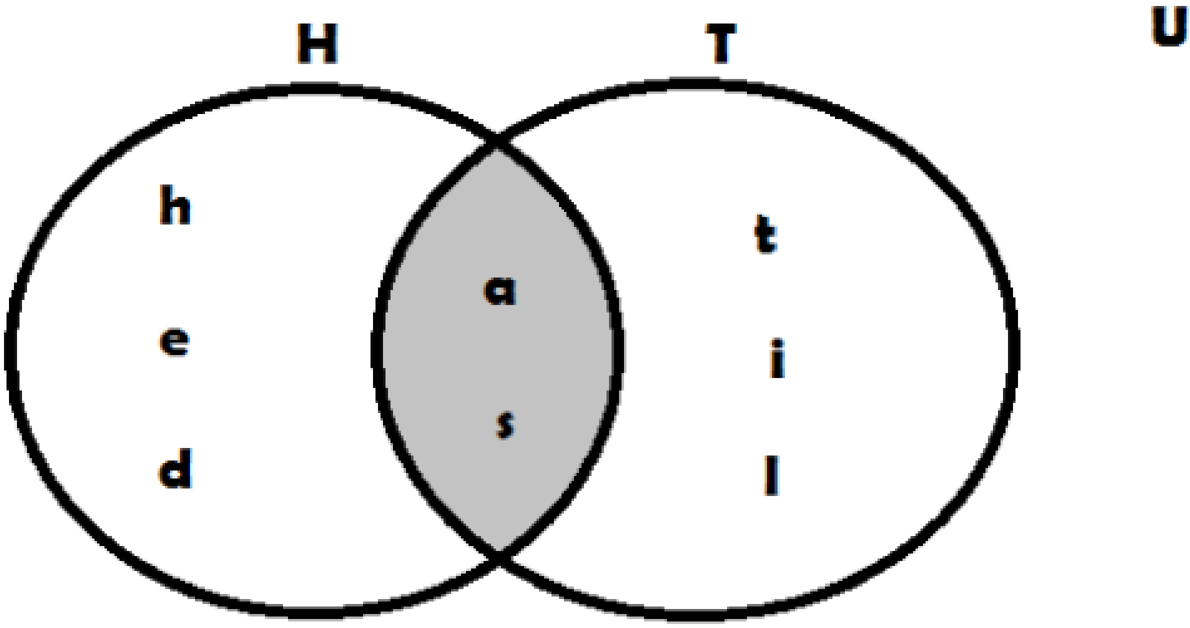
- If a set has a starting and end point both then it is finite but if it does not have a starting or end point then it is infinite set.
- If a set has a limited number of elements then it is finite but if its number of elements is unlimited then it is infinite.

What is Finite or Infinite?

Let’s try to determine a set whether it is finite or infinite with their elements.

No.	Examples	Finite or Infinite	Why?
1	A={5,10,15,20}	Finite	This set has both starting point and ending point and its number of elements are limited.
2	B= {5, 10, 15, 20,…}	Infinite	This set has a starting point but not an ending point. As the multiples of 5 could not be countable.
3	C= {…,-2,-1, 0, 1, 2,…}	Infinite	This set has no starting point and not even an end point, so its number of elements is uncountable.
4	D= {x:x W and 0<x<10}	Finite	This is the set of whole numbers between 0 and 10. So has limited number of elements.
5	E={x:x∈R and x-2=10}	Finite	This is a set of real numbers where the elements of this set are those where x-2=10.As the set of real numbers is infinite but here the equal sign make it finite.
6	F={x:x∈R and x+4>12}	Infinite	Here the elements of this set are x+4 is anything greater than 12. As the set of real numbers is infinite so there is no end point of this set.

Graphical Representation of Finite and Infinite Sets



Here in the above picture,

H = {a, s, h, e, d}

T = {a, s, t, i, l}

HUT = {a, s, h, e, d, t, i, l}

H∩T = {a, s}

Both H and T are finite sets as they have limited number of elements.

n(H) = 5 and n(T) = 5

HUT and H∩T are also finite.

This shows that we can easily represent the finite sets through venn diagram .

The union of two finite sets is finite.

The intersection of two finite sets is also finite.

But it is difficult to represent an infinite set with venn diagram, as it has unlimited number of elements and it can not be bounded in a circle to represent.

Algebra of sets:

Intuitively, a set is a “collection” of objects known as “elements.” But in the early 1900’s, a radical transformation occurred in mathematicians’ understanding of sets when the British philosopher Bertrand Russell identified a fundamental paradox inherent in this intuitive notion of a set (this paradox is discussed in exercises 66–70 at the end of this section). Consequently, in a formal set theory course, a set is defined as a mathematical object satisfying certain axioms. These axioms detail properties of sets and are used to develop an elegant and sophisticated theory of sets. This “axiomatic” approach to describing mathematical objects is relevant to the study of all areas of mathematics, and we begin exploring this approach later in this chapter. For now, we assume the existence of a suitable axiomatic framework for sets and focus on their basic relationships and operations. We first consider some examples.

Example 2.1.1

Each of the following collections of elements is a set.

• V = {cat, dog, fish}

• $W = \{1, 2\}$

• $X = \{1, 3, 5\}$

• $Y = \{n : n \text{ is an odd integer}\} = \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}$

In many settings, the upper case letters A, B, . . . , Z are used to name sets, and a pair of braces {,} is used to specify the elements of a set. In example 2.1.1, V is a finite set of three English words identifying common household pets. Similarly, W is finite set consisting of the integers 1 and 2, and X is a finite set consisting of the integers 1, 3, and 5. We have written Y using the two most common notations for an infinite set. As finite beings, humans cannot physically list every element of an infinite set one at a time. Therefore, we often use the descriptive set notation $\{n : P(n)\}$, where $P(n)$ is a predicate stating a property that characterizes the elements in the set. Alternatively, enough elements are listed to define implicitly a pattern and ellipses “. . .” are used to denote the infinite, unbounded nature of the set. This second notation must be used carefully, since people vary considerably in their perception of patterns, while clarity and precision are needed in mathematical exposition.

Certain sets are of widespread interest to mathematicians. Most likely, they are already familiar from your previous mathematics courses. The following notation, using “barred” upper case letters, is used to denote these fundamental sets of numbers.

Definition 2.1.1

• $\emptyset$ denotes the empty set { }, which does not contain any elements.

• N denotes the set of natural numbers $\{1, 2, 3, \ldots\}$.

• Z denotes the set of integers $\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$.

• Q denotes the set of rational numbers { p/q : p, q ∈ Z with q 6= 0 }.

• R denotes the set of real numbers consisting of directed distances from a designated point zero on the continuum of the real line.

• C denotes the set of complex numbers { a + bi : a, b ∈ R with i = $\sqrt{-1}$ }.

In this definition, various names are used for the same collection of numbers. For example, the natural numbers are referred to by the mathematical symbol "N," the English words "the natural numbers," and the set-theoretic notation "{1, 2, 3, . . .}." Mathematicians move freely among these different ways of referring to the same number system as the situation warrants. In addition, the mathematical symbols for these sets are "decorated" with the superscripts "∗," "+," and "−" to designate the corresponding sub collections of nonzero, positive, and negative numbers, respectively. For example, applying this symbolism to the integers Z = {. . . , −3, −2, −1, 0, 1, 2, 3, . . .}, we have

$$Z^* = \{\ldots, -3, -2, -1, 1, 2, 3, \ldots\},$$

$$Z^+ = \{1, 2, 3, \ldots\},$$

$$Z^- = \{-1, -2, -3, \ldots\}.$$

There is some discussion in the mathematics community concerning whether or not zero is a natural number. Many define the natural numbers in terms of the "counting" numbers 1, 2, 3, . . . (as we have done here) and refer to the set {0, 1, 2, 3, . . .} as the set of whole numbers. On the other hand, many mathematicians think of zero as a "natural" number. For example, the axiomatic definition of the natural numbers introduced by the Italian mathematician Giuseppe Peano in the late 1800s includes zero. Throughout this book, we use definition 2.1.1 and refer to the natural numbers as the set N = { 1, 2, 3, . . . }.

Our study of sets focuses on relations and operations of sets. The most fundamental relation associated with sets is the "element of" relationship that indicates when an object is a member of a set.

Definition 2.1.2

If a is an element of set A, then a ∈ A denotes "a is an element of A."

Example 2.1.2

As in example 2.1.1, let W = {1, 2} and recall that Q is the set of rationals.

• 1 is in W, and so 1 ∈ W.

• 3 is not in W, and so 3 6∈ W.

• 1/2 is rational, and so 1 2 ∈ Q.

• √ 2 is not rational (as we prove in section 3.4), and so √ 2 6∈ Q.

Question 2.1.1

Give an example of a finite set A with 2 ∈ A and an infinite set B with 2 6∈ B.

We now consider relationships between sets. We are particularly interested in describing when two sets are identical or equal. As it turns out, the identity relationship on sets is best articulated in terms of a more primitive “subset” relationship describing when all the elements of one set are contained in another set.

Definition 2.1.3

Let A and B be sets.

• A is a subset of B if every element of A is an element of B. We write A ⊆ B and show A ⊆ B by proving that if a ∈ A, then a ∈ B.

• A is equal to B if A and B contain exactly the same elements. We write A = B and show A = B by proving both A ⊆ B and B ⊆ A.

• A is a proper subset of B if A is a subset of B, but A is not equal to B. We write either A ⊂ B or A (B and show A ⊂ B by proving both A ⊆ B and B 6⊆ A.

Formally, the notation and the associated proof strategy are not part of the definition of these set relations. However, these facts are fundamental to working with sets and you will want to become adept at transitioning freely among definition, notation, and proof strategy.

Example 2.1.3

As in example 2.1.1, let W = {1, 2}, X = {1, 3, 5}, and Y ={n : n is an odd integer}. We first prove X ⊆ Y and then prove W 6⊆ Y.

Proof that X ⊆ Y We prove X ⊆ Y by showing that if a ∈ X, then a ∈ Y. Since X = {1, 3, 5} is finite, we prove this implication by exhaustion; that is, we consider every element of X one at a time and verify that each is in Y. Since 1 = 2 · 0 + 1, 3 = 2 · 1 + 1, and 5 = 2 · 2 + 1, each element of X is odd; in particular, each element of X has been expressed as 2k + 1 for some k ∈ Z). Thus, if a ∈ X, then a ∈ Y, and so X ⊆ Y.

Proof that W ⊆ Y We prove W ⊆ Y by showing that a ∈ W does not necessarily imply a ∈ Y. Recall that (p → q) is false precisely when [p ∧ (∼q)] is true; in this case, we need to identify a counterexample with a ∈ W and a 6∈ Y. Consider 2 ∈ W. Since 2 = 2 · 1 is even, we conclude 2 6∈ Y. Therefore, not every element of W is an element of Y.

Question 2.1.2

As in example 2.1.1, let $X = \{1, 3, 5\}$ and $Y = \{n : n$ is an odd integer $\}$. Prove that X is a proper subset of Y.

Example 2.1.4

The fundamental sets of numbers from definition 2.1.1 are contained in one another according to the following proper subset relationships.

$$\emptyset \subset N \subset Z \subset Q \subset R \subset C$$

When working with relationships among sets, we must be careful to use the notation properly so as to express true mathematical statements. One common misuse of set-theoretic notation is illustrated by working with the set $W = \{1, 2\}$. While it is true that $1 \in W$ since 1 is in W, the assertion that $\{1\} \in W$ is not true. In particular, W contains only numbers, not sets, and so the set{1}is not in W. In general, some sets do contain sets—W is just not one of these sets. Similarly, we observe that $\{1\} \subseteq W$ since $1 \in \{1, 2\} = W$, but $1 \subseteq W$ is not true; indeed, $1 \subseteq W$ is not a sensible mathematical statement since the notation $\subseteq$ is not defined between an element and a set, but only between sets.

Despite these distinctions, there is a strong connection between the "element of" relation $\in$ and the subset relation $\subseteq$, as you are asked to develop in the following question. In this way, we move beyond discussing relationships among specific sets of numbers to exploring more general, abstract properties that hold for every element and every set.

Question 2.1.3

Prove that $a \in A$ if and only if $\{a\} \subseteq A$.
Hint: Use definitions 2.1.2 and 2.1.3 to prove the two implications forming this "if-and-only-if" mathematical statement.

We now turn our attention to six fundamental operations on sets. These set operations manipulate a single set or a pair of sets to produce a new set. When applying the first three of these operations, you will want to utilize the close correspondence between the set operations and the connectives of sentential logic.

Definition 2.1.4 Let A and B be sets.

• AC denotes the complement of A and consists of all elements not in A, but in some prespecified universe or domain of all possible elements including those in A; symbolically, we define AC = {x : x 6∈ A}.

• $A \cap B$ denotes the intersection of A and B and consists of the elements in both A and B; symbolically, we define $A \cap B = \{x : x \in A \text{ and } x \in B\}$.

• $A \cup B$ denotes the union of A and B and consists of the elements in A or in B or in both A and B; symbolically, we define $A \cup B = \{x : x \in A \text{ or } x \in B\}$.

• $A \setminus B$ denotes the set difference of A and B and consists of the elements in A that are not in B; symbolically, we define $A \setminus B = \{x : x \in A \text{ and } x$ b$\in B\}$. We often use the identity $A \setminus B = A \cap B$ C.

• $A \times B$ denotes the Cartesian product of A and B and consists of the set of all ordered pairs with first-coordinate in A and second-coordinate in B; symbolically, we define $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$.

• P(A) denotes the power set of A and consists of all subsets of A; symbolically, we define $P(A) = \{X : X \subseteq A\}$. Notice that we always have $\emptyset \in P(A)$ and $A \in P(A)$.

Example 2.1.5

As above, we let $W = \{1, 2\}$, $X = \{1, 3, 5\}$ and $Y = \{n : n \text{ is an odd integer }\}$. In addition, we assume that the set of integers $Z = \{. . . , −2, −1, 0, 1, 2, . . .\}$ is the universe and we identify the elements of the following sets.

• WC = $\{. . . , −2, −1, 0, 3, 4, 5, . . .\}$

• Y C = $\{n : n \text{ is an even integer }\}$ by the parity property of the integers

• $W \cap X = \{1\}$, since 1 is the only element in both W and X

• $W \cup X = \{1, 2, 3, 5\}$, since union is defined using the inclusive-or

• $W \setminus X = \{2\}$ • $X \setminus W = \{3, 5\}$

• $Z * = Z \setminus \{0\} = \{. . . , −3, −2, −1, 1, 2, 3, . . .\}$

• $W \times X = \{(1, 1),(1, 3),(1, 5),(2, 1),(2, 3),(2, 5)\}$

• $P(W) = \{ \emptyset, \{1\}, \{2\}, \{1, 2\} \}$

The last two sets given in example 2.1.5 contain mathematical objects other than numbers; the power set is also an example of a set containing other sets. As we continue exploring mathematics, we will study sets of functions, matrices, and other more sophisticated mathematical objects.

Question 2.1.4

Working with W, X, and Y from example 2.1.5, identify the elements in the sets X C, $W \cap Y$, $W \cup Y$, $W \setminus Y$, $Y \setminus W$, $X \times W$, $W \times W$, $W \times Y$, and P(X). In addition, state six elements in P(Y); that is, state six subsets of Y.

The use of symbols to represent relationships and operations on mathematical objects is a standard feature of mathematics. Good choices in symbolism can facilitate mathematical understanding and insight, while poor choices can genuinely hinder the study and creation of mathematics. Historically, the symbols ∈ for “element of,” ∩ for “intersection,” and ∪ for “union” were introduced in 1889 by the Italian mathematician Giuseppe Peano. His work in formalizing and axiomatizing set theory and the basic arithmetic of the natural numbers remains of central importance. The Cartesian product × is named in honor of the French mathematician and philosopher René Descartes, who first formulated “analytic geometry” (an important branch of mathematics discussed in section 4.1).

Although we have presented the Cartesian product $A \times B$ as an operation on pairs of sets, this product extends to any finite number of sets. Mathematicians work with ordered triples $A \times B \times C = \{(a, b, c) : a \in A, b \in B, \text{ and } c \in C\}$, ordered quadruples $A \times B \times C \times D = \{(a, b, c, d) : a \in A, b \in B, c \in C, \text{ and } d \in D\}$, and even ordered n-tuples A1 × · · · × An = {(a1, . . . , an) : ai ∈ Ai for 1 ≤ i ≤ n}. While the use of n-tuples may at first seem to be of purely academic interest, models for science and business with tens (and even hundreds and thousands) of independent variables have become more common as computers have extended our capacity to analyze increasingly sophisticated events.

Along with considering the action of set-theoretic operations on specific sets of numbers, we are also interested in exploring general, abstract properties that hold for all sets. In this way we develop an algebra of sets, comparing various sets to determine when one is a subset of another or when they are equal. In developing this algebra, we adopt the standard approach of confirming informal intuitions and educated guesses with thorough and convincing proofs.

Example 2.1.6

For sets A and B, we prove $A \cap B \subseteq A$. Proof We prove $A \cap B \subseteq A$ by showing that if $a \in A \cap B$, then $a \in A$. We give a direct proof of this implication; we assume that $a \in A \cap B$ and show that $a \in A$. Since $a \in A \cap B$, both $a \in A$ and $a \in B$ from the definition of intersection. We have thus quickly obtained the goal of showing $a \in A$.

In example 2.1.6 we used a direct proof to show that one set is a subset of another. This strategy is very important: we prove $X \subseteq Y$ by assuming $a \in X$ and showing $a \in Y$. In addition, the process of proving $a \in X$ implies $a \in Y$ usually involves “taking apart” the sets X and Y and characterizing their elements based on the appropriate set-theoretic definitions. Once X and Y have been expanded in this way, our insights into sentential logic should enable us to understand the relationship between the two sets and to craft a proof (or disproof) of the claim. We illustrate this approach by verifying another set-theoretic identity.

Example 2.1.7 For sets A and B, we prove A \ B = A ∩ B C.

Proof In general, we prove two sets are equal by demonstrating that they are subsets of each other. In this case, we must show both A \ B ⊆ A ∩ B C and A ∩ B C ⊆ A \ B.

A \ B ⊆ A ∩ B C: We assume a ∈ A \ B and show a ∈ A ∩ B C. Since a ∈ A \ B, we know a ∈ A and a 6∈ B. The key observation is that a 6∈ B is equivalent to a ∈ B C from the definition of set complement. Since a ∈ A and a 6∈ B, we have both a ∈ A and a ∈ B C. Therefore, by the definition of intersection, a ∈ A ∩ B C. Thus, we have A \ B ⊆ A ∩ B C, completing the first half of the proof.

A ∩ B C ⊆ A \ B: We assume a ∈ A ∩ B C and show a ∈ A \ B. From the definition of intersection, we know a ∈ A ∩ B C implies both a ∈ A and a ∈ B C. Therefore, both a ∈ A and a 6∈ B from the definition of complement. This is exactly the definition of set difference, and so a ∈ A\B. Thus, A∩B C ⊆ A\B,completing the second half of the proof.

The proof of these two subset relationships establishes the desired equality A \ B = A ∩ B C for every set A and B.

Question 2.1.5 Prove that if A and B are sets with A ⊆ B, then B C ⊆ A C.

A whole host of set-theoretic identities can be established using the strategies illustrated in the preceding examples. As we have seen, the ideas and identities of sentential logic play a fundamental role in working with the set-theoretic operations. Recall that De Morgan’s laws are among the most important identities from sentential logic; consider the following set-theoretic version of these identities.

Example 2.1.8 De Morgan’s laws for sets

We prove one of De Morgan’s laws for sets: If A and B are sets, then both (A ∩ B) C = A C ∪ B C and (A ∪ B) C = A C ∩ B C.

Proof We prove the identity (A ∩ B) C = A C ∪ B C by arguing that each set is a subset of the other based on the following biconditionals:

a ∈ (A ∩ B) C	iff a 6∈ A ∩ B	Definition of complement
	iff a is not in both A and B	Definition of intersection
	iff either a 6∈ A or a 6∈ B	Sentential De Morgan’s laws
	iff either a ∈ A C or a ∈ B C	Definition of complement
	iff a ∈ A C ∪ B C	Definition of union

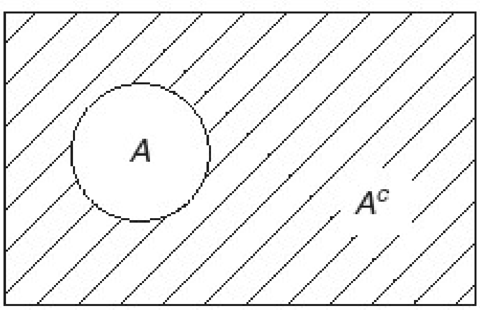
Working through these biconditionals from top to bottom, we have a ∈ (A ∩ B) C implies a ∈ A C ∪B C, and so (A∩B) C ⊆ A C ∪B C. Similarly, working through these biconditionals from bottom to top, we have a ∈ A C ∪ B C implies a ∈ (A ∩ B) C, and so A C ∪ B C ⊆ (A ∩ B) C. This proves one of De Morgan’s laws for sets, (A ∩ B) C = A C ∪ B C for every set A and B.

Question 2.1.6

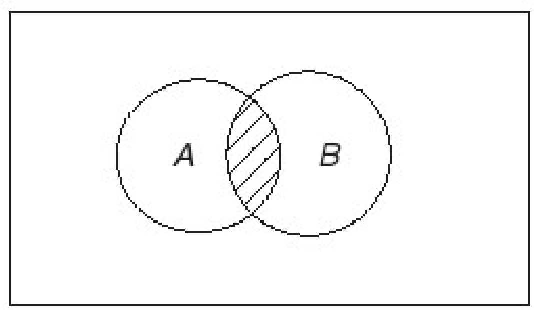
Prove the other half of De Morgan’s laws for sets; namely, prove that if A and B are sets, then $(A \cup B) C = A C \cap B C$. We end this section by discussing proofs that certain set-theoretic relations and identities do not hold. From section 1.7, we know that (supposed) identities can be disproved by finding a counterexample, exhibiting specific sets for which the given equality does not hold. To facilitate the definition of sets A, B,C with the desired properties, we introduce a visual tool for describing sets and set operations known as a Venn diagram. In a Venn diagram, the universe is denoted with a rectangle, and sets are drawn inside this rectangle using circles or ellipses. When illustrating two or more sets in a Venn diagram, we draw overlapping circles to indicate the possibility that the sets may share some elements in common. The Venn diagrams for the first four set operations from definition 2.1.4 are given in figure 2.1.

Example 2.1.9

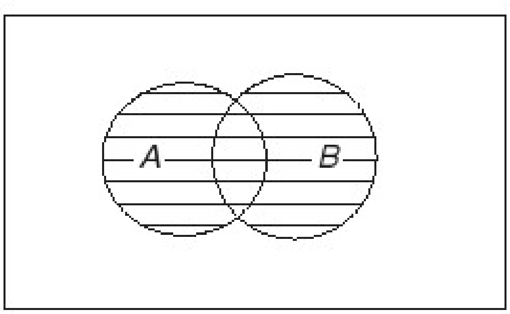
We disprove the false claim that if A, B, and C are sets, then $A \cap (B \cup C) = (A \cap B) \cup C$. This demonstrates that union and intersection operations are not associative when used together, and so we must be careful with the order of operation when “mixing” union and intersection.



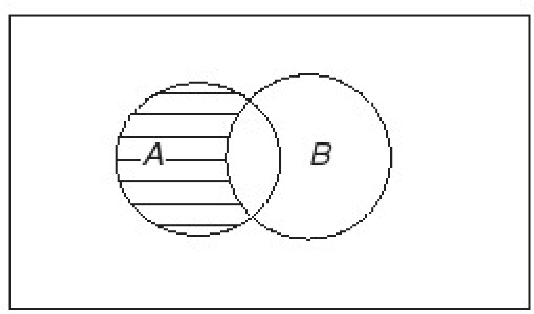
The shaded set is A^c.



The shaded set is $A \cap B$.



The shaded set is $A \cup B$.



The shaded set is $A \setminus B$.

Figure 2.1 Venn diagrams for basic set operations

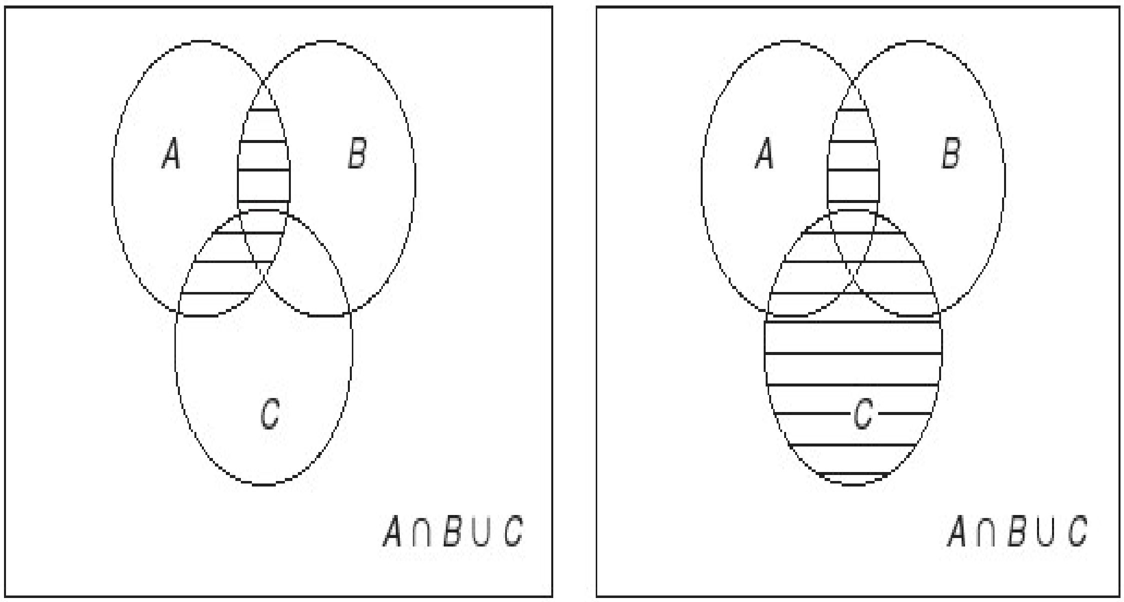


Figure 2.1 Venn diagrams for example 2.1.9 showing A ∩ (B ∪ C) 6= (A ∩ B) ∪ C

Examining the Venn diagrams, we see that if A, B,C are defined so that C contains an element that is in neither A nor B, the sets A ∩ (B ∪ C) and (A ∩ B) ∪ C will be different. Alternatively, we could define A, B,C so that B ∩ C contains an element that is not in A. Following the first approach, we choose to define the sets A = {1}, B = {1, 2}, and C = {1, 2, 3} and verify the desired inequality with the following computations.

$$A \cap (B \cup C) = \{1\} \cap \{1, 2, 3\} = \{1\}$$

$$(A \cap B) \cup C = \{1\} \cup \{1, 2, 3\} = \{1, 2, 3\}$$

Therefore these three sets provide a counterexample demonstrating that sometimes

$$A \cap (B \cup C) 6= (A \cap B) \cup C.$$

In example 2.1.9, the choice of sets A, B, and C is just one choice among many. We are certainly free to make other choices, and you might even think of constructing counterexamples as providing an opportunity to express your "mathematical personality."

Question 2.1.7

Guided by example 2.1.9, give another counterexample disproving the false claim that A ∩ (B ∪ C) = (A ∩ B) ∪ C for all sets A, B,C.

We highlight one subtlety that arises in this setting. In example 2.1.9 and question 2.1.7, the counterexamples only disprove the general claim that $A \cap (B \cup C) = (A \cap B) \cup C$ for all sets A, B,C. However, these counterexamples do not prove that we have inequality for every choice of sets. In fact, there exist many different cases in which equality does hold. For example, both $A = \emptyset$, $B = \emptyset$, $C = \emptyset$ and $A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1\}$ produce the equality $A \cap (B \cup C) = (A \cap B) \cup C$, but only because we are working with these specific sets. We therefore cannot make any general claims about the equality of $A \cap (B \cup C)$ and $(A \cap B) \cup C$, but must consider each possible setting on a case-by-case basis. In short, if we want to prove that a settheoretic identity does not always hold, then a counterexample accomplishes this goal; if we want to prove that a set-theoretic identity never holds, then we must provide a general proof and not just a specific (counter) example.

Question 2.1.8

Sketch the Venn diagram representing the following sets.

(a) $(A \cup B) \cap C$ (b) $A\ C \setminus B$

Question 2.1.9

Following the model given in example 2.1.9, disprove the false claim that the following identities hold for all sets A, B,C.

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

2.1.1 Reading Questions for Section 2.1

1. What is the intuitive definition of a set?
2. What is the intuitive definition of an element?
3. Describe two approaches to identifying the elements of an infinite set.
4. Name six important sets and the symbolic notation for these sets.
5. Define and give an example of the “element of” relation $a \in A$.
6. Define and give an example of the set relations: $A \subseteq B$, $A = B$, and $A \subset B$.
7. If A and B are sets, what strategy do we use to prove that $A \subseteq B$?
8. If A and B are sets, what strategy do we use to prove that $A = B$?
9. Define and give an example of the set operations: $A\ C$, $A \cap B$, $A \cup B$, $A \setminus B$, $A \times B$, and $P(A)$.
10. Define and give an example of a generalized Cartesian product $A1 \times A2 \times \cdots \times An$.

11. State both the sentential logic and the set-theoretic versions of De Morgan's laws.

12. Discuss the use of a Venn diagram for representing sets.

2.1.2 Exercises for Section 2.1 In exercises 1–14, identify the elements in each set, assuming A = {w, x, y, z} is the universe, B = { x, y }, C = { x, y, z }, and D = { x, z }.

1. B C 2. C C
3. B ∩ C 4. B ∩ D
5. B ∪ C 6. B ∪ D
7. B ∩ (C ∪ D) 8. (B ∩ C) ∪ D
9. B \ D 10. D \ B
11. B × C 12. B × D
13. P(B) 14. P(C)

In exercises 15–22, identify the elements in each set, assuming A = (0, 2) = {x : 0 < x ≤ 2} and B = [1, 3) = {x : 1 ≤ x < 3} are subsets of the real line R.

15. A C 16. B C

17. A ∩ B 18. A ∪ B

19. A \ B 20. B \ A

21. A C ∩ B C 22. A C ∪ B C

In exercises 23–27, give an example proving each subset relationship is proper.

23. ∅ ⊂ N 24. N ⊂ Z

25. Z ⊂ Q 26. Q ⊂ R

27. R ⊂ C

In exercises 28–41, prove each set-theoretic identity for sets A, B, and C.

28. {2, 2, 2} = {2} 29. {1, 2} = {2, 1}

30. {1} ∈ P({1}) 31. A ⊆ A (and so A ∈ P(A))

32. ∅ ⊆ A (and so ∅ ∈ P(A)) 33. A \ ∅ = A

34. [A C] C = A 35. A ∩ B ⊆ A

36. A ∩ ∅ = ∅ 37. A ⊆ A ∪ B

38. If A ⊆ B and B ⊆ C, then A ⊆ C. 39. If A ⊆ B and A ⊆ C, then A ⊆ B ∩ C.

40. (A ∪ B) \ C = (A \ C) ∪ (B \ C) 41. If A ⊆ B, then P(A) ⊆ P(B).

In exercises 42–45, disprove each false set-theoretic identity.

42. $1 = \{1\}$

43. $1 \subseteq \{1\}$

44. $\{1\} \in \{1\}$

45. $\{1\} \subseteq P(\{1\})$

For exercises 46–53, disprove the false claim that the following hold for all sets A, B, C by describing a counterexample.

46. If A 6⊆ B and B 6⊆ C, then A 6⊆ C.

47. If $A \subseteq B$, then $A\ C \subseteq B\ C$.

48. If $A\ C = B\ C$, then $A \cup B = \emptyset$.

49. If $A\ C = B\ C$, then $A \cap B = \emptyset$.

50. If $A \cup C = B \cup C$, then $A = B$.

51. If $A \cap C = B \cap C$, then $A = B$.

52. If $B = A \cup C$, then $A = B \setminus C$.

53. $(A \setminus B) \cup (B \setminus C) = A \setminus C$

Exercises 54–57 consider “disjoint” pairs of sets. We say that a pair of sets X and Y is disjoint when they have an empty intersection; that is, when $X \cap Y = \emptyset$.

In exercises 54–57, let $B = \{x, y\}$,$C = \{x, y, z\}$, $D = \{x, z\}$, $E = \{y\}$, and $F = \{w\}$ and identify the sets in this collection that are disjoint from the following sets.

54. B

55. C

56. D

57. E

Exercises 58–62 explore numeric properties of the power set operation.

58. State every element in $P(\emptyset)$. How many elements are in $P(\emptyset)$?

59. State every element in $P(\{1\})$. How many elements are in $P(\{1\})$?

60. State every element in $P(\{1, 2\})$. How many elements are in $P(\{1, 2\})$?

61. State every element in $P(\{1, 2, 3\})$. How many elements are in $P(\{1, 2, 3\})$?

62. Based on your answers to exercises 58–61, make a conjecture about how many elements are in $P(\{1, 2, 3, 4\})$. Extend your conjecture to $P(\{1, 2, \ldots, n\})$.

Exercises 63–65 consider how mathematicians have utilized set theory as a tool for defining the natural numbers. In particular, a correspondence between the nonnegative integers $\{0, 1, 2, 3, \ldots\}$ and certain sets is defined, beginning as follows.

$0 = \emptyset$

$1 = \{0\} = \{\emptyset\}$

$2 = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$

$3 = \{0, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$

63. Using this model as a guide, state the set corresponding to the integer 4.

64. Using this model as a guide, state the set corresponding to the integer 5.

65. For each natural number from 0 to 5, how many elements are in the corresponding set? Based on this observation make a conjecture of how many elements are in the set for the natural number 50.

Exercises 66–67 consider the Barber paradox that was introduced by Bertrand Russell in an effort to illuminate Russell's paradox (discussed in the exercises 68–70). The Barber paradox is based on the following question.

If the barber shaves everyone who doesn't shave themselves and only those who don't shave themselves, who shaves the barber?

66. Assume the barber does not shave himself and find a contradiction.

67. Assume the barber shaves himself and find a contradiction.

Exercises 68–70 consider Russell's paradox. A set N is said to be normal if the set does not contain itself; symbolically, we write N 6∈ N. Examples of normal sets include the set of all even integers (which is itself not an even integer) and the set of all cows (which is itself not a cow). An example of a set that is not normal is the set of all thinkable things (which is itself thinkable).

68. Give two more examples of normal sets and an example of a set that is not normal.

69. Let N be the set of all normal sets. Assume N is a normal set and find a contradiction.

70. Let N be the set of all normal sets. Assume N is not a normal set and find a contradiction.

Bertrand Russell pointed out this paradox in our intuitive understanding of sets in a letter to Gottlob Frege in 1903. This paradox holds when a set is defined as "any collection" of objects and highlights the interesting observation that not every collection is a set.

2.2 The Division Algorithm and Modular Addition

Our study of abstract algebra begins with the system of whole numbers known more formally as the integers. Recall that Z denotes the set of integers {. . . , −3, −2, −1, 0, 1, 2, 3, . . .}. From previous mathematics courses, we are already familiar with several operations on the integers, including addition, subtraction, multiplication, division, and exponentiation. In this chapter, we "push the boundaries" on these operations by studying certain subsets of the integers along with a modified addition operation known as modular addition. We use the division algorithm to define this new addition operation.

The division algorithm is actually the name of a theorem, but the standard proof of this result describes the long division algorithm for integers. The ancient Greek mathematician Euclid included the division algorithm in Book VII of Elements [73], a comprehensive survey of geometry and number theory. Traditionally, Euclid is believed to have taught and written at the Museum and Library of Alexandria in Egypt, but otherwise relatively little is known about him. And yet Elements is arguably the most important mathematics book ever written, appearing in more editions than any book other than the Christian Bible.

By the time Elements had appeared in 300 b.c.e., Greek mathematicians had recognized a duality in the fundamental nature of geometry. On the one hand, geometry is empirical, at least to the extent that it describes the physical space we inhabit. On the other hand, geometry is deductive because it uses axioms and reasoning to establish mathematically certain truths. Mathematicians and others continue to wonder at this duality. As Albert Einstein questioned, “How can it be that mathematics, being after all a product of human thought independent of experience, is so admirably adapted to the objects of reality?”

Mathematicians have a special affection for Euclid’s book because Elements is the first known comprehensive exposition of mathematics to utilize the deductive, axiomatic method. In addition, a Latin translation of Euclid’s Elements played a fundamental role in fostering the European mathematical renaissance of the sixteenth and seventeenth centuries. We now formally state the division algorithm.

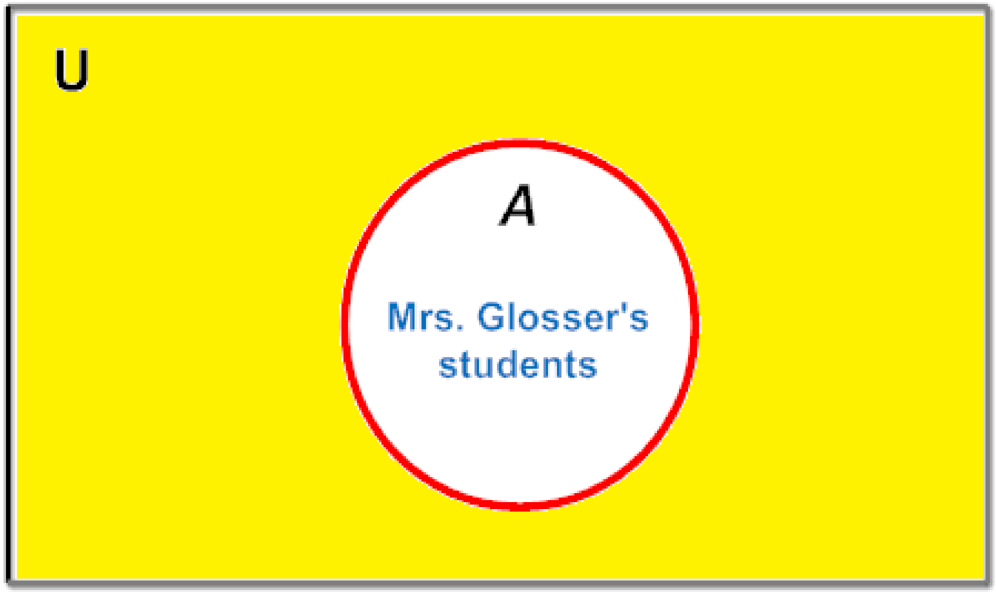
Complementation

Complement

In previous lessons, we learned that a set is a group of objects, and that Venn diagrams can be used to illustrate both set relationships and logical relationships.

Example 1: Given U = {students who attend The Kewl School} *and* A = {students in Mrs. Glosser's class}. What is the set of all students who attend The Kewl School that are not in Mrs. Glosser's class?

Analysis: The relationship between these sets is illustrated in the Venn diagram below.



Answer: The shaded area outside *A* represents all students who attend The Kewl School that are not in Mrs. Glosser's class.

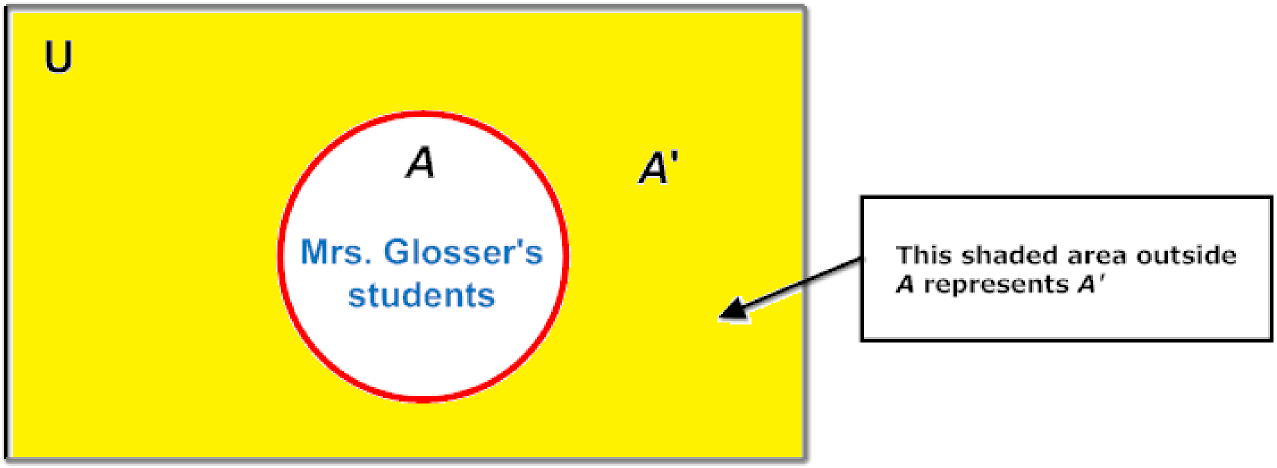
In example 1, the shaded area represents the complement of Set A. The complement of A, denoted by A', consists of all students in The Kewl school that are not in Mrs. Glosser's class. Recall that a Universal Set is the set of all elements under consideration, denoted by capital U, and that all other sets are subsets of the Universal Set. Now we can define the complement of a set.

Definition: The complement of a set *A*, denoted by *A'*, is the set of elements which belong to U but which do not belong to *A*.

The complement of set *A* is denoted by *A*', You can also say *"complement of A in* U*"*, or "A-prime". We can now label the sets in example 1 using this notation.

Example 1: Given U = {students who attend The Kewl School} *and* *A* = {students in Mrs. Glosser's class}. What is the set of all students who attend The Kewl School that are not in Mrs. Glosser's class?

Analysis: The relationship between these sets is illustrated in the Venn diagram below.



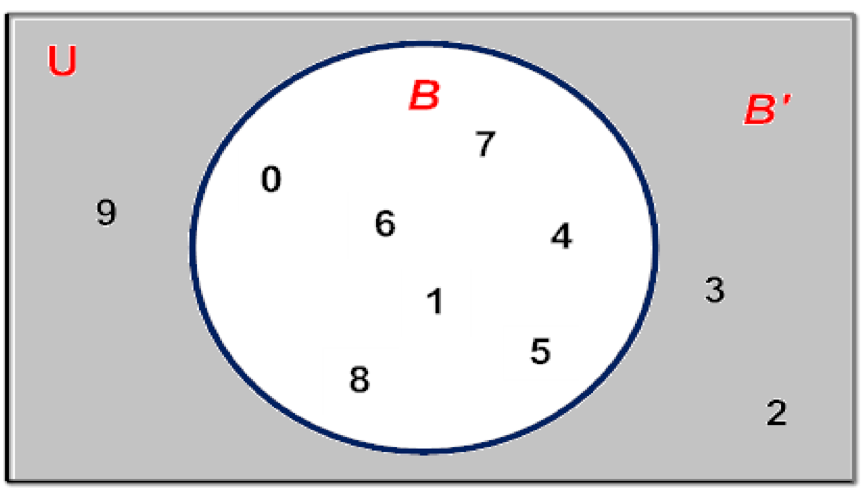
Answer: The shaded area outside *A* represents *A***'**, which is all students who attend The Kewl School that are not in Mrs. Glosser's class.

Another way to think of the complement of a set is as follow: Given set *A*,
the complement of *A* is the set of all elements in the universal set U, that are not in *A*.
Using set-builder notation, we can write:

$A' = \{ x \mid x \in U \text{ and } x \notin A \}$

Let's find the complement of a set of numbers.

Example 2: Given U = {single digits} and *B* = {0, 1, 4, 5, 6, 7, 8}, find the complement of *B*.



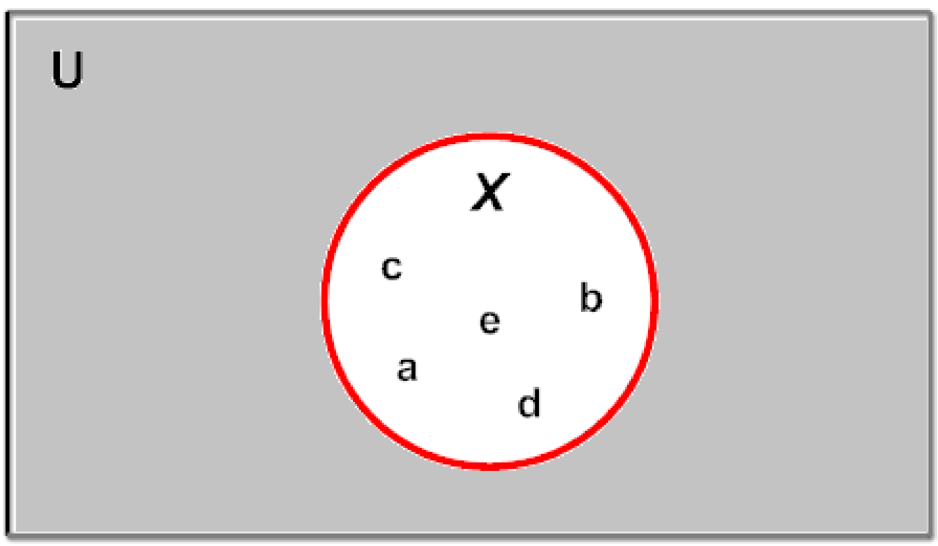
Answer: *B′* = {2, 3, 9}

Thus *B′* is the set of all numbers in $\mathcal{U}$ that are not in *B*. Using set-builder notation, we can write: *B′* = { x | x $\in$ $\mathcal{U}$ and x $\notin$ *B* }

In examples 3 through 5, the universal set is the English alphabet.

Example 3: Given $\mathcal{U}$ = {a, b, c, **...**, x, y, z} and *X* = {a, b, c, d, e}, find *X′*.

Analysis: *X′* would consist of all letters in the English alphabet that are not in *X.* This is shown in the Venn Diagram below.

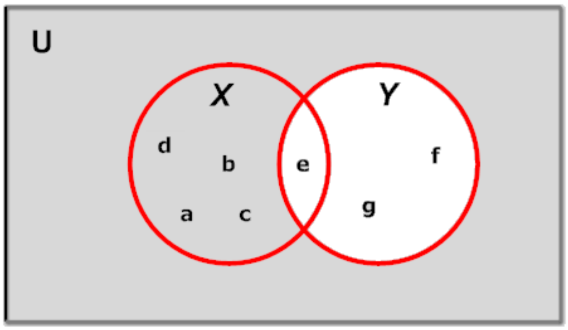


X′ is shaded

Answer: *X′* = {f, g, h, **...**, x, y, z}

Example 4: Given $\mathcal{U}$ = {a, b, c, **...**, x, y, z}*, X* = {a, b, c, d, e} *and* *Y* = {e, f, g}, find *Y′.*

Analysis: *Y′* would consist of all letters in the English alphabet that are not in *Y.* This is shown in the Venn Diagram below..

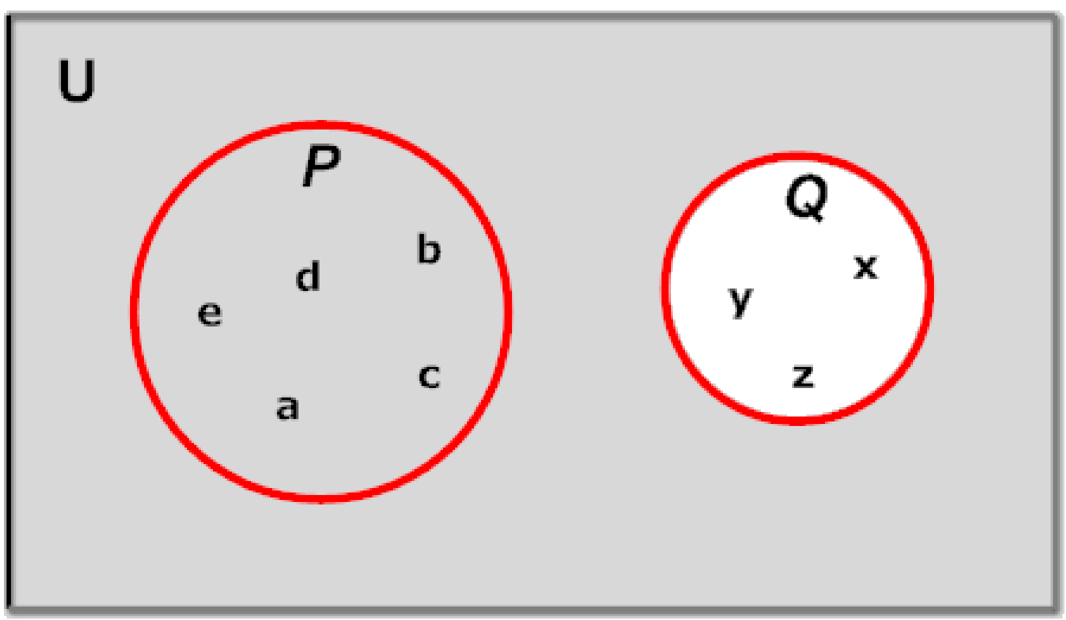


Y′ is shaded

Answer: *Y′* = {a, b, c, d, h, i, j, **...**, x, y, z}

Example 5: Given *U* = {a, b, c, **...**, x, y, z}, *P* = {a, b, c, d, e} *and* *Q* = {x, y, z}, find *Q′.*

Analysis: *Q′* consists of all the letters in the alphabet that are not in Q*.* This is shown in the Venn Diagram below..



Q′ is shaded

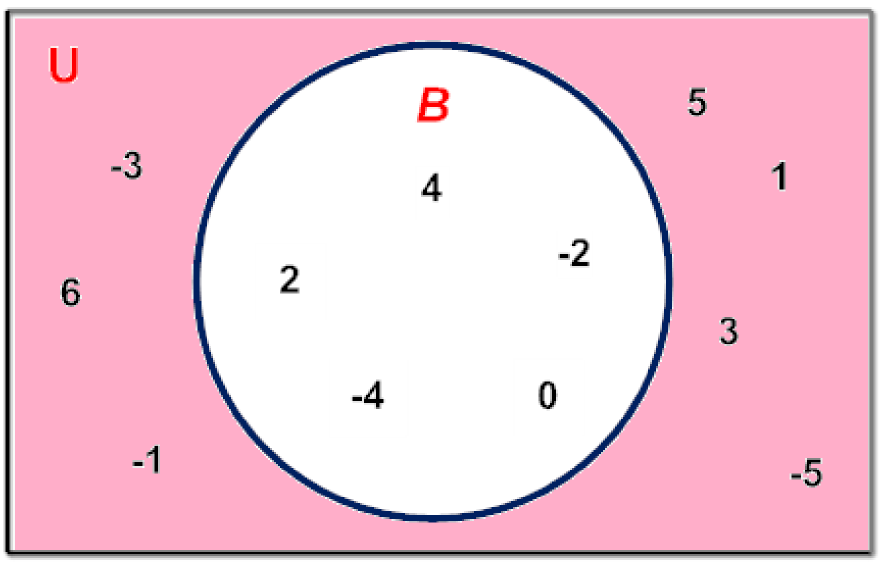
Answer: *Q′* = {a, b, c, d, e, f, g, h **...**, u, v, w}

Looking at the examples above, a set and its complement have no elements in common. The union of a set and its complement is the Universal Set. The intersection of a set and its complement is the null set. These statements are summarized below:

All of these notations have the same meaning. However, for the purpose of this instructional unit, we have chosen to use ***A′*,** read as *A*-prime.

Let's look at some examples of complement that involve set-builder notation and infinite sets.

Example 6: If U = { n **|** n ∈ Z and -6 < n < 7 } and B = { y **|** y even number; -5 < y < 6 }, then what is the complement of *B*?

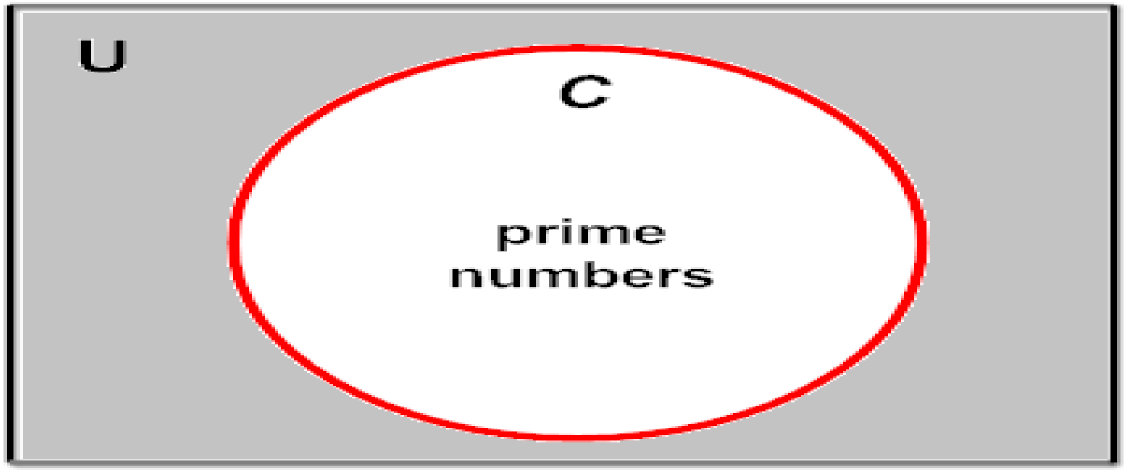


B′ **is shaded**

Answer: *B′* = {-5, -3, -1, 1, 3, 5, 6}

Example 7: Given U = {counting numbers > 1} and *C* = {prime numbers}, find *C′*.

Analysis: *C′* would consist of all counting numbers greater than 1 that are not prime. This is shown in the Venn Diagram below.



C' is shaded

Answer: *C'* = {composite numbers}

Summary: Given set *A*, the complement of *A* is the set of all element in the universal set U, that are not in *A*. The complement of set *A* is denoted as *A'* and is read as *A-prime*. The formal definition of complement is shown below.

$A' = \{ x \mid x \in U \text{ and } x \notin A \}$

Demorgan's law

If PP is some sentence or formula, then $\neg P \neg P$ is called the **denial** of PP. The ability to manipulate the denial of a formula accurately is critical to understanding mathematical arguments. The following tautologies are referred to as De Morgan's laws:

$$\neg(P \vee Q)\neg(P \wedge Q) \Leftrightarrow (\neg P \wedge \neg Q) \Leftrightarrow (\neg P \vee \neg Q)$$

These are easy to verify using truth tables, but with a little thought, they are not hard to understand directly. The first says that the only way that $P \vee Q P \vee Q$ can fail to be true is if both PP and QQ fail to be true. For example, the statements "I don't like chocolate or vanilla" and "I do not like chocolate and I do not like vanilla" clearly express the same thought. For a more mathematical example of the second tautology, consider "xx is not between 2 and 3." This can be written symbolically as $\neg((2<x) \wedge (x<3))\neg((2<x) \wedge (x<3))$, and clearly is equivalent to $\neg(2<x) \vee \neg(x<3), \neg(2<x) \vee \neg(x<3)$, that is, $(x \le 2) \vee (3 \le x)$.

We can also use De Morgan's laws to simplify the denial of $P \Rightarrow Q P \Rightarrow Q$:

$$\neg(P \Rightarrow Q) \Leftrightarrow \neg(\neg P \vee Q) \Leftrightarrow (\neg\neg P) \wedge (\neg Q) \Leftrightarrow P \wedge \neg Q \neg(P \Rightarrow Q) \Leftrightarrow \neg(\neg P \vee Q) \Leftrightarrow (\neg\neg P) \wedge (\neg Q) \Leftrightarrow P \wedge \neg Q$$

so the denial of $P\Rightarrow QP\Rightarrow Q$ is $P\wedge\neg QP\wedge\neg Q$. In other words, it is not the case that PP implies QQ if and only if PP is true and QQ is false. Of course, this agrees with the truth table for $P\Rightarrow QP\Rightarrow Q$ that we have already seen.

There are versions of De Morgan's laws for quantifiers:

$$\neg\forall xP(x)\neg\exists xP(x)\Leftrightarrow\exists x$$

$$\neg P(x)\Leftrightarrow\forall x\neg P(x)$$

You may be able to see that these are true immediately. If not, here is an explanation of $\neg\forall xP(x)\Rightarrow\exists x\neg P(x)\neg\forall xP(x)\Rightarrow\exists x\neg P(x)$ that should be convincing: If $\neg\forall xP(x)\neg\forall xP(x)$, then $P(x)P(x)$ is not true for every xx, which is to say that for some value aa, $P(a)P(a)$ is not true. This means that $\neg P(a)\neg P(a)$ is true. Since $\neg P(a)\neg P(a)$ is true, it is certainly the case that there is some value of xx that makes $\neg P(x)\neg P(x)$ true, which is to say that $\exists x\neg P(x)\exists x\neg P(x)$ is true. The other three implications may be explained in a similar way.

Here is another way to think of the quantifier versions of De Morgan's laws. The statement $\forall xP(x)\forall xP(x)$ is very much like a big conjunction. If the universe of discourse is the positive integers, for example, then it is equivalent to the statement that

$$P(1)\wedge P(2)\wedge P(3)\wedge\cdots P(1)\wedge P(2)\wedge P(3)\wedge\cdots$$

or, more concisely, we might write

$$\bigwedge x\in UP(x),$$

using notation similar to "sigma notation" for sums. Of course, this is not really a "statement" in our official mathematical logic, because we don't allow infinitely long formulas. In the same way, $\exists xP(x)\exists xP(x)$ can be thought of as

$$\bigvee x\in UP(x).\bigvee x\in UP(x).$$

Now the first quantifier law can be written

$$\neg\bigwedge x\in UP(x)\Leftrightarrow\bigvee x\in U(\neg P(x)),\neg\bigwedge x\in UP(x)\Leftrightarrow\bigvee x\in U(\neg P(x)),$$

which looks very much like the law

$$\neg(P\wedge Q)\Leftrightarrow(\neg P\vee\neg Q),\neg(P\wedge Q)\Leftrightarrow(\neg P\vee\neg Q),$$

but with an infinite conjunction and disjunction. Note that we can also rewrite De Morgan's laws for ∧∧ and ∨∨ as

$$\neg\bigwedge i=12(Pi(x))\neg\bigvee i=12(Pi(x))\Leftrightarrow\bigvee i=12(\neg Pi(x))\Leftrightarrow\bigwedge i=12(\neg Pi(x)).\neg\bigwedge i=12(Pi(x))\Leftrightarrow\bigvee i=12(\neg Pi(x))$$

$$\neg\bigvee i=12(Pi(x))\Leftrightarrow\bigwedge i=12(\neg Pi(x)).$$

This is more cumbersome, but it reflects the close relationship with the quantifier forms of De Morgan's laws.

Finally, general understanding is usually aided by specific examples: Suppose the universe is the set of cars. If $P(x)P(x)$ is "xx has four wheel drive," then the denial of "every car has four wheel drive" is "there exists a car which does not have four wheel drive." This is an example of the first law. If $P(x)P(x)$ is "xx has three wheels," then the denial of "there is a car with three wheels" is "every car does not have three wheels." This fits the pattern of the second law. In a more mathematical vein, a denial of the sentence "for every xx, $x2x2$ is positive" is "there is an xx such that $x2x2$ fails to be positive." A denial of "there is an xx such that $x2=-1x2=-1$" is "for every xx, $x2\neq-1x2\neq-1$."

It is easy to confuse the denial of a sentence with something stronger. If the universe is the set of all people, the denial of the sentence "All people are tall" is not the sentence "No people are tall." This might be called the **opposite** of the original sentence—it says more than simply "`All people are tall' is untrue." The correct denial of this sentence is "there is someone who is not tall," which is a considerably weaker statement. In symbols, the denial of $\forall x P(x)\forall x P(x)$ is $\exists x\ \neg P(x)\exists x\ \neg P(x)$, whereas the opposite is $\forall x\neg P(x)\forall x\neg P(x)$. ("Denial" is an "official" term in wide use; "opposite," as used here, is not widely used.)

De Morgan's laws can be used to simplify negations of the "some" form and the "all" form; the negations themselves turn out to have the same forms, but "reversed," that is, the negation of an "all" form is a "some" form, and vice versa.
Suppose $P(x)P(x)$ and $Q(x)Q(x)$ are formulas. We then have

$$\neg\forall x(P(x)\Rightarrow Q(x))\Leftrightarrow\ \exists x(P(x)\wedge\neg Q(x))\neg\forall x(P(x)\Rightarrow Q(x))\Leftrightarrow\ \exists x(P(x)\wedge\neg Q(x))$$

$$\neg\exists x(P(x)\wedge Q(x))\Leftrightarrow\ \forall x(P(x)\Rightarrow\neg Q(x))\neg\exists x(P(x)\wedge Q(x))\Leftrightarrow\ \forall x(P(x)\Rightarrow\neg Q(x))$$

The denial of the sentence "all lawn mowers run on gasoline" is the sentence "some lawn mower does not run on gasoline" (not "no lawn mowers run on gasoline," the opposite). We verify the first statement and leave the second for an exercise:

$$\neg\forall x(P(x)\Rightarrow Q(x))\Leftrightarrow\exists x\neg(P(x)\Rightarrow Q(x))\Leftrightarrow\exists x(P(x)\wedge\neg Q(x))\neg\forall x(P(x)\Rightarrow Q(x))\Leftrightarrow\exists x\neg(P(x)\Rightarrow Q(x))\Leftrightarrow\exists x(P(x)\wedge\neg Q(x))$$

A formula is usually simpler if $\neg\neg$ does not appear in front of any compound expression, that is, it appears only in front of simple statements such as $P(x)P(x)$. The following is an example of simplifying the denial of a formula using De Morgan's laws:

$$\neg\forall x(P(x)\vee\neg Q(x))\Leftrightarrow\exists x\neg(P(x)\vee\neg Q(x))\Leftrightarrow\exists x(\neg P(x)\wedge\neg\neg Q(x))\Leftrightarrow\exists x(\neg P(x)\wedge Q(x))\neg\forall x(P(x)\vee\neg Q(x))\Leftrightarrow\exists x\neg(P(x)\vee\neg Q(x))\Leftrightarrow\exists x(\neg P(x)\wedge\neg\neg Q(x))\Leftrightarrow\exists x(\neg P(x)\wedge Q(x))$$

Denials of formulas are extremely useful. In a later section we will see that the techniques called proof by contradiction and proof by contrapositive use them extensively. Denials can also be a helpful study device. When you read a theorem or a definition in mathematics it is frequently helpful to form the denial of that sentence to see what it means for the condition to fail. The more ways you think about a concept in mathematics, the clearer it should become.

Augustus De Morgan. (yy–1871; De Morgan himself noted that he was xx years old in the year x2x2.) De Morgan's father died when he was ten, after which he was raised by his mother, a devout member of the Church of England, who wanted him to be a minister. Far from becoming a minister, De Morgan developed a pronounced antipathy toward the Church, which would profoundly influence the course of his career.

De Morgan's interest in and talent for mathematics did not become evident until he was fourteen, but already at sixteen he entered Trinity College at Cambridge, where he studied algebra under George Peacock and logic under William Whewell. He was also an excellent flute player, and became prominent in musical clubs at Cambridge.

On graduation, De Morgan was unable to secure a position at Oxford or Cambridge, as he refused to sign the required religious test (a test not abolished until 1875). Instead, at the age of 22, he became Professor of Mathematics at London University, a new institution founded on the principle of religious neutrality.

De Morgan wrote prolifically about algebra and logic. Peacock and Gregory had already focused attention on the fundamental importance to algebra of symbol manipulation; that is, they established that the fundamental operations of algebra need not depend on the interpretation of the variables. De Morgan went one (big) step further: he recognized that the operations (++, −−, etc.) also need have no fixed meaning (though he made an exception for equality). Despite this view, De Morgan does seem to have thought that the only appropriate interpretations for algebra were familiar numerical domains, primarily the real and complex numbers. Indeed, he thought that the complex numbers formed the most general possible algebra, because he could not bring himself to abandon the familiar algebraic properties of the real and complex numbers, like commutativity.

One of De Morgan's most widely known books was *A Budget of Paradoxes*. He used the word `paradox' to mean anything outside the accepted wisdom of a subject. Though this need not be interpreted pejoratively, his examples were in fact of the `mathematical crank' variety—mathematically naive people who insisted that they could trisect the angle or square the circle, for example.

De Morgan's son George was himself a distinguished mathematician. With a friend, he founded the London Mathematical Society and served as its first secretary; De Morgan was the first president.

In 1866, De Morgan resigned his position to protest an appointment that was made on religious grounds, which De Morgan thought abused the principle of religious neutrality on which London University was founded. Two years later his son George died, and shortly thereafter a daughter died. His own death perhaps hastened by these events, De Morgan died in 1871 of `nervous prostration.'

The information here is taken from *Lectures on Ten British Mathematicians*, by Alexander Macfarlane, New York: John Wiley & Sons, 1916.

Exercises 1.3

Ex 1.3.1 Verify these tautologies using truth tables.

$$\neg(P\vee Q)\Leftrightarrow(\neg P\wedge\neg Q)\neg(P\vee Q)\Leftrightarrow(\neg P\wedge\neg Q)$$

$$\neg(P\wedge Q)\Leftrightarrow(\neg P\vee\neg Q)\neg(P\wedge Q)\Leftrightarrow(\neg P\vee\neg Q)$$

Ex 1.3.2 Suppose R(x)R(x) is the statement "xx is a rectangle," and S(x)S(x) is the statement "xx is a square." Write the following symbolically and decide which pairs of statements are denials of each other:

a) All rectangles are squares.

b) Some rectangles are squares.

c) Some squares are not rectangles.

d) No squares are rectangles.

e) No rectangles are squares.

f) All squares are rectangles.

g) Some squares are rectangles.

h) Some rectangles are not squares.

Ex 1.3.3 Write symbolically the following denials of definitions concerning a function ff:

a) ff is not increasing. c) ff is not constant.

b) ff is not decreasing. d) ff does not have a root.

Ex 1.3.4 Simplify the following expressions:

a) $\neg\forall x>0(x2>x)\neg\forall x>0(x2>x)$ c) $\neg\forall x\forall y(xy=y2\Rightarrow x=y)\neg\forall x\forall y(xy=y2\Rightarrow x=y)$

b) $\neg\exists x\in[0,1](x2+x<0)\neg\exists x\in[0,1](x2+x<0)$ d) $\neg\exists x\exists y(x>y\wedge y>x)\neg\exists x\exists y(x>y\wedge y>x)$

Ex 1.3.5 Verify the statement:

$$\neg\exists x(P(x)\wedge Q(x))\Leftrightarrow \forall x(P(x)\Rightarrow\neg Q(x))\neg\exists x(P(x)\wedge Q(x))\Leftrightarrow \forall x(P(x)\Rightarrow\neg Q(x))$$

Ex 1.3.6 Observe that

$$P\vee Q\Leftrightarrow\neg\neg(P\vee Q)\Leftrightarrow\neg(\neg P\wedge\neg Q)P\vee Q\Leftrightarrow\neg\neg(P\vee Q)\Leftrightarrow\neg(\neg P\wedge\neg Q)$$

so ∨∨ can be expressed in terms of ∧∧ and ¬¬.

a) Show how to express ⇒⇒ in terms of ∧∧ and ¬¬.

b) Show how to express ∧∧ in terms of ¬¬ and ∨∨.

c) Show how to express ∨∨ in terms of ¬¬ and ⇒⇒.

Ex 1.3.7 Express the universal quantifier ∀∀ in terms of ∃∃ and ¬¬. Express ∃∃ in terms of ∀∀ and ¬¬.

Ex 1.3.8 Compute the year yy of De Morgan's birth.

Common application of algebra of sets.

It’s easy to think of algebra as an abstract notion that has no use in real life. Understanding the history and the practical applications of algebra that are put into use every day might make you see it a little differently.

The main idea behind algebra is to replace numbers (or other specific objects) by symbols. This makes things a lot simpler: instead of saying “I’m looking for a number so that when I multiply it by 7 and add 3 I get 24”, you simply write $7x+3=24$, where x is the unknown number.

Algebra is a huge area in mathematics, and there are many mathematicians who spend their time thinking about what you can do with collections of abstract symbols. In real life, however, algebra merges into all other areas as a tool. Whenever life throws a maths problem at you, for example when you have to solve an equation or work out a geometrical problem, algebra is usually the best way to attack it. The equations you are learning about now are the ones that you’re most likely to come across in everyday life. This means that knowing how to solve them is very useful. If you’re planning to go into computer programming, however, the algebra you’ll need is more complicated and now’s the time to make sure you get the basics.

Did you know? The word algebra comes from the ancient Arabic word “al jebr”, which means the “reunion of broken parts”.

Solving equations

We have already seen how practical applications of algebra can be used to solve equations. You will often see equations like $3x+4=5$, where you want to find x.
Using algebra, you can give a recipe for solving any equation of this form:
if $ax+b=c$, then $x=(c-b)/a$.
So whenever you have to solve one of these, you don't have to go through the whole process of rearranging the equation. Instead you can just plug your numbers a, b and c into the recipe and get the answer. Read our linear equations article to see a practical application of algebra that you might already be familiar with.

Algebra in Geometry

Two-dimensional shapes can be represented using a co-ordinate system. Saying that a point has the co-ordinates $(4,2)$ for example, means that we get to that point by taking four steps into the horizontal direction and 2 in the vertical direction, starting from the point where the two axes meet.
Using algebra, we can represent a general point by the co-ordinates (x,y). You may have already learnt that a straight line is represented by an equation that looks like $y=mx+b$, for some fixed numbers m and b. There are similar equations that describe circles and more complicated curves. Using these algebraic expressions, we can compute lots of things without ever having to draw the shapes. For example we can find out if and where a circle and a straight line meet, or whether one circle lies inside another one. See the article on geometry to find out about its uses.

Algebra in computer programming

As we have seen, algebra is about recognizing general patterns. Rather than looking at the two equations $3x+1=5$ and $6x+2=3$ as two completely different things, Algebra sees them as being examples of the same general equation $ax+b=c$. Specific numbers have been replaced by symbols.
Computer programming languages, like C++ or Java, work along similar lines. Inside the computer, a character in a computer game is nothing but a string of symbols. The programmer has to know how to present the character in this way. Moreover, he or she only has a limited number of commands to tell the computer what to do with this string. Computer programming is all about representing a specific context, like a game, by

abstract symbols. A small set of abstract rules is used to make the symbols interact in the right way. Doing this requires algebra.

As Stanford University Education Professor William Damon says, schools need to give students a better understanding of why they are in school in the first place. In particular, students need to know why they are learning what is being taught. They need to understand how the knowledge and skills they are learning can help them accomplish their life goals. That is the only way to motivate students in a lasting way. (Tully, Susannah, "Helping Students Find a Sense of Purpose," The Chronicle Review, March 13, 2009, p. B14-B15.)

The "No Child Left Behind" Act has focused attention on reading and mathematics, assuming that students would understand that these subjects are important. But, as we have found in writing this booklet, neither students nor, surprisingly, their teachers are able to cite simple practical applications of elementary algebra. Hence it is no wonder that student motivation is so weak.

We believe that by presenting simple practical applications of algebra, students will gain a clearer understanding of why they are studying this subject.

As Professor Damon elaborates:

Students learn bits of knowledge that they may see little use for; and from time to time someone at a school assembly urges them to go and do great things in the world. When it comes to drawing connections between the two—that is, showing students how a math formula or a history lesson could be important for some purpose that a student may wish to pursue—schools too often leave their students flat.

If you visit a typical classroom and listen for the teacher's reasons for why the students should do their schoolwork, you will hear a host of narrow, instrumental goals, such as doing well in the course, getting good grades, and avoiding failure, or perhaps—if the students are lucky—the value of learning a specific skill for its own sake. But rarely (if ever) will you hear the teacher discuss with students broader purposes that any of these goals might lead to . . . How can schools expect that young people will find meaning in what they are doing if they so rarely draw their attention to considerations of the personal meaning and purpose of the work others do?

. . . most pervasive is a sense of emptiness that has ensnared many young people in long periods of drift during a time in their lives when they should be defining aspirations and making progress toward their fulfillment.

For too many young people today, apathy and anxiety have become the dominant moods, and disengagement or even cynicism has replaced the natural hopefulness of youth. That is not a problem that can be addressed by solutions advanced in the past. The message that young people do best when they are challenged to strive must be expanded to include an answer to the question: For what purpose? (ibid.)

This booklet was suggested by Brianne Blanton during her culminating oral examination for a master's degrees in educational administration at California State University-Bakersfield. Her husband, Peter, teaches algebra, and she had noticed that many students wanted to know "why?"

This booklet provides sample answers to first year algebra students who ask "why are we studying each of the California Algebra I Standards. It is written by graduate educational administration students who hope that it will be useful for algebra teachers.

Dr. Louis Wildman
Professor and Coordinator
Educational Administration Program
California State University-Bakersfield

After exploring the algebra of sets, we study two number systems denoted Zn and U(n) that are closely related to the integers. Our approach is based on a widely used strategy of mathematicians: we work with specific examples and look for general patterns. This study leads to the definition of modified addition and multiplication operations on certain finite subsets of the integers. We isolate key axioms, or properties, that are satisfied by these and many other number systems and then examine number systems that share the "group" properties of the integers. Finally, we consider an application of this mathematics to check digit schemes, which have become increasingly important for the success of business and telecommunications in our technologically based society. Through the study of these topics, we engage in a thorough introduction to abstract algebra from the perspective of the mathematician— working with specific examples to identify key abstract properties common to diverse and interesting mathematical systems.

The algebra of sets is an analysis of values. This lesson provides an overview of the properties of sets and laws of set theory and illustrates them with real-life examples.

Sets in Real Life

Do you have a favorite meal? Maybe it's a cheeseburger meal from your favorite hamburger restaurant. This meal probably includes a cheeseburger, French fries, a drink, some ketchup packets, and napkins. In real life, this is what we call a set.

The technical definition of a set is a collection of very specific objects. Let's go through the properties and laws of set theory in general.

Set Theory

A set of anything has to have specific criteria and be well defined. For example, one person may think that a cheeseburger dinner from a fast food restaurant is amazing, while someone else might be repulsed by its taste, making the criteria invalid. An example of a valid set would be edible foods that include bread, so the cheeseburger dinner would qualify. Let's make a list of foods and determine which ones are eligible for a set of edible items that include bread; we'll call our set "S."

S = {sandwich, hamburger, cheeseburger, toast, bread pudding}

The symbol ∈ indicates that something is part of a set. For example, grilled cheese ∈ S means that grilled cheese is part of set S. This is a true statement because grilled cheese is a sandwich. Ice cream ∉ S means that ice cream is not part of set S because it doesn't include bread.

Let's take a look at the properties of sets. The order of items in a set doesn't matter. In our set of edible foods that include bread, we could list toast first and sandwich last. If an item in a set is repeated, count it once. For example, let's say we have a set W that represents the letters in the word "cheeseburger". The example is here: W = {c, h, e, e, s, e, b, u, r, g, e, r}. As there are four e's and two r's, we can rewrite the set as W = {c, h, e, s, b, u, r, g}.

The Laws of Sets

Let's take a look at the different laws of sets one at a time.

1. Union of Sets

Let's say that we have two sets: S = {sandwich, hamburger, cheeseburger, toast, bread pudding} and B = {hamburger, cheeseburger}. We'll refer back to these sets throughout the rest of the lesson. The **union** of these sets is all items that are part of both sets, or ∪. The union of sets S and B is written as A ∪ B = {sandwich, hamburger, cheeseburger, toast, bread pudding}, which includes all of the items in both sets, but only one of each item if there are multiples.

2. Intersection of Sets

The intersection of sets defines what is common to both sets. For instance, in sets S and B, the hamburger and cheeseburger are common to both sets. The intersection of these sets is S ∩ B = {hamburger, cheeseburger}. This notation is similar to a Venn diagram of the two sets.

<u>Elementary Properties of Numbers:</u>

Mathematical Induction:-

Principle of Mathematical Induction

A class of integers is called hereditary if, whenever any integer x belongs to the class, the successor of x (that is, the integer $x + 1$) also belongs to the class. The principle of mathematical induction is then: If the integer 0 belongs to the class F and F is hereditary, every nonnegative integer belongs to F. Alternatively, if the integer 1 belongs to the class F and F is hereditary, then every positive integer belongs to F. The principle is stated sometimes in one form, sometimes in the other. As either form of the principle is easily proved as a consequence of the other, it is not necessary to distinguish between the two.

The principle is also often stated in intensional form: A property of integers is called hereditary if, whenever any integer x has the property, its successor has the property. If the integer 1 has a certain property and this property is hereditary, every positive integer has the property.

Proof by Mathematical Induction

An example of the application of mathematical induction in the simplest case is the proof that the sum of the first n odd positive integers is n^2—that is, that

$$\text{(1.)}\ 1 + 3 + 5 +\cdots+ (2n - 1) = n^2$$

for every positive integer n. Let F be the class of integers for which equation ***(1.)*** holds; then the integer 1 belongs to F, since $1 = 1^2$. If any integer x belongs to F, then

$$\textbf{(2.)}\ 1 + 3 + 5 +\cdots+ (2x - 1) = x^2.$$

The next odd integer after $2x - 1$ is $2x + 1$, and, when this is added to both sides of equation ***(2.)***, the result is

$$\textbf{(3.)}\ 1 + 3 + 5 +\cdots+ (2x + 1) = x^2 + 2x + 1 = (x + 1)^2.$$

Equation ***(2.)*** is called the hypothesis of induction and states that equation ***(1.)*** holds when n is x, while equation ***(3.)*** states that equation ***(1.)*** holds when n is $x + 1$. Since equation ***(3.)*** has been proved as a consequence of equation ***(2.)***, it has been proved that whenever x belongs to F the successor of x belongs to F. Hence by the principle of mathematical induction all positive integers belong to F.

The foregoing is an example of simple induction; an illustration of the many more complex kinds of mathematical induction is the following method of proof by double induction. To prove that a particular binary relation *F* holds among all positive integers, it is sufficient to show first that the relation *F* holds between 1 and 1; second that whenever *F* holds between *x* and *y*, it holds between *x* and *y* + 1; and third that whenever *F* holds between *x* and a certain positive integer *z* (which may be fixed or may be made to depend on *x*), it holds between *x* + 1 and 1.

The logical status of the method of proof by mathematical induction is still a matter of disagreement among mathematicians. Giuseppe Peano included the principle of mathematical induction as one of his five axioms for arithmetic. Many mathematicians agree with Peano in regarding this principle just as one of the postulates characterizing a particular mathematical discipline (arithmetic) and as being in no fundamental way different from other postulates of arithmetic or of other branches of mathematics.

Henri Poincaré maintained that mathematical induction is synthetic and a priori—that is, it is not reducible to a principle of logic or demonstrable on logical grounds alone and yet is known independently of experience or observation. Thus mathematical induction has a special place as constituting mathematical reasoning par excellence and permits mathematics to proceed from its premises to genuinely new results, something that supposedly is not possible by logic alone. In this doctrine Poincaré has been followed by the school of mathematical intuitionism which treats mathematical induction as an ultimate foundation of mathematical thought, irreducible to anything prior to it and synthetic a priori in the sense of Immanuel Kant.

Directly opposed to this is the undertaking of Gottlob Frege, later followed by Alfred North Whitehead and Bertrand Russell in *Principia Mathematica*, to show that the principle of mathematical induction is analytic in the sense that it is reduced to a principle of pure logic by suitable definitions of the terms involved.

Transfinite Induction

A generalization of mathematical induction applicable to any well-ordered class or domain *D*, in place of the domain of positive integers, is the method of proof by transfinite induction. The domain *D* is said to be well ordered if the elements (numbers or entities of any other kind) belonging to it are in, or have been put into, an order in such a way that: ***1.*** no element precedes itself in order; ***2.*** if *x* precedes *y* in order, and *y* precedes *z*, then *x* precedes *z*; ***3.*** in every non-empty subclass of D there is a first element (one that precedes all other elements in the subclass). From ***3.*** it follows in particular that the domain *D* itself, if it is not empty, has a first element.

When an element *x* precedes an element *y* in the order just described, it may also be said that *y* follows *x*. The successor of an element *x* of a well-ordered domain *D* is defined as the first element that follows *x* (since by ***3.***, if there are any elements that follow *x*, there must be a first among them). Similarly, the successor of a class *E* of elements of *D* is the first element that follows all members of *E*. A class *F* of elements of *D* is called hereditary if, whenever all the members of a class *E* of elements of D belong to F, the successor of E, if any, also belongs to F (and hence in particular, whenever an element x of D belongs to F, the successor of x, if any, also belongs to F). Proof by transfinite induction then depends on the principle that if the first element of a well-ordered domain D belongs to a hereditary class F, all elements of D belong to F.

One way of treating mathematical induction is to take it as a special case of transfinite induction. For example, there is a sense in which simple induction may be regarded as transfinite induction applied to the domain D of positive integers. The actual reduction of simple induction to this special case of transfinite induction requires the use of principles which themselves are ordinarily proved by mathematical induction, especially the ordering of the positive integers, and the principle that the successor of a class of positive integers, if there is one, must be the successor of a particular integer (the last or greatest integer) in the class. There is therefore also a sense in which mathematical induction is not reducible to transfinite induction.

The point of view of transfinite induction is, however, useful in classifying the more complex kinds of mathematical induction. In particular, double induction may be thought of as transfinite induction applied to the domain D of ordered pairs (x, y) of positive integers, where D is well ordered by the rule that the pair (x_1, y_1) precedes the pair (x_2, y_2) if $x_1 < x_2$ or if $x_1 = x_2$ and $y_1 < y_2$.

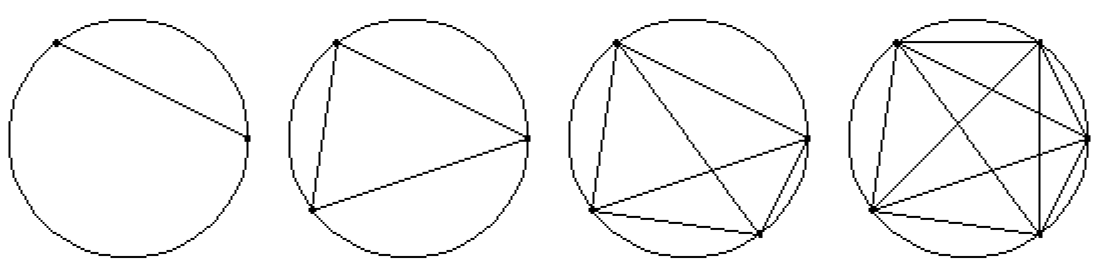
The need for proof

Most people today are lazy. We watch way too much television and are content to accept things as true without question.

If we see something that works a few times in a row, we're convinced that it works forever.

Regions of a Circle

Consider a circle with n points on it. How many regions will the circle be divided into if each pair of points is connected with a chord?



2 points	3 points	4 points	5 points
2 regions = 2^1	4 regions = 2^2	8 regions = 2^3	16 regions = 2^4

At this point, probably everyone would be convinced that with 6 points there would be 32 regions, but it's not proved, it's just conjectured. The conjecture is that the number of regions when n points are connected is 2^{n-1}.

Will finding the number of regions when there are six points on the circle prove the conjecture? No. If there are indeed 32 regions, all you have done is shown another example to support your conjecture. If there aren't 32 regions, then you have proved the conjecture wrong. In fact, if you go ahead and try the circle with six points on it, you'll find out that there aren't 32 regions.

You can never prove a conjecture is true by example.

You can prove a conjecture is false by finding a counter-example.

To prove a conjecture is true, you need some more formal methods of proof. One of these methods is the principle of mathematical induction.

Principle of Mathematical Induction (English)

1. Show something works the first time.
2. Assume that it works for this time,
3. Show it will work for the next time.
4. Conclusion, it works all the time

Principle of Mathematical Induction (Mathematics)

1. Show true for n = 1
2. Assume true for n = k
3. Show true for n = k + 1
4. Conclusion: Statement is true for all n >= 1

The key word in step 2 is assume. You are not trying to prove it's true for n = k, you're going to accept on faith that it is, and show it's true for the next number, n = k + 1. If it later turns out that you get a contradiction, then the assumption was wrong.

Annotated Example of Mathematical Induction

Prove 1 + 4 + 9 + ... + n^2 = n (n + 1) (2n + 1) / 6 for all positive integers n.

Another way to write "for every positive integer n" is $\forall n \in Z^+$. This works because Z is the set of integers, so Z^+ is the set of positive integers. The upside down A is the symbol for "for all" or "for every" or "for each" and the symbol that looks like a weird e is the "element of" symbol. So technically, the statement is saying "for every n that is an element of the positive integers", but it's easier to say "for every positive integer n".

Identify the general term and n^{th} partial sum before beginning the problem

The general term, a_n, is the last term on the left hand side. $\mathbf{a_n = n^2}$
The n^{th} partial sum, S_n, is the right hand side. $\mathbf{S_n = n (n + 1) (2n + 1) / 6}$

Find the next term in the general sequence and the series

The next term in the sequence is a_{k+1} and is found by replacing n with k+1 in the general term of the sequence, a_n.

$\mathbf{a_{k+1} = (k + 1)^2}$

The next term in the series is S_{k+1} and is found by replacing n with k+1 in the n^{th} partial sum, S_n. You may wish to simplify the next partial sum, S_{k+1}
$S_{k+1} = (k+1) [(k+1) + 1] [2(k+1) + 1] / 6$
$\mathbf{S_{k+1} = (k + 1) (k + 2) (2k +3) / 6}$ (This will be our Goal in step 3)

We will use these definitions later in the mathematical induction process. We're now ready to begin.

1. Show the statement is true for n = 1, that is, Show that $a_1 = S_1$.

a_1 is the first term on the left or you can find it by substituting n=1 into the formula for the general term, a_n.
S_1 is found by substituting n=1 into the formula for the n^{th} partial sum, S_n.

lhs: $a_1 = 1$
rhs: $S_1 = 1 (1+1) [2(1) + 1] / 6 = 1(2)(3) / 6 = 1$

So, you can see that the left hand side equals the right hand side for the first term, so we have established the first condition of mathematical induction.

2. Assume the statement is true for n = k

The left hand side is the sum of the first k terms, so we can write that as S_k. The right hand side is found by substituting n=k into the S_n formula.

Assume that $S_k = k (k + 1) (2k + 1) / 6$

3. Show the statement is true for n = k+1

What we are trying to show is that $S_{k+1} = (k + 1) (k + 2) (2k +3) / 6$. This was our goal from earlier.

We begin with something that we know (assume) is true and add the next term, a_{k+1}, to both sides.

$S_k + \mathbf{a_{k+1}} = k (k + 1) (2k + 1) / 6 + \mathbf{a_{k+1}}$

On the left hand side, $S_k + a_{k+1}$ means the "sum of the first k terms" plus "the k+1 term", which gives us the sum of the first k+1 terms, S_{k+1}.

This often gives students difficulties, so lets think about it this way. Assume k=10. Then S_k would be S_{10}, the sum of the first 10 terms and a_{k+1} would be a_{11}, the 11^{th} term in the sequence. $S_{10} + a_{11}$ would be the sum of the 10 terms plus the 11^{th} term which would be the sum of the first 11 terms.

On the right hand side, replace a_{k+1} by $(k+1)^2$, which is what you found it was before beginning the problem.

$\mathbf{S_{k+1}} = k (k + 1) (2k + 1) / 6 + \mathbf{(k + 1)^2}$

Now, try to turn your right hand side into goal of (k + 1) (k + 2) (2k +3) / 6.
You need to get a common denominator, so multiply the last term by 6/6.

$S_{k+1} = k (k + 1) (2k + 1) / 6 + \mathbf{6} (k + 1)^2 / \mathbf{6}$

Now simplify. It is almost always easier to factor rather than expand when simplifying. This is especially aided by the fact that your goal is in factored form. You can use that to help you factor. You know that you want a (k+1) (k+2) (2k+3) in the final form. We see right now that there is a (k+1) that is common to both of those, so let's begin by factoring it out.

$S_{k+1} = (k + 1) [k (2k + 1) + 6 (k + 1)] / 6$

What's left inside the brackets [ ] doesn't factor, so we expand and combine like terms.

$S_{k+1} = (k + 1) (2k^2 + k + 6k + 6) / 6$
$S_{k+1} = (k + 1) (2k^2 + 7k + 6) / 6$

Now, try to factor $2k^2 + 7k + 6$, keeping in mind that you need a (k+2) and (2k+3) in the goal that you don't have yet.

$S_{k+1} = (k + 1) (k + 2) (2k + 3) / 6$

Hey! That's what our goal was. That's what we were trying to show. That means we did it!

Conclusion

The conclusion is found by saying "Therefore, by the principle of mathematical induction" and restating the original claim.

Therefore, by the principle of mathematical induction,
$1 + 4 + 9 + ... + n^2 = n (n + 1) (2n + 1) / 6$ for all positive integers n.

Summations

Earlier in the chapter we had some summation formulas that were very melodious. In the following examples, c is a constant, and x and y are functions of the index.

You can factor a constant out of a summation

$\sum cx = c\sum x$

The sum of a constant times a function is the constant times the sum of a function.

The sum of a sum is the sum of the sums

$\sum(x+y) = \sum x + \sum y$

The summation symbol can distribute over addition.

The sum of a difference is the difference of the sums

$\sum(x-y) = \sum x - \sum y$

The summation symbol can distribute over subtraction.

Sum of the Powers of the Integers

Now, we're going to look at the sum of the whole number powers of the natural numbers.

Sigma Notation = Closed Form	Expanded
$\sum_{k=1}^{n} 1 = n$	1 + 1 + 1 + ... + 1 (n times)
$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$	1 + 2 + 3 + ... + n
$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$	1 + 4 + 9 + ... + n^2
$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}$	1 + 8 + 27 + ... + n^3

$$\sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$$ $1 + 16 + 81 + ... + n^4$

$$\sum_{k=1}^{n} k^5 = \frac{n^2(n+1)^2(2n^2+2n-1)}{12}$$ $1 + 32 + 243 + ... + n^5$

The closed form for a summation is a formula that allows you to find the sum simply by knowing the number of terms.

Finding Closed Form

Find the sum of : $1 + 8 + 22 + 42 + ... + (3n^2-n-2)$

The general term is $a_n = 3n^2-n-2$, so what we're trying to find is $\sum(3k^2-k-2)$, where the $\sum$ is really the sum from k=1 to n, I'm just not writing those here to make it more accessible.

Take the original, open form of the summation, $\sum(3k^2-k-2)$

Distribute the summation sign, $\sum 3k^2 - \sum k - \sum 2$

Factor out any constants, $3\sum k^2 - \sum k - 2\sum 1$

Replace each summation by the closed form given above. The closed form is a formula for a sum that doesn't include the summation sign, only n.

$$3\left[\frac{n(n+1)(2n+1)}{6}\right]-\left[\frac{n(n+1)}{2}\right]-2[n]$$

Now get a common denominator, in this case, 2.

$$\frac{n(n+1)(2n+1)}{2}-\frac{n(n+1)}{2}-\frac{4n}{2}$$

Remember that the word factor begins with the letter F and anytime you have a choice of doing something in mathematics that starts with the letter F, that's probably where you should start. So, do not expand, factor instead.

The common factor is n so we'll factor that out of each term. The whole expression is over 2.

$n [(n+1)(2n+1) - (n+1) - 4] / 2$

Now expand inside the brackets [ ].

$n [2n^2 + 3n + 1 - n - 1 - 4] / 2$

Simplify like terms.

$n (2n^2 + 2n - 4) / 2$

Notice the common factor of 2 inside the parentheses, let's factor that out.

$2 n (n^2 + n - 2) / 2$

The 2 in the numerator and the 2 in the denominator divide out and we can factor the rest to get the closed form for the sum.

$n (n + 2) (n - 1)$

Isn't that beautiful? At this point, we could write a mathematical induction problem similar to those in the book for this problem. It would read ...

Prove:

$$\sum_{k=1}^{n}\left(3k^2 - k - 2\right) = n\left(n+2\right)\left(n-1\right), \forall n \in Z^+$$

We're not going to prove that statement, but that is how the book came up with many of the problems that you're asked to prove. You take an open form, find the closed form, and then put it in the text as a problem for the student to prove.

Pattern Recognition

Sometimes you have to figure out what the general term of a sequence is. Here are some guidelines.

1. Calculate the first several terms of the sequence. Sometimes it helps to write the term in factored and expanded form.
2. Try to find a recognizable pattern. Here are some things to look for
    1. Linear patterns: $an + b$ (will have a common difference)
    2. Quadratic pattern: $an^2 + b$ (the perfect squares plus/minus a constant)
    3. Cubic pattern: $an^3 + b$ (the perfect cubes plus/minus a constant)
    4. Exponential patterns: $2^n + b$, $3^n + b$ (powers of 2 or 3 plus/minus a constant)
    5. Factorial patterns: $n!$, $(2n)!$, $(2n-1)!$ (factoring these really helps)

3. After you have your pattern, then you can use mathematical induction to prove the conjecture is correct.

Finite Differences

Finite differences can help you find the pattern if you have a polynomial sequence.

The first differences are found by subtracting consecutive terms. If the first differences are all the same, then the pattern is linear.

The second differences are found by subtracting consecutive first differences. If the second differences are all the same, then the pattern is quadratic. Remember that you can find a quadratic model by taking the equation $y=ax^2+bx+c$ with three points. Then solve the system of equations that results. The analogy here is that you can find $a_n=an^2+bn+c$ by substituting in three terms in the sequence for a_n and their corresponding position in the sequence for n. Then solve the system of linear equations.

This can be extended to third differences by subtracting consecutive second differences. If the third differences are all the same, the pattern is cubic. You can fit a cubic model with four points and the model $a_n=an^3+bn^2+cn+d$.

Finite Differences Example:

Find the general term of the sequence 1, -2, -1, 4, 13, 26, 43, 64, 89, ...

The first finite differences are found by subtracting consecutive terms in the original sequnce. That is, take -2-1=-3, -1-(-2)=1, 4-(-1)=5, 13-4=9, 26-13=13, etc.

The first finite differences are: -3, 1, 5, 9, 13, 17, 21, 25, ...

Since these aren't all the same, your sequence is not a linear (first degree) polynomial.

Move on to the second finite differences. These are the differences in the consecutive terms of the first finite differences. 1-(-3)=4, 5-1=4, 9-5=4, 13-9=4, 17-13=4, etc.

The second finite differences are 4, 4, 4, 4, 4, 4, 4, ...

Since the second finite differences are all the same, your sequence is that of a second degree polynomial and you can write the general term as $a_n = An^2 + Bn + C$.

Plug in the values 1, 2, 3 (since there are three variables) into the expression.

1. When n=1, 1A + 1B + 1C = 1 (the first term is 1)
2. When n=2, 4A + 2B + 1C = -2 (the second term is -2)
3. When n=3, 9A + 3B + 1C = -1 (the third term is -1)

Now, solve that system of linear equations (I recommend using matrix inverses AX=B, so $X=A^{-1}B$). If you need a refresher on how to do that, visit the section on matrix inverses in chapter 6.

In our case, we get A=2, B=-9, and C=8, so the general term of our sequence is $a_n=2n^2-9n+8$.

If you want to check it, pick any value for n and plug it into the general term. You should get the n^{th} number in the sequence. For example, if n = 6, then $a_6 = 2(6)^2 - 9(6) + 8 = 26$. Check the sequence, sure enough, the 6^{th} number is 26.

Division Algorithm

Value of Polynomial and Division Algorithm

Arithmetic operations like addition, subtraction, multiplication and division play a huge and most basic rule in Mathematics. Maths is made by these operations. All other operations go easy with the polynomials except the division operation, which gets complex when dealt with polynomials. But this section will explain to you the division of polynomials and the division algorithm related to it, from basics.

So, what's the basic formula we are learning from the day we solved our first division problem? This is:

Dividend = Quotient × Divisor + Remainder

Example: Divide the polynomial $2x^2+3x+1$ by polynomial x+2.
Solution: Divisor= x+2
Dividend=$2x^2 + 3x + 1$
Note: Put the dividend under the division sign and divisor outside the sign.

Steps for Division of Polynomials

- **Step 1:** Firstly, Arrange the divisor as well as dividend individually in decreasing order of their degree of terms.
- **Step 2:** In case of division we seek to find the quotient. To find the very first term of the quotient, divide the first term of the dividend by the highest degree term in the divisor. In the current case,

$$2x^2/x = 2x.$$

- **Step 3:** Write 2x in place of the quotient.

- **Step 4:** Multiply the divisor by the quotient obtained. Put the product underneath the dividend.
- **Step 5:** Subtract the product obtained as happens in case of a division operation.
- **Step 6:** Write the result obtained after drawing another bar to separate it from prior operations performed.
- **Step 7:** Bring down the remaining terms of the dividend.
- **Step 8:** Again divide the dividend by the highest degree term of the remaining divisor. Follow the same prior procedure until either the remainder becomes zero or its degree is less than the degree of the divisor.
- **Step 9:** At this stage, the quotient obtained is our answer.

Quotient Obtained = $2x + 1$

Note: Division Algorithms for Polynomials is same as the Long Division Algorithm In Polynomials

Download NCERT Solutions for Class 10 Maths

Division Algorithm For Polynomials

Division algorithm for polynomials states that, suppose f(x) and g(x) are the two polynomials, where $g(x)\neq 0$, we can write:

$$\mathbf{f(x) = q(x)\, g(x) + r(x)}$$

which is same as the Dividend = Divisor * Quotient + Remainder and where r(x) is the remainder polynomial and is equal to 0 and degree r(x) < degree g(x).

Verification of Division Algorithm

Take the above example and verify it.

Divisor = $x+2$
Dividend = $2x^2 + 3x + 1$
Quotient = $2x – 1$
Remainder = 0

Applying the Algorithm:

$$2x^2 + 3x + 1 = (x + 2)(2x + 1) + 0$$
$$2x^2 + 3x + 1 = 2x^2 + 3x + 1$$

Hence verified.

Finding Factors of Polynomials with Division Algorithm

Long division algorithm is used to find out factors of polynomials of degree greater than equal to two. We'll be describing the steps to find out the factors along with an example.

Example: Find roots of cubic polynomial $P(x)=3x^3 – 5x^2 – 11x – 3$

Solution

- **Step 1:** Use the factor theorem to find a factor of the polynomial.
- **Step 2:** First divide the whole equation by the coefficient of the highest degree term of the dividend.

$$P(x)=3x^3 – 5x^2 – 11x – 3$$

On dividing the whole equation by 3,

$$P(x) = x^3 – (5/3)x^2 – (11/3)x – 1$$

- **Step 3:** Find out factors of the constant term so obtained. In the present case, factors of the constant term are 1 and -1.
- **Step 4:** Put the value of x in $P(x) = 3x^3 – 5x^2 – 11x – 3$ equal to 1 and find the remainder. Again put the value of remainder equal to -1 in and find the remainder using remainder theorem. Find the value of x for which remainder is zero for the cubic polynomial.

$$P(1) = 3(1)^3 – 5(1)^2 – 11(1) – 3 = -16$$

$$P(-1) = 3(-1)^3 – 5(-1)^2 – 11(-1) – 3 = 0$$

- **Step 5:** Remainder is zero for x = -1. So, (x + 1) is a root of the polynomial.
- **Step 6:** By Division Algorithm, find out the quotient. It comes out: $3x^2 – 8x – 3$
- **Step 7:** Now, Quotient = $3x^2 – 8x – 3$

Dividend = (Divisor) * (Quotient) + Remainder

In present case,

$$3x^3 - 5x^2 - 11x - 3 = (x + 1)(3x^2 - 8x - 3) + 0$$

By factorizing the quadratic polynomial we shall be able to find out remaining factors of the cubic polynomial.

- **Step 8:** Break middle term in terms of a pair of numbers such that its product is equal to -9 and summation equal to -3.
- **Step 9:** On factorizing, possible pair of number satisfying both conditions is (-9, 1). Breaking the middle term,

$$f(x) = 3x^2 - 8x - 3$$

$$= 3x^2 - 9x + x - 3$$

- **Step 10:** Form pairs of terms and factor out GCD of the two pairs separately. Then again factor out GCD of the remaining two products.
- **Step 11:**

$$f(x) = 3x^2 - 8x - 3 = 3x^2 - 9x + x - 3$$

$$= 3x(x - 3) + 1(x - 3) = (x - 3)(3x + 1)$$

Now,

$$3x^3 - 5x^2 - 11x - 3 = (x + 1)(3x^2 - 8x - 3) + 0$$

$$= (x + 1)(x - 3)(3x + 1)$$

Factors of cubic polynomial are -1, 3 and -1/3.

Solved Example for You

Question 1: What is the division algorithm formula?

Answer: It states that for any integer, a and any positive integer b, there exists a unique integer q and r such that a = bq + r. Here r is greater than or equal to 0 and less than b. Moreover, a is the dividend, b is the divisor, q is the quotient and r is the remainder.

Question 2: Explain Euclid's division algorithm?

Answer: It refers to a technique to compute the Highest Common Factor (HCF) of two given positive integers. In addition, let us remind you that the HCF of two positive integers a and b is the largest positive integer d that divides both a and b.

Question 3: How does the division algorithm work?

Answer: It refers to an algorithm that gives two integers a and b, and when we compute their quotient and/ or remainder the result of Euclidean division. In addition, we apply some of them by hand, whereas digital circuit designs and software employ others.

Question 4: State the main difference between Lemma and algorithm?

Answer: A Lemma refers to a proven statement that we use to prove another statement. On the other hand, an algorithm refers to a series of well-defined steps that gives a procedure for solving a type of problem.

The Greatest Common Divisor

Euclidean algorithm, Euclidian Algorithm: GCD (Greatest Common Divisor) Explained with C++ and Java Examples. For this topic you must know about Greatest The numbers that these two lists share in common are the common divisors of 54 and 24: 1 , 2 , 3 , 6. {\displaystyle 1,2,3,6.\,} The greatest of these is 6. That is, the greatest common divisor of 54 and 24. One writes: gcd (54 , 24) = 6. {\displaystyle \gcd (54,24)=6.\,}

Greatest common divisor, GCD (Greatest Common Divisor) or HCF (Highest Common Factor) of two An efficient solution is to use Euclidean algorithm which is the main algorithm used The greatest common divisor polynomial g(x) of two polynomials a(x) and b(x) is defined as the product of their shared irreducible polynomials, which can be identified using the Euclidean algorithm. The basic procedure is similar to that for integers.

Code for Greatest Common Divisor in Python, GCD of two numbers is the largest number that divides both of them. A simple way to find GCD is to factorize both numbers and multiply common factors. GCD. In algebra, the greatest common divisor of two polynomials is a polynomial, of the highest possible degree, that is a factor of both the two original polynomials. This concept is analogous to the greatest common divisor of two integers. In the important case of univariate polynomials over a field the polynomial GCD may be computed, like for the integer GCD, by Euclid's algorithm using long division. The polynomial GCD is defined only up to the multiplication by an invertible constant. The simila

Greatest common divisor calculator

GCD Calculator, The GCD calculator allows you to quickly find the greatest common divisor of a set of numbers. You may enter between two and ten non-zero integers between -2147483648 and 2147483647. The numbers must be separated by commas, spaces or tabs or may be entered on separate lines. Free Greatest Common Divisor (GCD) calculator - Find the gcd of two or more numbers step-by-step This website uses cookies to ensure you get the best experience. By using this website, you agree to our Cookie Policy.

Greatest Common Factor (GCF, HCF, GCD , Calculate the GCF, GCD or HCF and see work with steps. Learn how to find the greatest common factor using factoring, prime factorization and the Euclidean Greatest Common Divisor Calculator Calculate the GCD of a set of numbers. The GCD calculator allows you to quickly find the greatest common divisor of a set of numbers.

GCD Calculator - Greatest Common Divisor, Second number: The gcd of two numbers is their greatest common divisor, i.e. the largest number that The greatest common divisor (GCD) allows us to find out the common number that you can divide a group of determined whole numbers by without having anything left over. The greatest common divisor, amongst other things, allows us to simplify fractions so it is easier to work with them. How does the GCD calculator work?

Greatest common divisor java

Java Program to Find GCD of Two Numbers, The GCD (Greatest Common Divisor) of two numbers is the largest positive integer number that divides both the numbers without leaving any remainder. For example. GCD of 30 and 45 is 15. GCD also known as HCF (Highest Common Factor). Given that BigInteger is a (mathematical/functional) superset of int, Integer, long, and Long, if you need to use these types, convert them to a BigInteger, do the GCD, and convert the result back. private static int gcdThing(int a, int b) { BigInteger b1 = BigInteger.valueOf(a); BigInteger b2 = BigInteger.valueOf(b); BigInteger gcd = b1.gcd(b2); return gcd.intValue(); }

The **java.lang.Math** class contains methods for performing basic numeric operations such as the elementary exponential, logarithm, square root, and trigonometric functions.

Class methods

Sr.No.	Method & Description
1	static double abs(double a) This method returns the absolute value of a double value.
2	static float abs(float a) This method returns the absolute value of a float value.
3	static int abs(int a) This method returns the absolute value of an int value.

4	static long abs(long a) This method returns the absolute value of a long value.
5	static double acos(double a) This method returns the arc cosine of a value; the returned angle is in the range 0.0 through pi.
6	static double asin(double a) This method returns the arc sine of a value; the returned angle is in the range -pi/2 through pi/2.
7	static double atan(double a) This method returns the arc tangent of a value; the returned angle is in the range -pi/2 through pi/2.
8	static double atan2(double y, double x) This method returns the angle theta from the conversion of rectangular coordinates (x, y) to polar coordinates (r, theta).
9	static double cbrt(double a) This method returns the cube root of a double value.
10	static double ceil(double a) This method returns the smallest (closest to negative infinity) double value that is greater than or equal to the argument and is equal to a mathematical integer.
11	static double copySign(double magnitude, double sign) This method returns the first floating-point argument with the sign of the second floating-point argument.
12	static float copySign(float magnitude, float sign) This method returns the first floating-point argument with the sign of the second floating-point argument.

13	static double cos(double a) This method returns the trigonometric cosine of an angle.
14	static double cosh(double x) This method returns the hyperbolic cosine of a double value.
15	static double exp(double a) This method returns Euler's number e raised to the power of a double value.
16	static double expm1(double x) This method returns e^x -1.
17	static double floor(double a) This method returns the largest (closest to positive infinity) double value that is less than or equal to the argument and is equal to a mathematical integer.
18	static int getExponent(double d) This method returns the unbiased exponent used in the representation of a double.
19	static int getExponent(float f) This method returns the unbiased exponent used in the representation of a float.
20	static double hypot(double x, double y) This method returns sqrt(x^2 +y^2) without intermediate overflow or underflow.
21	static double IEEEremainder(double f1, double f2) This method computes the remainder operation on two arguments as prescribed by the IEEE 754 standard.
22	static double log(double a) This method returns the natural logarithm (base e) of a double value.

23	static double log10(double a) This method returns the base 10 logarithm of a double value.
24	static double log1p(double x) This method returns the natural logarithm of the sum of the argument and 1.
25	static double max(double a, double b) This method returns the greater of two double values.
26	static float max(float a, float b) This method returns the greater of two float values.
27	static int max(int a, int b) This method returns the greater of two int values.
28	static long max(long a, long b) This method returns the greater of two long values.
29	static double min(double a, double b) This method returns the smaller of two double values.
30	static float min(float a, float b) This method returns the smaller of two float values.
31	static int min(int a, int b) This method returns the smaller of two int values.
32	static long min(long a, long b) This method returns the smaller of two long values.
33	static double nextAfter(double start, double direction)

	This method returns the floating-point number adjacent to the first argument in the direction of the second argument.
34	static float nextAfter(float start, double direction) This method returns the floating-point number adjacent to the first argument in the direction of the second argument.
35	static double nextUp(double d) This method returns the floating-point value adjacent to d in the direction of positive infinity.
36	static float nextUp(float f) This method returns the floating-point value adjacent to f in the direction of positive infinity.
37	static double pow(double a, double b) This method returns the value of the first argument raised to the power of the second argument.
38	static double random() This method returns a double value with a positive sign, greater than or equal to 0.0 and less than 1.0.
39	static double rint(double a) This method returns the double value that is closest in value to the argument and is equal to a mathematical integer.
40	static long round(double a) This method returns the closest long to the argument.
41	static int round(float a) This method returns the closest int to the argument.
42	static double scalb(double d, int scaleFactor)

	This method returns d × $2^{scaleFactor}$ rounded as if performed by a single correctly rounded floating-point multiply to a member of the double value set.
43	static float scalb(float f, int scaleFactor) This method return f × $2^{scaleFactor}$ rounded as if performed by a single correctly rounded floating-point multiply to a member of the float value set.
44	static double signum(double d) This method returns the signum function of the argument; zero if the argument is zero, 1.0 if the argument is greater than zero, -1.0 if the argument is less than zero.
45	static float signum(float f) This method returns the signum function of the argument; zero if the argument is zero, 1.0f if the argument is greater than zero, -1.0f if the argument is less than zero.
46	static double sin(double a) This method returns the hyperbolic sine of a double value.
47	static double sinh(double x) This method Returns the hyperbolic sine of a double value.
48	static double sqrt(double a) This method returns the correctly rounded positive square root of a double value.
49	static double tan(double a) This method returns the trigonometric tangent of an angle.r
50	static double tanh(double x) This method returns the hyperbolic tangent of a double value.
51	static double toDegrees(double angrad) This method converts an angle measured in radians to an approximately

	equivalent angle measured in degrees.
52	static double toRadians(double angdeg) This method converts an angle measured in degrees to an approximately equivalent angle measured in radians.
53	static double ulp(double d) This method returns the size of an ulp of the argument.
54	static double ulp(float f) This method returns the size of an ulp of the argument.

Finding Greatest Common Divisor in Java, we'll look at three approaches to find the Greatest Common Divisor (GCD) of two integers. Further, we'll look at their implementation in Java. The GCD (Greatest Common Divisor) of two numbers is the largest positive integer number that divides both the numbers without leaving any remainder. For example. GCD of 30 and 45 is 15. GCD also known as HCF (Highest Common Factor). In this tutorial we will write couple of different Java programs to find out the GCD of two numbers.

Program to find GCD or HCF of two numbers, Java program to find GCD of two numbers. class Test. {. // Recursive function to return gcd of a and b. static int gcd(int a, int b). {. // Everything divides 0. if (a == 0). In mathematics, the greatest common divisor (gcd), sometimes known as the greatest common factor (gcf) or highest common factor (hcf), of two non-zero integers, is the largest positive integer that divides both numbers. The greatest common divisor of a and b is written as gcd(a, b), or sometimes simply as (a, b).

Greatest common divisor c++

C Program to Find GCD of two Numbers, The HCF or GCD of two integers is the largest integer that can exactly divide both numbers (without a remainder). There are many ways to find the greatest common divisor in C programming. There are many ways to find the greatest common divisor in C programming. Example #1: GCD Using for loop and if Statement In this program, two integers entered by the user are stored in variable n1 and n2.Then, for loop is iterated until i is less than n1 and n2.

C Program to Find G.C.D Using Recursion, In this C programming example, you will learn to find the GCD (Greatest Common Divisor) of two positive integers entered by the user using recursion. The Greatest Common Divisor (GCD) of two numbers is the largest number that divides both of them. For example: Let's say we have two numbers are 45 and 27. 45 = 5 * 3 * 3 27 = 3 * 3 * 3. So, the GCD of 45 and 27 is 9. A program to find the GCD of two numbers is given as follows.

Program to find GCD or HCF of two numbers, C program to find GCD of two numbers. #include <stdio.h>. // Recursive function to return gcd of a and b. int gcd(int a, int b). {. // Everything divides 0. if (a == 0). In mathematics, the greatest common divisor (gcd) of two or more integers, which are not all zero, is the largest positive integer that divides each of the integers. For example, the gcd of 8 and 12 is 4.

Greatest common divisor python

gcd() function Python, has to be found with the resulting remainder as zero. The greatest common divisor (GCD) of a and b is the largest number that divides both of them with no remainder. One way to find the GCD of two numbers is Euclid's algorithm, which is based on the observation that if r is the remainder when a is divided by b , then gcd(a, b) = gcd(b, r) .

gcd() in Python, gcd() in Python. The Highest Common Factor (HCF) , also called gcd, can be computed in python using a single function offered by math module and hence can Write a Python program to compute the greatest common divisor (GCD) of two positive integers. Pictorial Presentation: Sample Solution:-. Python Code: def gcd(x, y): gcd = 1 if x % y == 0: return y for k in range(int(y / 2), 0, -1): if x % k == 0 and y % k == 0: gcd = k break return gcdprint(gcd(12, 17))print(gcd(4, 6))

Code for Greatest Common Divisor in Python, It's in the standard library. >>> from fractions import gcd >>> gcd(20,8) 4. Source code from the inspect module in Python 2.7: >>> print inspect.getsource(gcd) Greatest common divisor or gcd is a mathematical expression to find the highest number which can divide both the numbers whose gcd has to be found with the resulting remainder as zero. It has many mathematical applications. Python has a inbuilt gcd function in the math module which can be used for this purpose.

Greatest common divisor tutorial

How to Find the Greatest Common Divisor by Using the Euclidian , This tutorial demonstrates how the euclidian algorithm can be used to find the greatest common Duration: 4:10 Posted: Dec 16, 2012 This tutorial demonstrates how the euclidian algorithm can be used to find the greatest common denominator of two large numbers. Learn Math Tutorials Booksto

The Greatest Common Divisor made easy, The solution to a typical GCD exam question. See my other videos https://www.youtube.com Duration: 4:18 Posted: Feb 24, 2014 Greatest Common Divisor | Euclidean Algorithm | Code Tutorial by various websites and I have tried to make some of my own changes for the sake of this tutorial. Greatest common divisor

Undergraduate Mathematics/Greatest common divisor, or HCF is simple. fo r loop continue till the value of i is equal to x or y and if condition checks whether the remainder of x and y when divided by i is equal to 0 or not. In this tutorial, we'll look at three approaches to find the Greatest Common Divisor (GCD) of two integers. Further, we'll look at their implementation in Java. 2.

Gcd geeksforgeeks

Program to find GCD or HCF of two , GCD (Greatest Common Divisor) or HCF (Highest Common Factor) of two numbers is the largest number that divides both of them. For example GCD of 20 and 28 GCD (Greatest Common Divisor) or HCF (Highest Common Factor) of two numbers is the largest number that divides both of them. For example GCD of 20 and 28 is 4 and GCD of 98 and 56 is 14. Recommended: Please solve it on “ PRACTICE ” first, before moving on to the solution.

Mathematical Algorithms, A Computer Science portal for geeks. It contains well written, well thought and well explained computer science and programming articles, quizzes and Using gcd() can compute the same gcd with just one line. math.gcd(x, y) Parameters : x : Non-negative integer whose gcd has to be computed. y : Non-negative integer whose gcd has to be computed. Return Value : This method will return an absolute/positive integer value after calculating the GCD of given parameters x and y.

An algorithm in mathematics is a procedure, a description of a set of steps that can be used to solve a mathematical computation: but they are much more common than that today. Algorithms are used in many branches of science (and everyday life for that matter), but perhaps the most common example is that step-by-step procedure used in long division.

The process of resolving a problem in such as "what is 73 divided by 3" could be described by the following algorithm:

- How many times does 3 go into 7?
- The answer is 2
- How many are left over? 1

- Put the 1(ten) in front of the 3.
- How many times does 3 go into 13?
- The answer is 4 with a remainder of one.
- And of course, the answer is 24 with a remainder of 1.

The step by step procedure described above is called a long division algorithm.

Why Algorithms?

While the description above might sound a bit detailed and fussy, algorithms are all about finding efficient ways to do the math. As the anonymous mathematician says, 'Mathematicians are lazy so they are always looking for shortcuts.' Algorithms are for finding those shortcuts.

A baseline algorithm for multiplication, for example, might be simply adding the same number over and over again. So, 3,546 times 5 could be described in four steps:

- How much is 3546 plus 3546? 7092
- How much is 7092 plus 3546? 10638
- How much is 10638 plus 3546? 14184
- How much is 14184 plus 3546? 17730

Five times 3,546 is 17,730. But 3,546 multiplied by 654 would take 653 steps. Who wants to keep adding a number over and over again? There are a set of multiplication algorithms for that; the one you choose would depend on how large your number is. An algorithm is usually the most efficient (not always) way to do the math.

Common Algebraic Examples

FOIL (First, Outside, Inside, Last) is an algorithm used in algebra that is used in multiplying polynomials: the student remembers to solve a polynomial expression in the correct order:

To resolve (4x + 6)(x + 2), the FOIL algorithm would be:

- Multiply the first terms in the parenthesis (4x times x = 4x2)
- Multiply the two terms on the outside (4x times 2 = 8x)
- Multiply the inside terms (6 times x = 6x)
- Multiply the last terms (6 times 2 = 12)
- Add all the results together to get 4x2 + 14x + 12)

BEDMAS (Brackets, Exponents, Division, Multiplication, Addition and Subtraction.) is another useful set of steps and is also considered a formula. The BEDMAS method refers to a way to order a set of mathematical operations.

Teaching Algorithms

Algorithms have an important place in any mathematics curriculum. Age-old strategies involve rote memorization of ancient algorithms; but modern teachers have also begun to develop curriculum over the years to effectively teach the idea of algorithms, that there are multiple ways of resolving complex issues by breaking them into a set of procedural steps. Allowing a child to creatively invent ways of resolving problems is known as developing algorithmic thinking.

When teachers watch students do their math, a great question to pose to them is "Can you think of a shorter way to do that?" Allowing children to create their own methods to resolve issues stretches their thinking and analytical skills.

Outside of Math

Learning how to operationalize procedures to make them more efficient is an important skill in many fields of endeavor. Computer science continually improves upon arithmetic and algebraic equations to make computers run more efficiently; but so do chefs, who continually improve their processes to make the best recipe for making a lentil soup or a pecan pie.

Other examples include online dating, where the user fills out a form about his or her preferences and characteristics, and an algorithm uses those choices to pick a perfect potential mate. Computer video games use algorithms to tell a story: the user makes a decision, and the computer bases the next steps on that decision. GPS systems use algorithms to balance readings from several satellites to identify your exact location and the best route for your SUV. Google uses an algorithm based on your searches to push appropriate advertising in your direction.

Some writers today are even calling the 21st century the Age of Algorithms. They are today a way to cope with the massive amounts of data we are generating daily.

Stein's Algorithm for finding GCD, Stein's algorithm or binary GCD algorithm is an algorithm that computes the greatest common divisor of two non-negative integers. Stein's algorithm replaces A Computer Science portal for geeks. It contains well written, well thought and well explained computer science and programming articles, quizzes and practice/competitive programming/company interview Questions.

Find the greatest common divisor of 9118 12173 and 33182

The Euclidean Algorithm for finding Greatest Common Divisor, The GCD calculator allows you to quickly find the greatest common divisor of a set of numbers. You may enter between two and ten non-zero integers between The greatest common divisor (also known as greatest common factor, highest common divisor or highest common factor) of a set of numbers is the largest positive integer number that devides all the numbers in the set without remainder. It is the biggest multiple of all numbers in the set. The GCD is most often calculated for two numbers, when it is used to reduce fractions to their lowest terms.

GCD Calculator, To find the greatest common factor of two numbers just type them in and get the To get the Greates Common Factor (GCF) of 33182 and 12173 we need to To calculate the greatest common divisor of 3 different numbers, we can use this prinicple: gcd(a, b, c) = gcd(a, gcd(b, c)) So we apply the Euclidean algorithm twice. Let's see if this works! gcd(9118, 12173, 33182) = gcd(9118, gcd(12173, 33182)) First, use the Euclidean algorithm to find the inner piece. = gcd(12173, 33182)

GCD Calculator - Greatest Common Divisor, A METHOD FOR FINDING THE GREATEST COMMON DIVISOR FOR TWO The greatest common divisor (gcd) of two integers, a and b, is the largest integer. GCD stands for Greatest Common Divisor. GCD is largest number that divides the given numbers. The GCD is sometimes called the greatest common factor (GCF). GCD Example. Find the GCD (GCF) of 45 and 54. Step 1: Find the divisors of given numbers: The divisors of 45 are : 1, 3, 5, 9, 15, 45. The divisors of 54 are : 1, 2, 3, 6, 9, 18, 27, 54

The Euclidean Algorithm

The Euclidean algorithm is one of the oldest known algorithms (it appears in Euclid’s Elements) yet it is also one of the most important, even today.

Not only is it fundamental in mathematics, but it also has important applications in computer security and cryptography.

The algorithm provides an extremely fast method to compute the greatest common divisor (gcd) of two integers.

Definition. Let a, b be two integers. A common divisor of the pair a, b is any integer d such that d | a and d | b.

Reminder: To say that d | a means that $\exists c \in Z$ such that $a = d \cdot c$. I.e., to say that d | a means that a is an integral multiple of d.

Example. The common divisors of the pair 12, 150 include ±1, ±2, ±3, ±6. These are ALL the common divisors of this pair of integers.

Question: How can we be sure there aren't any others?

• Divisors of 12 are ±1, ±2, ±3, ±4, ±6, ±12 and no others.

• Divisors of 150 are ±1, ±2, ±3, ±5, ±6, ±10, ±15, ±25, ±30, ±50, ±75, ±150 and no others.

• Now take the intersection of the two sets to get the common divisors.

Definition: The greatest common divisor (written as gcd(a, b)) of a pair a, b of integers is the biggest of the common divisors.

In other words, the greatest common divisor of the pair a, b is the maximum element of the set of common divisors of a, b.

Example: From our previous example, we know the set of common divisors of the pair 12, 150 is the set {±1, ±2, ±3, ±6}. Thus, gcd(12, 150) = 6, since 6 is the maximum element of the set.

The gcd(a, b) always exists, except in one case: gcd(0, 0) is undefined. Why?

Because any positive integer is a common divisor of the pair 0, 0 and the set of positive integers has no maximum element.

Why is the gcd defined for every other pair of integers?

• Hint: Can you prove that if at least one of the integers a, b is non-zero, then the set of common divisors has an upper bound?

• Why is that enough to prove the claim?

• How is the existence of said maximum related to the well-ordering principle, if it is? If you can't figure out the answers to these questions, then you don't understand the definitions yet!

TEST: What is gcd(a, 0) for any integer a 6= 0?

COMMENT: Rosen defines gcd(0, 0) = 0. Do you think that is reasonable? What is wrong, if anything, with allowing gcd(0, 0) to be undefined? Would defining gcd(0, 0) = ∞ be more reasonable?

The following observation means that we may as well confine our attention to pairs of non-negative integers when we study the gcd.

Lemma. For any integers a, b we have gcd(a, b) = gcd(|a|, |b|).

The proof is left as an exercise for you. Here's a hint: How does the list of divisors of a differ from that of |a|?.

At this point, we have an infallible method for computing the gcd of a given pair of numbers:

1. Find the set of positive divisors of each number. (Why is it enough to find just the positive divisors?)

2. Find the intersection of the two sets computed in the previous step.

3. The maximum element of the intersection is the desired gcd.

How efficient is this method? How long do you think it would take to compute all the positive divisors of a larger number such as a = 1092784930198374849278478587371?

For large numbers a, we would essentially be forced to try dividing by each number up to the square root of a, in the worst case. (The worst case turns out to be the case where a is prime — we will say more about primes later.) 2 Suppose that a has 200 decimal digits. Then 10199 ≤ a < 10200, so 3 · 1099 < √ a < 10100. Dividing by every number up to the square root would involve doing at least 3 · 1099 divisions.

Suppose we use a supercomputer that can do a billion (109) divisions per second. Then the number of seconds it would take the supercomputer to do all the needed divisions (in the worst case) would be at least

3 · 1099/109 = 3 · 1090 seconds.

How many seconds is that? Well, there are 60 · 60 · 24 seconds in a day, and 60 · 60 · 24 · 365 = 31536000 seconds in a year. That's roughly 3.2 · 108 seconds per year. So the number of years it would take the supercomputer to do all the needed divisions (in the worst case) would be at least

3 · 1090/(3.2 · 108) = 9.375 · 1081 years.

This is rather alarming, once you look up the age of the universe: 14.6 billion years.

CONCLUSION: It would take MUCH longer than the age of the universe for a fast supercomputer to perform that many divisions!!

Nevertheless, I can find the gcd of a pair a 200 digit numbers on my Macbook (which is NOT a supercomputer) in a couple of seconds.

THERE MUST BE A BETTER METHOD THAN MAKING LISTS OF DIVISORS!

The better method is called the Euclidean algorithm, of course. It is based on the division algorithm. Let's see how it works on a small example.

Example (Find gcd(10319, 2312)). Divide 10319 by 2312: $10319 = 4 \cdot 2312 + 1071$.

Divide 2312 by 1071: $2312 = 2 \cdot 1071 + 170$.

Divide 1071 by 170: $1071 = 6 \cdot 170 + 51$.

Divide 170 by 51: $170 = 3 \cdot 51 + 17$.

Divide 51 by 17: $51 = 3 \cdot 17 + 0$ STOP!

CONCLUSION: gcd(10319, 2312) = 17 (the last non-zero remainder).

In the example, we found the gcd with just five divisions. Try making lists of divisors of the two numbers to compute the gcd. We stopped when we did because we had to: the next step would involve division by zero!

Theorem (Euclidean algorithm). Given positive integers a, b with $a \geq b$. Put $r0 = a$ and $r1 = b$. For each $j \geq 0$, apply the division algorithm to divide rj by $rj+1$ to obtain an integer quotient $qj+1$ and remainder $rj+2$, so that:

$$rj = rj+1qj+1 + rj+2 \text{ with } 0 \leq rj+2 < rj+1.$$

This process terminates when a remainder of 0 is reached, and the last nonzero remainder in the process is gcd(a, b).

The proof requires a small lemma, which we state and prove first.

Lemma. Given integers d, e such that $e = dq + r$, where q, r are integers, we have that $\gcd(e, d) = \gcd(d, r)$.

Proof. Let c be any common divisor of the pair d, e. Then c must divide the left hand side of $e - dr = r$, so c must divide r. Thus c is a common divisor of the pair d, r.

On the other hand, let c be any common divisor of the pair d, r. Then c divides the right hand side of $e = dq + r$, so c divides e. Thus c is a common divisor of the pair d, e.

This shows that the pair e, d and the pair d, r have the same set of common divisors. It follows that the maximum is the same, too, in other words, gcd(e, d) = gcd(d, r).

Now we can prove the theorem:

Proof. By the lemma, we have that at each stage of the Euclidean algorithm, $\gcd(r_j, r_{j+1}) = \gcd(r_{j+1}, r_{j+2})$. The process in the Euclidean algorithm produces a strictly decreasing sequence of remainders $r_0 > r_1 > r_2 > \cdots > r_{n+1} = 0$. This sequence must terminate with some remainder equal to zero since as long as the remainder is positive the process could be continued.

If r_n is the last non-zero remainder in the process, then we have

$$r_n = \gcd(r_n, 0) = \gcd(r_{n-1}, r_n) = \cdots = \gcd(r_0, r_1) = \gcd(a, b).$$

Each successive pair of remainders in the process is the same. The proof is complete.

We can prove more. Let $g = \gcd(a, b) = r_n$. Solving for the remainder r_n in the last equation $r_{n-2} = r_{n-1}q_{n-1} + r_n$ with non-zero remainder gives us that

$$g = r_n = r_{n-2} - r_{n-1}q_{n-1}$$

which shows that g can be expressed as a linear combination of the two preceding remainders in the sequence of remainders. By backwards induction, this is true at each step along the way, all the way back to the pair $r_0 = a$, $r_1 = b$. For instance, since $r_{n-1} = r_{n-3} - r_{n-2}q_{n-2}$ by substituting into the above equation we get

$$g = r_{n-2} - r_{n-1}q_{n-1} = r_{n-2} - (r_{n-3} - r_{n-2}q_{n-2})q_{n-1}$$

$$= q_{n-1}r_{n-3} + (1 + q_{n-2}q_{n-1})r_{n-2},$$

which is another linear combination, as claimed.

This analysis proves the following result, and it also provides a method for finding such a linear combination.

Theorem (Bezout's theorem). Let g = gcd(a, b) where a, b are positive integers. Then there are integers x, y such that g = ax + by.

In other words, the gcd of the pair a, b is always expressible as some integral linear combination of a, b. By substituting backwards successively in the equations from the Euclidean algorithm, we can always find such a linear combination.

Example (gcd(10319, 2312) = 17 revisited). We want to find integers x, y such that 17 = 10319x+2312y. Let's recall that when we computed this gcd earlier in this lecture, we got 10319, 2312, 1071, 170, 51, 17, 0 for the sequence of remainders. So $r_0 = 10319$,

$r_1 = 2312$, $r_2 = 1071$, $r_3 = 170$, $r_4 = 51$, $r_5 = 17$, and $r_6 = 0$. The equations we got before, written in reverse order, are in the first column below, and the calculation of x, y is shown in the second column:

$$
\begin{array}{ll}
r_3 = 3r_4 + r_5 & \Rightarrow 17 = r_5 = r_3 - 3r_4 \\
r_2 = 6r_3 + r_4 & \Rightarrow 17 = r_3 - 3(r_2 - 6r_3) = -3r_2 + 19r_3 \\
r_1 = 2r_2 + r_3 & \Rightarrow 17 = -3r_2 + 19(r_1 - 2r_2) = 19r_1 - 41r_2 \\
r_0 = 4r_1 + r_2 & \Rightarrow 17 = 19r_1 - 41(r_0 - 4r_1) = -41r_0 + 183r_1.
\end{array}
$$

Remembering that $r_0 = 10319$, $r_1 = 2312$ this calculation proves that $17 = (10319)(-41) + (2312)(183)$, so $x = -41$ and $y = 183$.

Theorem. Let $g = \gcd(a, b)$ where a, b are integers, not both 0. Then g is the least positive integer which is expressible as an integral linear combination of a, b.

Proof. (Sketch) Let S be the set of all positive integers expressible in the form $ax + by$ for integers x, y. By the well-ordering principle, the set S has a least element, call it d.

Apply the division algorithm to show that $d \mid a$ and $d \mid b$. This shows that d is a common divisor of the pair a, b.

Now assume that c is any other common divisor of the pair a, b. Since d is expressible in the form $ax + by$, you can show that c must divide d. This shows that $c \le d$. It follows that d is the greatest common divisor, so $d = g$, as desired.

This theorem implies Bezout's theorem (again). It also gives a new characterization of the gcd.

The Diophantine Equation

Diophantine equations

Algebraic equations, or systems of algebraic equations with rational coefficients, the solutions of which are sought for in integers or rational numbers. It is usually assumed that the number of unknowns in Diophantine equations is larger than the number of equations; thus, they are also known as indefinite equations. In modern mathematics the concept of a Diophantine equation is also applied to algebraic equations the solutions of which are sought for in the algebraic integers of some algebraic extension of the field QQ of rational numbers, of the field of pp- adic numbers, etc.

The study of Diophantine equations is on the border-line between number theory and algebraic geometry (cf. Diophantine geometry).

Finding solutions of equations in integers is one of the oldest mathematical problems. As early as the beginning of the second millennium B.C. ancient Babylonians succeeded in solving systems of equations with two unknowns. This branch of mathematics flourished to the greatest extent in Ancient Greece. The principal source is Aritmetika by Diophantus (probably the 3rd century A.D.), which contains different types of equations and systems. In this book, Diophantus (hence the name "Diophantine equations") anticipated a number of methods for the study of equations of the second and third degrees which were only fully developed in the 19th century [1]. The creation of the theory of rational numbers by the scientists of Ancient Greece led to the study of rational solutions of indefinite equations. This point of view is systematically followed by Diophantus in his book. Even though his work contains solutions of specific Diophantine equations only, there is reason to believe that he was also familiar with a few general methods.

The study of Diophantine equations usually involves major difficulties. Moreover, it is possible to specify, explicitly, polynomials

$$F(x,y1\ldots yn)F(x,y1\ldots yn)$$

with integer coefficients such that no algorithm exists by which it would be possible to tell, for any given xx, whether the equation

$$F(x,y1\ldots yn)=0F(x,y1\ldots yn)=0$$

is solvable for $y_1 \dots y_n$ (cf. Diophantine equations, solvability problem of). Examples of such polynomials may be explicitly written down; no exhaustive description of their solutions can be given (if the Church thesis is accepted).

The simplest Diophantine equation

$$ax+by=1,$$

where a and b are relatively prime integers, has infinitely many solutions (if x_0, y_0 form a solution, then the pair of numbers $x=x_0+bn$ and $y=y_0-an$, where n is an arbitrary integer, will also be a solution). Another example of a Diophantine equation is

$$x^2+y^2=z^2.$$

Positive integral solutions of this equation represent the lengths of the small sides x, y and of the hypotenuse z of right-angled triangles with integral side lengths; these numbers are known as Pythagorean numbers. All triplets of relatively prime Pythagorean numbers are given by the formulas

$$x=m^2-n^2,\ y=2mn,\ z=m^2+n^2,$$

where m and n are relatively prime integers ($m>n>0$).

Diophantus in his Aritmetika deals with the search for rational (not necessarily integral) solutions of special types of Diophantine equations. The general theory of solving of Diophantine equations of the first degree was developed by C.G. Bachet in the 17th century; for more details on this subject see Linear equation. P. Fermat, J. Wallis, L. Euler, J.L. Lagrange, and C.F. Gauss in the early 19th century mainly studied Diophantine equations of the form

$$ax^2+bxy+cy^2+dx+ey+f=0,$$

where a, b, c, d, e, and f are integers, i.e. general inhomogeneous equations of the second degree with two unknowns. Lagrange used continued fractions in his study of general inhomogeneous Diophantine equations of the second degree with two unknowns. Gauss developed the general theory of quadratic forms, which is the basis of solving certain types of Diophantine equations.

In studies on Diophantine equations of degrees higher than two significant success was attained only in the 20th century. It was established by A. Thue that the Diophantine equation

$$a_0x^n+a_1x^{n-1}y+\cdots+a_ny^n=c,$$

where $n\geq3$, $a_0\dots a_n,c$ are integers, and the polynomial $a_0t^n+\cdots+a_n$ is irreducible in the field of rational numbers, cannot have an infinite number of integer solutions. However, Thue's method fails to yield either a bound on the solutions or on their number. A. Baker obtained effective theorems giving bounds on solutions of certain equations of this kind. B.N. Delone proposed another method of investigation, which is applicable to a narrower class of Diophantine equations, but which yields a bound for the number of solutions. In particular, Diophantine equations of the form

$$ax^3+y^3=1$$

are fully solvable by this method.

The theory of Diophantine equations has many directions. Thus, a well-known problem in this theory is Fermat's problem — the hypothesis according to which there are no non-trivial solutions of the Diophantine equation

$$x^n+y^n=z^n \quad (1)$$

if $n\geq3$. The study of integer solutions of equation (1) is a natural generalization of the problem of Pythagorean triplets. Euler obtained a positive solution of Fermat's problem for $n=4$. Owing to this result, Fermat's problem is reduced to the proof of the absence of non-zero integer solutions of equation (1) if n is an odd prime. At the time of writing (1988) the study concerned with solving (1) has not been completed. The difficulties involved in solving it are due to the fact that prime factorization in the ring of algebraic integers is not unique. The theory of divisors in rings of algebraic integers makes it possible to confirm the validity of Fermat's theorem for many classes of prime exponents n.

The arithmetic of rings of algebraic integers is also utilized in many other problems in Diophantine equations. For instance, such methods were applied in a detailed solution of an equation of the form

$$N(\alpha_1x_1+\cdots+\alpha_nx_n)=m, \quad (2)$$

where $N(\alpha)$ is the norm of the algebraic number α, and integral rational numbers $x_1\dots x_n$ which satisfy equation (2) are to be found. Equations of this class include, in particular, the Pell equation $x^2-dy^2=1$. Depending on the values of $\alpha_1\dots\alpha_n$ which appear in (2), these equations are subdivided into two types. The first type — the so-called complete forms — comprises equations in which among the α_i there are m linearly independent numbers over the field of rational numbers $\mathbb{Q}$, where $m=[\mathbb{Q}(\alpha_1\ldots\alpha_n):\mathbb{Q}]$ is the degree of the algebraic number field $\mathbb{Q}(\alpha_1\ldots\alpha_n)$ over $\mathbb{Q}$. Incomplete forms are those in which the maximum number of linearly independent numbers α_i is less than m. The case of complete forms is simpler and its study has now, in principle, been completed. It is possible, for example, to describe all solutions of any complete form [2].

The second type — the incomplete forms — is more complicated, and the development of its theory is still (1988) far from being completed. Such equations are studied with the aid of Diophantine approximations. They include the equation

$$F(x,y)=C,$$

where $F(x,y)$ is an irreducible homogeneous polynomial of degree $n\geq 3$. This equation may be written as

$$\prod_{j=1}^{n}(x-\alpha_j y)=C, \tag{3}$$

where α_j are all the roots of the polynomial $F(z,1)=0$. The existence of an infinite sequence of integral solutions of equation (3) would lead to relationships of the form

$$\left|\frac{x_i}{y_i}-\alpha_j\right|\leq\frac{C_1(F)}{y_i^n} \tag{4}$$

for some α_j. Without loss of generality, one may assume that $y_i\to\infty$. Accordingly, if y_i is sufficiently large, inequality (4) will be in contradiction with the Thue–Siegel–Roth theorem, from which follows that the equation $F(x,y)=C$, where F is an irreducible form of degree three or higher, cannot have an infinite number of solutions.

Equations such as (2) constitute a fairly narrow class among all Diophantine equations. For instance, their simple appearance notwithstanding, the equations

$$x^3+y^3+z^3=N \tag{5}$$

and

$$x^2+y^2+z^2+u^2=N \tag{6}$$

are not in this class. The study of the solutions of equation (6) is a fairly thoroughly investigated branch of Diophantine equations — the representation of numbers by quadratic forms. The Lagrange theorem states that (6) is solvable for all natural N. Any natural number not representable in the form $4^a(8k-1)$, where aa and kk are non-negative integers, can be represented as a sum of three squares (Gauss' theorem). Criteria are known for the existence of rational or integral solutions of equations of the form

F(x1…xn)=a,F(x1…xn)=a,

where F(x1…xn)F(x1…xn) is a quadratic form with integer coefficients. Thus, according to Minkowski–Hasse theorem, the equation

∑i,jaijxixj=b,∑i,jaijxixj=b,

where aijaij and bb are rational, has a rational solution if and only if it is solvable in real numbers and in pp- adic numbers for each prime number pp.

The representation of numbers by arbitrary forms of the third degree or higher has been studied to a lesser extent, because of inherent difficulties. One of the principal methods of study in the representation of numbers by forms of higher degree is the method of trigonometric sums (cf. Trigonometric sums, method of). In this method the number of solutions of the equation is explicitly written out in terms of a Fourier integral, after which the circle method is employed to express the number of solutions of the equation in terms of the number of solutions of the corresponding congruences. The method of trigonometric sums depends less than do other methods on the algebraic peculiarities of the equation.

There exists a large number of specific Diophantine equations which are solvable by elementary methods [5].

References

1. I.M. Vinogradov, "Fundamentals of number theory" , Moscow (1972) (In Russian)

2. Z.I. Borevich, I.R. Shafarevich, "Number theory" , Acad. Press (1966) (Translated from Russian) (German translation: Birkhäuser, 1966) MR0195803 Zbl 0145.04902

3. L.E. Dickson, "History of the theory of numbers" , **1** , Chelsea, reprint (1934) MR1793101 MR1720467 MR0245501 MR0245500 MR0245499 MR1520248 MR1519706 MR1519382 Zbl 60.0817.03

4. I.G. Bashmakov, "Diophant und diophantische Gleichungen" , Deutsch. Verlag Wissenschaft. (1974) (Translated from Russian)

5. W. Sierpiński, "On solution of equations in integers" , PWN (1956) (In Polish)

.

Comments

The most outstanding recent result in the study of Diophantine equations was the proof by G. Falting of the Mordell conjecture, stating that curves of genus >1>1(cf. Genus of a curve) over algebraic fields have no more than a finite number of rational points (cf. [a1]). From this result it follows, in particular, that the Fermat equation xn+yn+zn=0xn+yn+zn=0 has only a finite number of rational solutions for n>3n>3.

In the last decade there was also some progress in dealing with cubic forms (cf. Cubic form) and systems of equations consisting of pairs of quadratic forms (cf. Quadratic form). This development was based on cohomological methods that provide an obstruction to the Hasse principle. These methods were suggested by Yu.I. Manin (cf. [a2]) and are now called the Brauer–Manin obstruction to the Hasse principle. It was conjectured in [a3] that the Brauer–Manin obstruction is the only one to the Hasse principle for rational surfaces. This was verified in many cases, for example, for all cubic equations ax3+by3+cz3+dz3=0ax3+by3+cz3+dz3=0 where aa, bb, cc, dd are positive integers less than 100 ([a5]). By application of suitable hyperplane sections the problem of existence of rational solutions for cubic equations with N>4N>4 variables, or for a pair of quadratic equations with N>5N>5 variables, can be reduced to the problem for rational surfaces (cf. Rational surface) for which the existence of rational points (or, equivalently, of rational solutions for a corresponding system of equations) can be effectively verified. In particular, this method gives lower bounds for NN for which the system of two quadratic equations has solutions that are better than those obtained by the present circle method ([a4]).

Applications of transcendental number theory to Diophantine equations can be found in [a11], [a12]. Diophantine equations from the point of view of algebraic geometry are treated in [a6], [a13]. Monographs dealing specifically with Fermat's equation (cf. also Fermat great theorem) are [a8] and [a14].

Unit-II Matrix

Matrix

Matrix, a set of numbers arranged in rows and columns so as to form a rectangular array. The numbers are called the elements, or entries, of the matrix. Matrices have wide applications in engineering, physics, economics, and statistics as well as in various branches of mathematics. Historically, it was not the matrix but a certain number associated with a square array of numbers called the determinant that was first recognized. Only gradually did the idea of the matrix as an algebraic entity emerge. The term *matrix* was introduced by the 19th-century English mathematician James Sylvester, but it was his friend the mathematician Arthur Cayley who developed the algebraic aspect of matrices in two papers in the 1850s. Cayley first applied them to the study of systems of linear equations, where they are still very useful. They are also important because, as Cayley recognized, certain sets of matrices form algebraic systems in which many of the ordinary laws of arithmetic (e.g., the associative and distributive laws) are valid but in which other laws (e.g., the commutative law) are not valid. Matrices have also come to have important applications in computer graphics, where they have been used to represent rotations and other transformations of images.

If there are *m* rows and *n* columns, the matrix is said to be an "*m* by *n*" matrix, written "$m \times n$." For example, is a 2×3 matrix. A matrix with *n* rows and *n* columns is called a square matrix of order *n*. An ordinary number can be regarded as a 1×1 matrix; thus, 3 can be thought of as the matrix [3].

In a common notation, a capital letter denotes a matrix, and the corresponding small letter with a double subscript describes an element of the matrix. Thus, a_{ij} is the element in the *i*th row and *j*th column of the matrix *A*. If *A* is the 2×3 matrix shown above, then $a_{11} = 1$, $a_{12} = 3$, $a_{13} = 8$, $a_{21} = 2$, $a_{22} = -4$, and $a_{23} = 5$. Under certain conditions, matrices can be added and multiplied as individual entities, giving rise to important mathematical systems known as matrix algebras.

Matrices occur naturally in systems of simultaneous equations. In the following system for the unknowns *x* and *y*,

$$\begin{aligned} 2x + 3y &= 7 \\ 3x + 4y &= 10, \end{aligned}$$

the array of numbers

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

is a matrix whose elements are the coefficients of the unknowns. The solution of the equations depends entirely on these numbers and on their particular arrangement. If 3 and 4 were interchanged, the solution would not be the same.

Two matrices A and B are equal to one another if they possess the same number of rows and the same number of columns and if $a_{ij} = b_{ij}$ for each i and each j.
If A and B are two $m \times n$ matrices, their sum $S = A + B$ is the $m \times n$ matrix whose elements $s_{ij} = a_{ij} + b_{ij}$. That is, each element of S is equal to the sum of the elements in the corresponding positions of A and B.

A matrix A can be multiplied by an ordinary number c, which is called a scalar. The product is denoted by cA or Ac and is the matrix whose elements are ca_{ij}.

The multiplication of a matrix A by a matrix B to yield a matrix C is defined only when the number of columns of the first matrix A equals the number of rows of the second matrix B. To determine the element c_{ij}, which is in the ith row and jth column of the product, the first element in the ith row of A is multiplied by the first element in the jth column of B, the second element in the row by the second element in the column, and so on until the last element in the row is multiplied by the last element of the column; the sum of all these products gives the element c_{ij}. In symbols, for the case where A has m columns and B has m rows,

$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}$. The matrix C has as many rows as A and as many columns as B.

Unlike the multiplication of ordinary numbers a and b, in which ab always equals ba, the multiplication of matrices A and B is not commutative. It is, however, associative and distributive over addition. That is, when the operations are possible, the following equations always hold true: $A(BC) = (AB)C$, $A(B + C) = AB + AC$, and $(B + C)A = BA + CA$. If the 2×2 matrix A whose rows are (2, 3) and (4, 5) is multiplied by itself, then the product, usually written A^2, has rows (16, 21) and (28, 37).

A matrix O with all its elements 0 is called a zero matrix. A square matrix A with 1s on the main diagonal (upper left to lower right) and 0s everywhere else is called a unit matrix. It is denoted by I or I_n to show that its order is n. If B is any square matrix and I and O are the unit and zero matrices of the same order, it is always true that $B + O = O + B = B$ and $BI = IB = B$. Hence O and I behave like the 0 and 1 of ordinary arithmetic. In fact, ordinary arithmetic is the special case of matrix arithmetic in which all matrices are 1×1.

Associated with each square matrix A is a number that is known as the determinant of A, denoted det A. For example, for the 2×2 matrix

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, det $A = ad - bc$. A square matrix B is called nonsingular if det $B \neq 0$. If B is nonsingular, there is a matrix called the inverse of B, denoted B^{-1}, such that $BB^{-1} = B^{-1}B = I$. The equation $AX = B$, in which A and B are known matrices and X is an unknown matrix, can be solved uniquely if A is a nonsingular matrix, for then A^{-1} exists and both sides of the equation can be multiplied on the left by it: $A^{-1}(AX) = A^{-1}B$. Now $A^{-1}(AX) = (A^{-1}A)X = IX = X$; hence the solution is $X = A^{-1}B$. A system of m linear equations in n unknowns can always be expressed as a matrix equation $AX = B$ in which A is the $m \times n$ matrix of the coefficients of the unknowns, X is the $n \times 1$ matrix of the unknowns, and B is the $n \times 1$ matrix containing the numbers on the right-hand side of the equation.

A problem of great significance in many branches of science is the following: given a square matrix A of order n, find the $n \times 1$ matrix X, called an n-dimensional vector, such that $AX = cX$. Here c is a number called an eigenvalue, and X is called an eigenvector. The existence of an eigenvector X with eigenvalue c means that a certain transformation of space associated with the matrix A stretches space in the direction of the vector X by the factor.

Submatix

Calling Sequence

SubMatrix(A, r, c, options)
SubVector(V, i, options)

Parameters

A -Matrix

r -integer, range with integer endpoints, or list of integers and/or ranges with integer endpoints; the indices of the Matrix rows used to construct the submatrix

c -integer, range with integer endpoints, or list of integers and/or ranges with integer endpoints; the indices of the Matrix columns used to construct the submatrix

V -Vector

i -integer, range with integer endpoints, or list of integers and/or ranges with integer endpoints; the indices of the Vector elements used to construct the subvector

options-(optional); constructor options for the result object

Description

- The SubMatrix(A, r, c) function returns a Matrix created by using the entries of A that are in the intersection of the rows and columns specified by **r** and **c**. For more information regarding parameters **r** and **c**, see Matrix and Vector Entry Selection.

- The SubVector(V, i) function returns a Vector created by using the entries of V that are specified by i. The orientation of the resulting subvector is the same as the orientation of V. For more information regarding parameter i, see Matrix and Vector Entry

Selection.

- The constructor options provide additional information (readonly, shape, storage, order, datatype, and attributes) to the Matrix or Vector constructor that builds the result. These options may also be provided in the form outputoptions=[...], where [...] represents a Maple list. If a constructor option is provided in both the calling sequence directly and in an outputoptions option, the latter takes precedence (regardless of the order).

- This function is part of the LinearAlgebra package, and so it can be used in the form SubMatrix(..) only after executing the command with(LinearAlgebra). However, it can always be accessed through the long form of the command by using LinearAlgebra [SubMatrix] (..).

Types of matrices such as symmetric

Matrices are distinguished on the basis of their order, elements and certain other conditions. There are different types of matrices but the most commonly used are discussed below. Let’s find out the types of matrices in the field of mathematics.

Different types of Matrices and their forms are used for solving numerous problems. Some of them are as follows:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

1) Row Matrix

A row matrix has only one row but any number of columns. A matrix is said to be a row matrix if it has only one row. For example,

A=[−1/2√523]A=[−1/2√523]

is a row matrix of order 1 × 4. In general, $A = [a_{ij}]_{1 \times n}$ is a row matrix of order 1 × n.

2) Column Matrix

A column matrix has only one column but any number of rows. A matrix is said to be a column matrix if it has only one column. For example,

is a column matrix of order 4 × 1. In general, $B = [b_{ij}]_{m \times 1}$ is a column matrix of order m × 1.

3) Square Matrix

A square matrix has the number of columns equal to the number of rows. A matrix in which the number of rows is equal to the number of columns is said to be a square matrix. Thus an m × n matrix is said to be a square matrix if m = n and is known as a square matrix of order ‘n’. For example,

A=⎡⎣⎢3−103/2√3/2143−1⎤⎦⎥A=[3−103/2√3/2143−1]

is a square matrix of order 3. In general, $A = [a_{ij}]$ m × m is a square matrix of order m.

4) Rectangular Matrix

A matrix is said to be a rectangular matrix if the number of rows is not equal to the number of columns. For example,

A=⎡⎣⎢ 3−103/2√3/2143−17/22−5⎤⎦⎥

⎣⎦A=[3−103/2√3/2143−17/22−5]

is a matrix of the order 4 × 3

5) Diagonal matrix

A square matrix B = [bij] m × m is said to be a diagonal matrix if all its non-diagonal elements are zero, that is a matrix $B =[b_{ij}]_{m\times m}$ is said to be a diagonal matrix if $b_{ij} = 0$, when $i \neq j$. For example,

A=[4][−1002]⎡⎣⎢3000−50002⎤⎦⎥A=[4][−1002][3000−50002]

are diagonal matrices of order 1, 2, 3, respectively.

6) Scalar Matrix

A diagonal matrix is said to be a scalar matrix if all the elements in its principal diagonal are equal to some non-zero constant. A diagonal matrix is said to be a scalar matrix if its diagonal elements are equal, that is, a square matrix $B = [b_{ij}]_{n \times n}$ is said to be a scalar matrix if

- $b_{ij} = 0$, when $i \neq j$

- b_{ij} = k, when i = j, for some constant k.

For example,A=[4][−100−1]⎡⎢⎣300030003⎤⎥⎦A=[4][−100−1][300030003] are scalar matrices of order 1, 2 and 3, respectively.

7) Zero or Null Matrix

A matrix is said to be zero matrix or null matrix if all its elements are zero.
For Example,

A=[0][0000]⎡⎢⎣000000000⎤⎥⎦A=[0][0000][000000000]

are all zero matrices of the order 1, 2 and 3 respectively. We denote zero matrix by O.

8) Unit or Identity Matrix

If a square matrix has all elements 0 and each diagonal elements are non-zero, it is called identity matrix and denoted by I.
Equal Matrices: Two matrices are said to be equal if they are of the same order and if their corresponding elements are equal to the square matrix A = $[a_{ij}]_{n \times n}$ is an identity matrix if

- a_{ij} = 1 if i = j
- a_{ij} = 0 if i ≠ j

We denote the identity matrix of order n by I_n. When the order is clear from the context, we simply write it as I. For example,

A=[1][1001]⎡⎢⎣100010001⎤⎥⎦A=[1][1001][100010001]

are identity matrices of order 1, 2 and 3, respectively. Observe that a scalar matrix is an identity matrix when k = 1. But every identity matrix is clearly a scalar matrix.

9) Upper Triangular Matrix

A square matrix in which all the elements below the diagonal are zero is known as the upper triangular matrix. For example,

A=⎡⎢⎣3−57040009⎤⎥⎦A=[3−57040009]

10) Lower Triangular Matrix

A square matrix in which all the elements above the diagonal are zero is known as the upper triangular matrix. For example,

$A=\begin{bmatrix}3&0&0\\0&4&0\\-5&7&9\end{bmatrix}$

Skew Symmetric

A symmetric matrix and skew-symmetric matrix both are square matrices. But the difference between them is, the symmetric matrix is equal to its transpose whereas skew-symmetric matrix is a matrix whose transpose is equal to its negative.

If A is a symmetric matrix, then $A = A^T$ and if A is a skew-symmetric matrix then $A^T = -A$.

Also, read:

- Upper Triangular Matrix
- Diagonal Matrix
- Identity Matrix

Symmetric Matrix

To understand if a matrix is a symmetric matrix, it is very important to know about transpose of a matrix and how to find it. If we interchange rows and columns of an $m \times n$ matrix to get an $n \times m$ matrix, the new matrix is called the transpose of the given matrix. There are two possibilities for the number of rows (m) and columns (n) of a given matrix:

- If $m = n$, the matrix is square
- If $m \neq n$, the matrix is rectangular

For the second case, the transpose of a matrix can never be equal to it. This is because, for equality, the order of the matrices should be the same. Hence, the only case where the transpose of a matrix can be equal to it, is when the matrix is square. But this is only the first condition. Even if the matrix is square, its transpose may or may not be equal to it. For example:

If $A=\begin{bmatrix}1&3\\2&4\end{bmatrix}$, then $A'=\begin{bmatrix}1&2\\3&4\end{bmatrix}$. Here, we can see that $A \neq A'$.
Let us take another example.

$B=\begin{bmatrix}1&2&17\\2&5&-11\\17&-11&9\end{bmatrix}$

If we take the transpose of this matrix, we will get:

$B'=\begin{bmatrix}1&2&17\\2&5&-11\\17&-11&9\end{bmatrix}$

We see that $B = B'$. Whenever this happens for any matrix, that is whenever transpose of a matrix is equal to it, the matrix is known as a symmetric matrix. But how can we find whether a matrix is symmetric or not without finding its transpose? We know that:

If A = [aij]m×n then A’ = [aij]n×m (for all the values of i and j)
So, if for a matrix A,aij = aji (for all the values of i and j) and m = n, then its transpose is equal to itself. A symmetric matrix will hence always be square. Some examples of symmetric matrices are:
P=[1511−3]
Q=⎡⎣⎢−101125712100123572 3−10001⎤⎦⎥

Properties of Symmetric Matrix

- Addition and difference of two symmetric matrices results in symmetric matrix.
- If A and B are two symmetric matrices and they follow the commutative property, i.e. AB =BA, then the product of A and B is symmetric.
- If matrix A is symmetric then A^n is also symmetric, where n is an integer.
- If A is a symmetrix matrix then A^{-1} is also symmetric.

Skew Symmetric Matrix

A matrix can be skew symmetric only if it is square. If the transpose of a matrix is equal to the negative of itself, the matrix is said to be skew symmetric. This means that for a matrix to be skew symmetric,

A’=-A

Also, for the matrix,aji = – aij(for all the values of i and j). The diagonal elements of a skew symmetric matrix are equal to zero. This can be proved in following way:
The diagonal elements are characterized by the general formula,

aij , where i = j
If i = j, then aij = aii = ajj
If A is skew symmetric, then

$a_{ji} = – a_{ji}$

$\Rightarrow a_{ii} = – a_{ii}$

$\Rightarrow 2.a_{ii} = 0$

$\Rightarrow a_{ii} = 0$

So, $a_{ij} = 0$, when i = j (for all the values of i and j)

Some examples of skew symmetric matrices are:

P=[05−50]
Q=⎡⎣⎢0−2720−3−730⎤⎦⎥

Properties of Skew Symmetric Matrix

- When we add two skew-symmetric matrices then the resultant matrix is also skew-symmetric.

- Scalar product of skew-symmetric matrix is also a skew-symmetric matrix.
- The diagonal of skew symmetric matrix consists of zero elements and therefore the sum of elements in the main diagonals is equal to zero.
- When identity matrix is added to skew symmetric matrix then the resultant matrix is invertible.
- The determinant of skew symmetric matrix is non-negative

Determinant of Skew Symmetric Matrix

If A is a skew-symmetric matrix, which is also a square matrix, then the determinant of A should satisfy the below condition:

Det (A^T) = det (-A) = $(-1)^n$ det(A)

The inverse of skew-symmetric matrix does not exist because the determinant of it having odd order is zero and hence it is singular.

Eigenvalue of Skew Symmetric Matrix

If A is a real skew-symmetric matrix then its eigenvalue will be equal to zero. Alternatively, we can say, non-zero eigenvalues of A are non-real.

Every square matrix can be expressed in the form of sum of a symmetric and a skew symmetric matrix, uniquely. Learn various concepts in maths & science by visiting our site BYJU'S.

Hermitian

Hermitian: denoting or relating to a matrix in which those pairs of elements that are symmetrically placed with respect to the principal diagonal are complex conjugates

I have thought that Hermitian was synonymous with "real", meaning, if the matrix (A, for example) is Hermitian then that means there are no complex values in the matrix. I also believe it means the complex conjugate of the matrix is equal to the matrix like so:

$$A=A^\dagger.A=A^\dagger.$$

However, there also exist Hermitian functions (which are complex?!) and the Hermitian operator (does not have to be real). Could someone please tell me what does the word "Hermitian" mean and what are the differences between the three: Hermitian matrix, Hermitian function, and Hermitian Operator? I am confused.

(PS: Please feel free to correct me if I have tagged this question incorrectly.)

Skew Hermitian

Skew-Hermitian matrices can be understood as the complex versions of real skew-symmetric matrices, or as the matrix analogue of the purely imaginary numbers.[2] The set of all skew-Hermitian matrices forms the Lie algebra, which corresponds to the Lie group U(*n*). The concept can be generalized to include linear transformations of any complex vector space with a sesquilinear norm.

Note that the adjoint of an operator depends on the scalar product considered on the dimensional complex or real space . If denotes the scalar product on , then saying is skew-adjoint means that for all one has Imaginary numbers can be thought of as skew adjoint (since they are like matrices), whereas real numbers correspond to self-adjoint operators.

Nilpotent

An element aa of a ring or semi-group with zero AA such that an=0an=0 for some natural number nn. The smallest such nn is called the nilpotency index of aa. For example, in the residue ring modulo pnpn(under multiplication), where pp is a prime number, the residue class of pp is nilpotent of index nn; in the ring of (2×2)(2×2)-matrices with coefficients in a field KK the matrix

$$\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}$$

is nilpotent of index 2; in the group algebra Fp[G]Fp[G], where FpFp is the field with pp elements and GG the cyclic group of order pp generated by σσ, the element 1−σ1−σ is nilpotent of index p.
If aa is a nilpotent element of index nn, then

1=(1−a)(1+a+⋯+an−1),1=(1−a)(1+a+⋯+an−1),

that is, (1−a)(1−a) is invertible in AA and its inverse can be written as a polynomial in aa.
In a commutative ring AA an element aa is nilpotent if and only if it is contained in all prime ideals of the ring. All nilpotent elements form an ideal JJ, the so-called nil radical of the ring; it coincides with the intersection of all prime ideals of AA. The ring A/JA/J has no non-zero nilpotent elements.

In the interpretation of a commutative ring AA as the ring of functions on the space SpecASpecA(the spectrum of AA, cf. Spectrum of a ring), the nilpotent elements correspond to functions that vanish identically. Nevertheless, the consideration of nilpotent elements frequently turns out to be useful in algebraic geometry because it makes it possible to obtain purely algebraic analogues of a number of concepts in analysis and differential geometry (infinitesimal deformations, etc.).

Involutary

In mathematics, an **involution**, or an **involutory function**, is a function f that is its own inverse,

$$f(f(x)) = x$$

for all x in the domain of f.[1] Equivalently, applying f twice produces the original value.

The term **anti-involution** refers to involutions based
n antihomomorphisms (see § Quaternion algebra, groups, semigroups below)

$$f(xy) = f(y)\, f(x)$$

such that

$$xy = f(f(xy)) = f(\, f(y)\, f(x)\,) = f(f(x))\, f(f(y)) = xy.$$

A simple example of an involution of the three-dimensional Euclidean space is reflection through a plane. Performing a reflection twice brings a point back to its original coordinates.

Another involution is reflection through the origin; not a reflection in the above sense, and so, a distinct example.

These transformations are examples of affine involutions.

Orthogonal

"Orthogonal" redirects here. For the trilogy of novels by Greg Egan, see Orthogonal (novel). For software design concept, see Orthogonality (programming).
In mathematics, orthogonality is the generalization of the notion of perpendicularity to the linear algebra of bilinear forms. Two elements u and v of a vector space with bilinear form B are orthogonal when B(u, v) = 0. Depending on the bilinear form, the vector space may contain nonzero self-orthogonal vectors. In the case of function spaces, families of orthogonal functions are used to form a basis.

By extension, orthogonality is also used to refer to the separation of specific features of a system. The term also has specialized meanings in other fields including art and chemistry.

Singular and Non singular matrices

In this article, we will discuss singular matrix and non-singular matrix. The matrix is used in different variants in the mathematical calculations. They are used to make the complex calculations simpler. The application and multiplication of matrices help us understand the properties of the matrices. Matrices are the result of binary calculations, and they are used to solve complex calculations always by using some rules. A matrix always consists of rows, columns, and they are always added and multiplied to get a definite result. It can be a 2*2 matrix or a 3*3 matrix as well.

The matrix is the ordered arrangement of rectangular array of functions or the numbers that are written in between the square brackets. In the matrix, row and column include the values or the expressions that are called elements or entries. Here the total number of rows by the number of columns describes the size or dimension of a matrix. This can be better represented in the pictorial diagram.

The matrices are classified into different types. They are being classified as a row matrix, column matrix, identity matrix, square matrix, rectangular matrix. In this process, the matrices are being able to identify the same as well.

When we speak about the matrices, there are various ways in which matrices are being represented. We will just like to give a pictorial representation of the same for better memorization.

Addition and subtraction of matrices

Before going into matrix addition, let us have a brief idea of what are matrices. In mathematics, a matrix is a rectangular array of numbers, expression or symbols, arranged in rows and columns. Horizontal Rows are denoted by “m” whereas the Vertical Columns are denoted by “n.” Thus a matrix (m x n) has m and n numbers of rows and columns respectively. We also know about different types of matrices Such as square matrix, row matrix, null matrix, diagonal matrix, scalar matrix, identity matrix, diagonal matrix, triangular matrix, etc. Now, let us now focus on how to perform the basic operation on matrices such as matrix addition and subtraction with examples.

By recalling the small concept of addition of algebraic expressions, we know that while the addition of algebraic expressions can only be done with the corresponding like terms, similarly the addition of two matrices can be done by addition of corresponding terms in the matrix.

There are basically two criteria which define the addition of matrix. They are as follows:

1. Consider two matrices A & B. These matrices can be added iff(if and only if) the order of the matrices are equal, i.e. the two matrices have the same number of rows and columns. For example, say matrix A is of the order 3×4, then the matrix B can be added to matrix A if the order of B is also 3×4.

2. The addition of matrices is not defined for matrices of different sizes.

Matrix subtraction is exactly the same as matrix addition. All the constraints valid for addition are also valid for matrix subtraction. Matrix subtraction can only be done when the two matrices are of the same size. Subtraction cannot be defined for matrices of different sizes. Mathematically,

$P–Q=P+(−Q)$

In other words, it can be said that matrix subtraction is an addition of the inverse of a matrix to the given matrix, i.e. if matrix Q has to be subtracted from matrix P, then we will take the inverse of matrix Q and add it to matrix P.

Rank of matrices

The rank of a matrix is the dimension of the subspace spanned by its rows. As we will prove in Chapter 15, the dimension of the column space is equal to the rank. This has important consequences; for instance, if A is an $m \times n$ matrix and $m \geq n$, then rank (A) ≤ n, but if $m < n$, then rank (A) ≤ m. It follows that if a matrix is not square, either its columns or its rows must be linearly dependent.

For small square matrices, perform row elimination in order to obtain an upper-triangular matrix. If a row of zeros occurs, the rank of the matrix is less than n, and it is singular. As we will see in Chapters 7, 15, and 23, finding the rank of an arbitrary matrix is somewhat complex and relies on the computation of what are termed its singular values.

For any $m \times n$ matrix, rank (A) + nullity (A) = n. Thus, if A is $n \times n$, then for A to be nonsingular, nullity (A) must be zero.

Matrix Equation

Conventional digital computers can execute advanced operations by a sequence of elementary Boolean functions of 2 or more bits. As a result, complicated tasks such as solving a linear system or solving a differential equation require a large number of computing steps and an extensive use of memory units to store individual bits. To accelerate the execution of such advanced tasks, in-memory computing with resistive memories provides a promising avenue, thanks to analog data storage and physical computation in the memory. Here, we show that a cross-point array of resistive memory devices can directly solve a system of linear equations, or find the matrix eigenvectors. These operations are completed in just one single step, thanks to the physical computing with Ohm’s and Kirchhoff’s laws, and thanks to the negative feedback connection in the cross-point circuit. Algebraic problems are demonstrated in hardware and applied to classical computing tasks, such as ranking webpages and solving the Schrödinger equation in one step.

Linear algebra problems, such as solving systems of linear equations and computing matrix eigenvectors, lie at the heart of modern scientific computing and data-intensive tasks. Traditionally, these problems in forms of matrix equations are solved by matrix factorizations or iterative matrix multiplications (1, 2), which are computationally expensive with polynomial time complexity, e.g., $O(N^3)$ where N is the size of the problem. As conventional computers are increasingly challenged by the scaling limits of the complementary metal-oxide-semiconductor (CMOS) technology (3), and by the energy and latency burdens of moving data between the memory and the computing units (4), improving the computing performance with increasing hardware resources becomes difficult and noneconomic. To get around these fundamental limits, in-memory computing has recently emerged as a promising technique to conduct computing in situ, i.e., within the memory unit (5). One example is computing within cross-point arrays, which can accelerate matrix-vector multiplication (MVM) by Ohm's law and Kirchhoff's law with analog and reconfigurable resistive memories (5⇓⇓–8). In-memory MVM has been adopted for several tasks, including image compression (5), sparse coding (6), and the training of deep neural networks (7, 8). However, solving matrix equations, such as a linear system $Ax = b$, in a single operation remains an open challenge. Here, we show that a feedback circuit including a reconfigurable cross-point resistive array can provide the solution to algebraic problems such as systems of linear equations, matrix eigenvectors, and differential equations in just one step.

Resistive memories are two-terminal elements that can change their conductance in response to applied voltage stimuli (9, 10). Owing to their nonvolatile and reconfigurable behavior, resistive memories have been widely investigated and developed for storage-class memory (11, 12), stateful logic (13⇓–15), in-memory computing (5, 6, 16, 17), and neuromorphic computing applications (7, 8, 18, 19). Resistive memories include various device concepts, such as resistive switching memory (RRAM, refs. 9⇓⇓–12), phase-change memory (PCM, ref. 20), and spin-transfer torque magnetic memory (21). Implemented in the cross-point array architecture, resistive memories can naturally accelerate data-intensive operations with enhanced time/energy efficiencies compared with classical digital computing (5, 6, 17). It has also been shown recently that iterated MVM operations with resistive cross-point arrays can solve systems of linear equations, in combination with digital floating-point computers (22). The higher the desired accuracy of the solution, the more iterations are needed to complete the operation. However, iteration raises a fundamental limit toward achieving high computing performance in terms of energy and latency.

Solution by Cramer's rule and Gauss Elimination method

We have learned how to solve systems of equations in two variables and three variables, and by multiple methods: substitution, addition, Gaussian elimination, using the inverse of a matrix, and graphing. Some of these methods are easier to apply than others and are more appropriate in certain situations. In this section, we will study two more strategies for solving systems of equations.

Evaluating the Determinant of a 2×2 Matrix

A determinant is a real number that can be very useful in mathematics because it has multiple applications, such as calculating area, volume, and other quantities. Here, we will use determinants to reveal whether a matrix is invertible by using the entries of a square matrix to determine whether there is a solution to the system of equations. Perhaps one of the more interesting applications, however, is their use in cryptography. Secure signals or messages are sometimes sent encoded in a matrix. The data can only be decrypted with an invertible matrix and the determinant. For our purposes, we focus on the determinant as an indication of the invertibility of the matrix. Calculating the determinant of a matrix involves following the specific patterns that are outlined in this section.

Find the Determinant of a 2 × 2 Matrix

We will now introduce a final method for solving systems of equations that uses determinants. Known as Cramer's Rule, this technique dates back to the middle of the 18th century and is named for its innovator, the Swiss mathematician Gabriel Cramer (1704-1752), who introduced it in 1750 in Introduction à l'Analyse des lignes Courbes algébriques. Cramer's Rule is a viable and efficient method for finding solutions to systems with an arbitrary number of unknowns, provided that we have the same number of equations as unknowns.

Cramer's Rule will give us the unique solution to a system of equations, if it exists. However, if the system has no solution or an infinite number of solutions, this will be indicated by a determinant of zero. To find out if the system is inconsistent or dependent, another method, such as elimination, will have to be used.

To understand Cramer's Rule, let's look closely at how we solve systems of linear equations using basic row operations. Consider a system of two equations in two variables.

We eliminate one variable using row operations and solve for the other. Say that we wish to solve for If equation (2) is multiplied by the opposite of the coefficient of in equation (1), equation (1) is multiplied by the coefficient of in equation (2), and we add the two equations, the variable will be eliminated.

Unit-III Vectors

Vectors

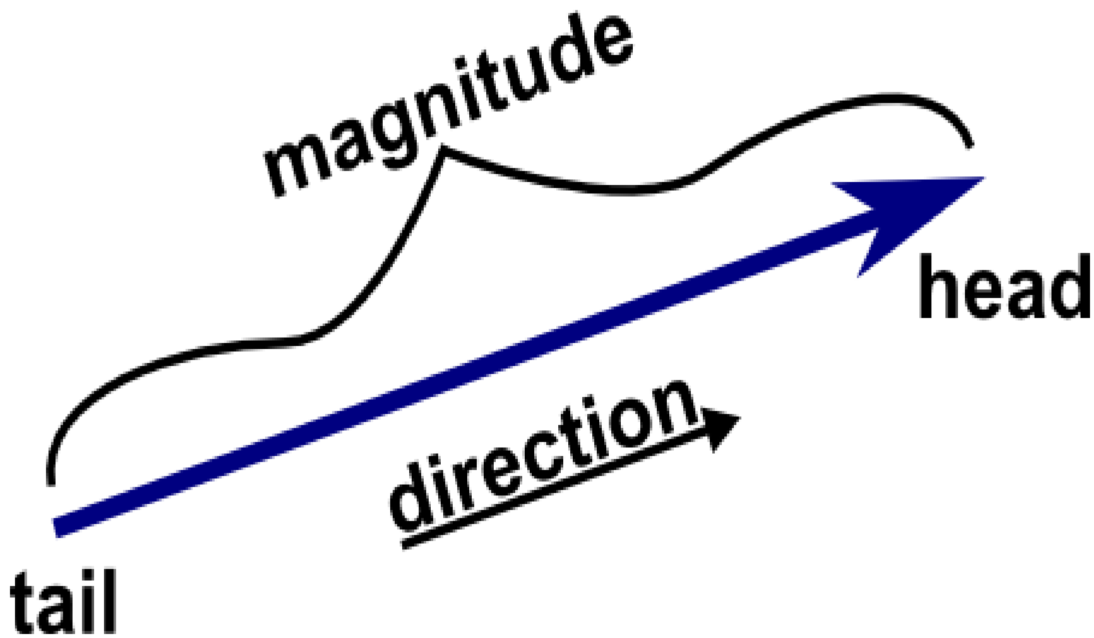
A vector is an object that has both a magnitude and a direction. Geometrically, we can picture a vector as a directed line segment, whose length is the magnitude of the vector and with an arrow indicating the direction. The direction of the vector is from its tail to its head.

Two vectors are the same if they have the same magnitude and direction. This means that if we take a vector and translate it to a new position (without rotating it), then the vector we obtain at the end of this process is the same vector we had in the beginning.

Two examples of vectors are those that represent force and velocity. Both force and velocity are in a particular direction. The magnitude of the vector would indicate the strength of the force or the speed associated with the velocity.
We denote vectors using boldface as in aa or bb. Especially when writing by hand where one cannot easily write in boldface, people will sometimes denote vectors using arrows as in $\vec{a}$ a→ or $\vec{b}$ b→, or they use other markings. We won't need to use arrows here. We denote the magnitude of the vector aa by ‖a‖‖a‖. When we want to refer to a number and stress that it is not a vector, we can call the number a scalar. We will denote scalars with italics, as in aa or bb.
You can explore the concept of the magnitude and direction of a vector using the below applet. Note that moving the vector around doesn't change the vector, as the position of the vector doesn't affect the magnitude or the direction. But if you stretch or turn the vector by moving just its head or its tail, the magnitude or direction will change. (This applet also shows the coordinates of the vector, which you can read about in another page.)

Definition of a vector

A vector is an object that has both a magnitude and a direction. Geometrically, we can picture a vector as a directed line segment, whose length is the magnitude of the vector and with an arrow indicating the direction. The direction of the vector is from its tail to its head.



Two vectors are the same if they have the same magnitude and direction. This means that if we take a vector and translate it to a new position (without rotating it), then the vector we obtain at the end of this process is the same vector we had in the beginning.

Two examples of vectors are those that represent force and velocity. Both force and velocity are in a particular direction. The magnitude of the vector would indicate the strength of the force or the speed associated with the velocity.
We denote vectors using boldface as in aa or bb. Especially when writing by hand where one cannot easily write in boldface, people will sometimes denote vectors using arrows as in $\vec{a}$ a→ or $\vec{b}$ b→, or they use other markings. We won't need to use arrows here. We denote the magnitude of the vector aa by ‖a‖‖a‖. When we want to refer to a number and stress that it is not a vector, we can call the number a scalar. We will denote scalars with italics, as in aa or bb.

You can explore the concept of the magnitude and direction of a vector using the below applet. Note that moving the vector around doesn't change the vector, as the position of the vector doesn't affect the magnitude or the direction. But if you stretch or turn the vector by moving just its head or its tail, the magnitude or direction will change. (This applet also shows the coordinates of the vector, which you can read about in another page.)

The magnitude and direction of a vector. The blue arrow represents a vector aa. The two defining properties of a vector, magnitude and direction, are illustrated by a red bar and a green arrow, respectively. The length of the red bar is the magnitude ‖a‖‖a‖ of the vector aa. The green arrow always has length one, but its direction is the direction of the vector aa. The one exception is when aa is the zero vector (the only vector with zero magnitude), for which the direction is not defined. You can change either end of aa by dragging it with your mouse. You can also move aa by dragging the middle of the vector; however, changing the position of the aa in this way does not change the vector, as its magnitude and direction remain unchanged.

Vector algebra

Vector algebra is one of the essential topics of algebra. It studies the algebra of vector quantities. As we know, there are two types of physical quantities, scalars and vectors. The scalar quantity has only magnitude, whereas the vector quantity has both magnitude and direction. Learn about Magnitude Of A Vector here.

Algebra is a significant subject in Maths where we use universal symbols or letters to signify the quantities, numbers and variables. These symbols are later used in many expressions, equations and formulae, to perform algebraic operations. It has many branches.

In essence, vector algebra is an algebra where the essential elements usually denote vectors. We perform algebraic operations on vectors and vector spaces. This branch has rules and hypotheses based on the properties and behaviour of vectors. Here, you will learn various concepts based on the basics of vector algebra and some solved examples.

Definition

A vector is an object which has both magnitudes and direction. It is usually represented by an arrow which shows the direction(→) and its length shows the magnitude. The arrow which indicates the vector has an arrowhead and its opposite end is the tail. It is denoted as $\vec{V}$. The magnitude of the vector is represented as |V|. Two vectors are said to be equal if they have equal magnitudes and equal direction.

Vector Algebra Operations

Just like in usual Algebra, we also perform arithmetic operations such as addition, subtraction, multiplication on vectors. However, in the case of multiplication, vectors have two terminologies, such as dot product and cross product.

Addition of Vectors

Let us consider there are two vectors P and Q, then the sum of these two vectors can be performed when the tail of vector Q meets with the head of vector A. And during this addition, the magnitude and direction of the vectors should not change. The vector addition follows two important laws, which are;

- Commutative Law: P + Q = Q + P
- Associative Law: P + (Q + R) = (P + Q) + R

Subtraction Of Vectors

Here, the direction of other vectors is reversed and then the addition is performed on both the given vectors. If P and Q are the vectors, for which the subtraction method has to be performed, then we invert the direction of another vector say for Q, make it -Q. Now, we need to add vector P and -Q. Thus, the direction of the vectors are opposite each other, but the magnitude remains the same.

- P – Q = P + (-Q)

Multiplication of Vectors

If k is a scalar quantity and it is multiplied by a vector A, then the scalar multiplication is given by kA. If k is positive then the direction of the vector kA is the same as vector A, but if the value of k is negative, then the direction of vector kA will be opposite to the direction of vector A. And the magnitude of the vector kA is given by |kA|.

Dot Product

The dot product is often called a scalar product. It is represented using a dot(.) between two vectors. Here, two coordinate vectors of equal length are multiplied in such a way that they result in a single number. So basically when we take the scalar product of two vectors, the result is either a number of a scalar quantity. Suppose P and Q are two vectors, then the dot product for both the vectors is given by;

- P.Q = |P| |Q| cos θ

If P and Q are both in the same direction, i.e. θ = 0°, then;

- P.Q = |P| |Q|

If P and Q are both orthogonal, i.e. θ = 90°, then;

- P.Q = 0 [since cos 90° = 0]

In vector algebra, if two vectors are given as;

$P = [P_1, P_2, P_3, P_4, \ldots, P_n]$ and $Q = [Q_1, Q_2, Q_3, Q_4, \ldots, Q_n]$

Then their dot product is given by;

- $P.Q = P_1Q_1+P_2Q_2+P_3Q_3+\ldots\ldots P_nQ_n$

Addition and Subtraction of Vectors
ectors have both magnitude and direction, one cannot simply add two vectors to obtain their sum. The addition of vectors is not as straightforward as the addition of scalars. To better understand this, let us consider an example of a car travelling 10 miles North and 10 miles South. Here, the total distance travelled is 20 miles but the displacement is zero. The North and South displacements are each vector quantities, and the opposite directions cause the individual displacements to cancel each other out. In this article, let us explore ways to carry out the addition and subtraction of vectors.

Vector Addition: Triangle and Parallelogram Law of Vectors

As already discussed, vectors cannot be simply added algebraically. Following are a few points to remember while adding vectors:

- Vectors are added geometrically and not algebraically.
- Vectors whose resultant have to be calculated behave independently of each other.
- Vector Addition is nothing but finding the resultant of a number of vectors acting on a body.
- Vector Addition is commutative. This means that the resultant vector is independent of the order of vectors.

Triangle Law of Vector Addition

The vector addition is done based on the Triangle law. Let us see what triangle law of vector addition is:

Suppose there are two vectors: a→ and b→
Now, draw a line AB representing a→ with A as the tail and B as the head. Draw another line BC representing (b→) with B as the tail and C as the head. Now join the line AC with A as the tail and C as the head. The line AC represents the resultant sum of the vectors a→ and b→
The line AC represents a→ + b→
The magnitude of a→ + b→ is:
a2 + b2 + 2ab cos θ−−−−−−−−−−−−−−−−−√
Where,

a = magnitude of vector a→
b = magnitude of vector b→
θ = angle between a→ and b→
Let the resultant make an angle of ϕ with a→, then:
tanϕ = b sin θa + b cos θ
Let us understand this by the means of an example. Suppose there are two vectors having equal magnitude A, and they make an angle θ with each other. Now, to find the magnitude and direction of the resultant, we will use the formulas mentioned above.
Let the magnitude of the resultant vector be B

B = A2 + A2 + 2AA cos θ−−−−−−−−−−−−−−−−−−√ = 2 A cos θ2
Let’s say that the resultant vector makes an angle Ɵ with the first vector
tan ϕ = A sin θA + A cos θ = tan θ2
Or,

Ɵ = θ2

Parallelogram Law of Vector Addition

The vector addition may also be understood by the law of parallelogram. The law states that “If two vectors acting simultaneously at a point are represented in magnitude and direction by the two sides of a parallelogram drawn from a point, their resultant is given in magnitude and direction by the diagonal of the parallelogram passing through that point.”

According to this law, if two vectors →P and →Q are represented by two adjacent sides of a parallelogram both pointing outwards as shown in the figure below, then the diagonal drawn through the intersection of the two vectors represent the resultant.

Scalar and vector product of two vectors

A vector can be multiplied by another vector but may not be divided by another vector. There are two kinds of products of vectors used broadly in physics and engineering. One kind of multiplication is a *scalar multiplication of two vectors*. Taking a scalar product of two vectors results in a number (a scalar), as its name indicates. Scalar products are used to define work and energy relations. For example, the work that a force (a vector) performs on an object while causing its displacement (a vector) is defined as a scalar product of the force vector with the displacement vector. A quite different kind of multiplication is a *vector multiplication of vectors*. Taking a vector product of two vectors returns as a result a vector, as its name suggests. Vector products are used to define other derived vector quantities. For example, in describing rotations, a vector quantity called *torque* is defined as a vector product of an applied force (a vector) and its distance from pivot to force (a vector). It is important to distinguish between these two kinds of vector multiplications because the scalar product is a scalar quantity and a vector product is a vector quantity.

Simple application of vectors

In the Vectors episode of NBC Learn's "The Science of NFL Football" you see that quarterbacks must account for their own motion when throwing a pass, and that both the player's movement and the path of the ball can be represented by arrows known as vectors.

Vectors are used in science to describe anything that has both a direction and a magnitude. They are usually drawn as pointed arrows, the length of which represents the vector's magnitude. A quarterback's pass is a good example, because it has a direction (usually somewhere downfield) and a magnitude (how hard the ball is thrown).

Off the field, vectors can be used to represent any number of physical objects or phenomena. Wind, for instance, is a vectorial quantity, because at any given location it has a direction (such as northeast) and a magnitude (say, 45 kilometers per hour). You could make a map of airflow at any point in time, then, by drawing wind vectors for a number of different geographic locations.

Many properties of moving objects are also vectors. Take, for instance, a billiard ball rolling across a table. The ball's velocity vector describes its movement—the direction of the vector arrow marks the ball's direction of motion, and the length of the vector represents the speed of the ball.

The billiard ball's momentum is also a vectorial quantity, because momentum is equal to mass times velocity. Therefore, the ball's momentum vector points in the same direction as its velocity vector, and the momentum vector's magnitude, or length, is the multiplication product of the ball's speed and its mass.

Momentum vectors are useful when you want to predict what will happen when two objects come into contact. Recall from the video that vectors can be added together by joining them to make a shape called a parallelogram and finding the diagonal of that parallelogram. The diagonal is the sum of the two vectors that form the sides of the parallelogram.

Let's say that a rolling billiard ball is moving toward a glancing collision with a stationary billiard ball. On impact, the moving ball transfers some of its momentum to the stationary ball, and both roll away from the collision in different directions. Following the impact, both balls have velocity and hence momentum. In fact, the sum of the momentum vectors of the two balls after the collision is equal to the first ball's momentum vector before the collision, ignoring small losses due to friction as well as sound and heat energy produced during the impact.

So, with an understanding of vectors, billiards players can predict where both balls will go following a collision, allowing them to sink more target balls while keeping the cue ball safely on the table.

Vectors are geometric representations of magnitude and direction which are often represented by straight arrows, starting at one point on a coordinate axis and ending at a different point. All vectors have a length, called the magnitude, which represents some quality of interest so that the vector may be compared to another vector. Vectors, being arrows, also have a direction. This differentiates them from scalars, which are mere numbers without a direction.

A vector is defined by its magnitude and its orientation with respect to a set of coordinates. It is often useful in analyzing vectors to break them into their component parts. For two-dimensional vectors, these components are horizontal and vertical. For three dimensional vectors, the magnitude component is the same, but the direction component is expressed in terms of xx, yy and zz.

To visualize the process of decomposing a vector into its components, begin by drawing the vector from the origin of a set of coordinates. Next, draw a straight line from the origin along the x-axis until the line is even with the tip of the original vector. This is the horizontal component of the vector. To find the vertical component, draw a line straight up from the end of the horizontal vector until you reach the tip of the original vector. You should find you have a right triangle such that the original vector is the hypotenuse.

Decomposing a vector into horizontal and vertical components is a very useful technique in understanding physics problems. Whenever you see motion at an angle, you should think of it as moving horizontally and vertically at the same time. Simplifying vectors in this way can speed calculations and help to keep track of the motion of objects.

Physical quantities can usually be placed into two categories, vectors and scalars. These two categories are typified by what information they require. Vectors require two pieces of information: the magnitude and direction. In contrast, scalars require only the magnitude. Scalars can be thought of as numbers, whereas vectors must be thought of more like arrows pointing in a specific direction.

Vectors require both a magnitude and a direction. The magnitude of a vector is a number for comparing one vector to another. In the geometric interpretation of a vector the vector is represented by an arrow. The arrow has two parts that define it. The two parts are its length which represents the magnitude and its direction with respect to some set of coordinate axes. The greater the magnitude, the longer the arrow. Physical concepts such as displacement, velocity, and acceleration are all examples of quantities that can be represented by vectors. Each of these quantities has both a magnitude (how far or how fast) and a direction. In order to specify a direction, there must be something to which the direction is relative. Typically this reference point is a set of coordinate axes like the x-y plane.

Scalars differ from vectors in that they do not have a direction. Scalars are used primarily to represent physical quantities for which a direction does not make sense. Some examples of these are: mass, height, length, volume, and area. Talking about the direction of these quantities has no meaning and so they cannot be expressed as vectors.

One of the ways in which representing physical quantities as vectors makes analysis easier is the ease with which vectors may be added to one another. Since vectors are graphical visualizations, addition and subtraction of vectors can be done graphically.

The graphical method of vector addition is also known as the head-to-tail method. To start, draw a set of coordinate axes. Next, draw out the first vector with its tail (base) at the origin of the coordinate axes. For vector addition it does not matter which vector you draw first since addition is commutative, but for subtraction ensure that the vector you draw first is the one you are subtracting *from*. The next step is to take the next vector and draw it such that its tail starts at the previous vector's head (the arrow side). Continue to place each vector at the head of the preceding one until all the vectors you wish to add are joined together. Finally, draw a straight line from the origin to the head of the final vector in the chain. This new line is the vector result of adding those vectors together.

Unit-IV Differentiation

Differentiation of Functions as polynomials

Polynomials are some of the simplest functions we use. We need to know the derivatives of polynomials such as x^4+3x, $8x^2+3x+6$, and 2. Let's start with the easiest of these, the function $y=f(x)=c$, where c is any constant, such as 2, 15.4, or one million and four (10^6+4). It turns out that the derivative of any constant function is zero. This makes sense if you think about the derivative as the slope of a tangent line. To use the definition of a derivative, with $f(x)=c$,

$$\begin{aligned}\frac{d}{dx}(c) &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\\ &= \lim_{h\to 0}\frac{c-c}{h}\\ &= \lim_{h\to 0}\frac{0}{h}\\ &= 0.\end{aligned}$$

For completeness, now consider $y=f(x)=x$. This is the equation of a straight line with slope 1, and we expect to find this from the definition of the derivative. We are not disappointed:

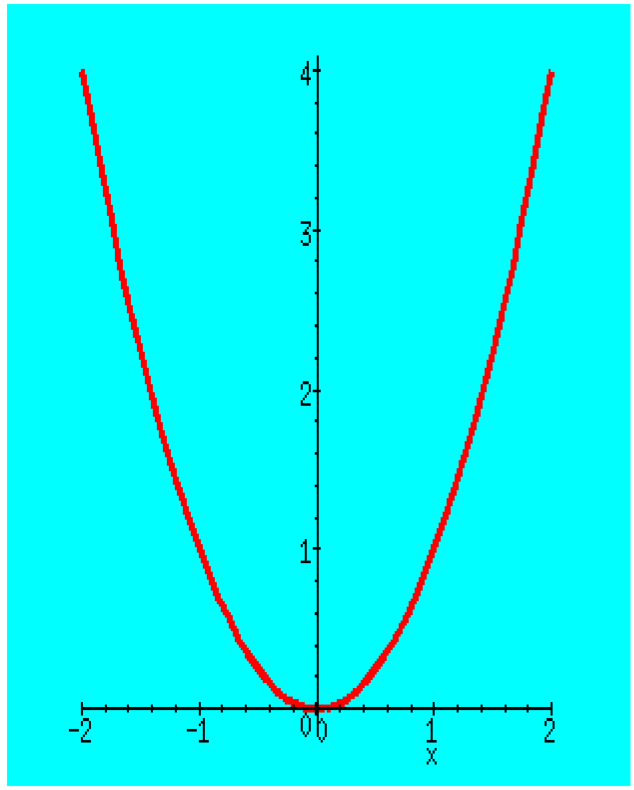
$$\begin{aligned}\frac{dy}{dx} = \frac{dx}{dx} &= \lim_{h\to 0}\frac{f(x+h)-x}{h}\\ &= \frac{(x+h)-x}{h}\\ &= \frac{h}{h}\\ &= 1.\end{aligned}$$

Two things to note in the above:

- It may be tempting to ``cancel'' the term ``*dx*'' in the intermediate step. This is valid, but only in this simple case.

- It will never be as easy as this again, although it won't be much harder.

  Before going to the most general case, consider $y=f(x)=x^2$. This is the most basic parabola, as shown. The derivative of $f(x)$ may still be found from basic algebra:

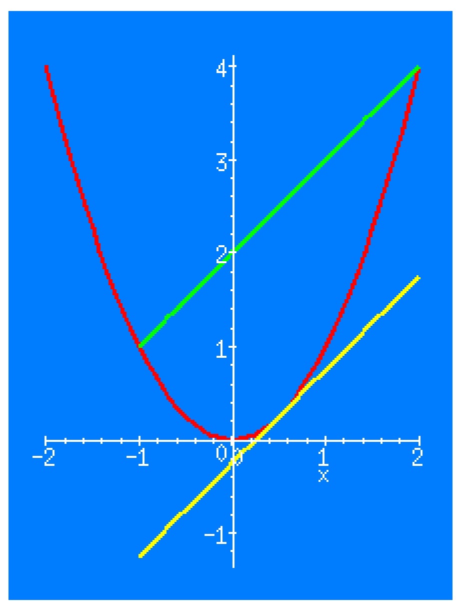


$$
\begin{aligned}
\frac{dy}{dx} = \frac{dx}{dx} &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\\
&= \frac{(x+h)^2-x^2}{h}\\
&= \lim_{h\to 0}\frac{(x^2+2\,x\,h+h^2)-x^2}{h}\\
&= \lim_{h\to 0}\frac{2\,x\,h+h^2}{h}\\
&= \lim_{h\to 0}(2\,x+h)\\
&= 2\,x.
\end{aligned}
$$

This tells us exactly what we expect; the derivative is zero at $x=0$, has the same sign as x, and becomes steeper (more negative or positive) as x becomes more negative or positive.

An interesting result of finding this derivative is that the slope of the secant line is the slope of the function at the midpoint of the interval. Specifically,

$$\frac{\Delta y}{\Delta x} = 2x + h = 2\left(x+\frac{h}{2}\right) = f'\left(x+\frac{h}{2}\right).$$



(In the figure shown, x = -1 and h = 3, so (x+h/2) = +1/2.
Please note that parabolic functions are the *only* functions (aside from linear or constant functions) for which this is always true.

From here, we can and should consider $y=f(x)=x^n$ for any positive integer n. There are many ways to do this, with varying degrees of formality.

To start, consider that for n a positive integer, the binomial theorem allows us to express $f(x+h)$ as

$$\begin{aligned} f(x+h) &= (x+h)^n \\ &= x^n + n\,x^{n-1}\,h + (n(n-1)/2)\,x^{n-2}\,h^2 + \cdots \\ &\quad + n\,x\,h^{n-1} + h^n. \end{aligned}$$

(In the above, there will always be no more than n+1 nonzero terms.) Then, algebra again gives us

$$\begin{aligned} \frac{d}{dx}x^n &= \lim_{h\to 0}\frac{\left(x^n + n\,x^{n-1}\,h + \cdots + +h^n\right) - x^n}{h} \\ &= \lim_{h\to 0}\left(n\,x^{n-1}\cdots + h^{n-1}\right) \\ &= n\,x^{n-1}. \end{aligned}$$

This very convenient form is seen to reproduce the above results for n=1, n=2 and even n=0, which is the case c=1.

The above result could be found from an inductive process, using the product rule, but the inductive step is similar to that which allows extension of the binomial theorem to all positive integers, and adds little to this presentation.

The extension from $f(x)=x^n$ to arbitrary polynomials (only finite order will be considered here) needs only two straightforward, perhaps even obvious results:

- The derivative of the sum of two function is the sum of the derivatives.
- The derivative of a function multiplied by a constant is the derivative of the fuctnion multiplied by the same constant.

In symbols, these results are

$$\frac{d}{dx}\left[f(x)+g(x)\right]=\frac{d\,f}{dx}+\frac{dg}{dx}$$

$$\frac{d}{dx}\left(c\,f(x)\right)=c\frac{d\,f}{dx}.$$

In the above, c is a constant, and differentiability of the functions at the desired points is assumed.

Combining all of these results, we can see that for the coefficients a_k all constants,

$$\text{If} \qquad f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{m-1} x^{m-1} + a_m x^m$$

$$\text{then} \qquad \frac{d}{dx} f(x) = f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + (m-1)a_{m-1} x^{m-2} + m a_m x^{m-1}.$$

This is often seen in summation notation as

$$\text{If} \qquad f(x) = \sum_{k=0}^{m} a_k x^k$$

$$\text{then} \qquad \frac{d}{dx} f(x) = f'(x) = \sum_{k=0}^{m} k a_k x^{k-1}.$$

Rationales

Journal Information

This lively journal is produced five times per year and includes contributions from mathematics practitioners. It reflects the best of current thinking and practice. In addition to articles covering mathematics teaching, it provides practical advice on general teaching methods, information on the analysis of official reports and reviews of classroom based projects and surveys. It is aimed at teachers of secondary pupils, students in training and all those with a professional interest in mathematics education. (Teachers of primary pupils may also find this journal useful.) Regular features include Mathematical Resources, Book Reviews and the Correspondence Column. Mathematics in School is published by the Mathematical Association, an organization based out of the UK.

Publisher Information

The Association exists to bring about improvements in the teaching of mathematics and its applications, and to provide a means of communication among students and teachers of mathematics. Its work is carried out through its Council

Mathematics is the study of order, relation, pattern, uncertainty and generality and is underpinned by observation, logical reasoning and deduction. From its origin in counting and measuring, its development throughout history has been catalysed by its utility in explaining real-world phenomena and its inherent beauty. It has evolved in highly sophisticated ways to become the language now used to describe many aspects of the modern world.

Mathematics is an interconnected subject that involves understanding and reasoning about concepts and the relationships between those concepts. It provides a framework for thinking and a means of communication that is powerful, logical, concise and precise.

The Mathematics Stage 6 syllabuses are designed to offer opportunities for students to think mathematically. Mathematical thinking is supported by an atmosphere of questioning, communicating, reasoning and reflecting and is engendered by opportunities to generalise, challenge, identify or find connections and think critically and creatively.

All Mathematics Stage 6 syllabuses provide opportunities for students to develop 21st-century knowledge, skills, understanding, values and attitudes. As part of this, in all courses students are encouraged to learn with the use of appropriate technology and make appropriate choices when selecting technologies as a support for mathematical activity.

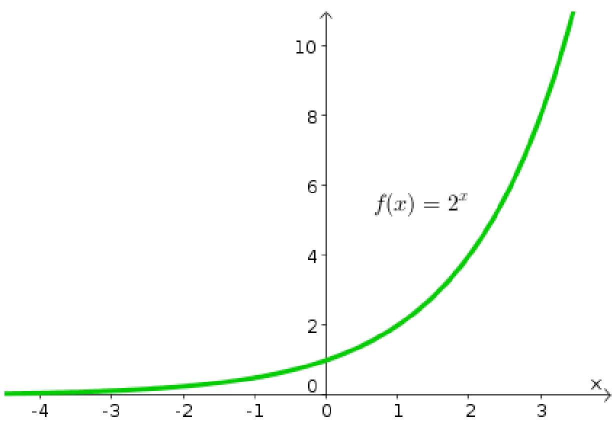
The Mathematics Advanced, Mathematics Extension 1 and Mathematics Extension 2 courses form a continuum to provide opportunities at progressively higher levels for students to acquire knowledge, skills and understanding in relation to concepts within the area of mathematics that have applications in an increasing number of contexts. These concepts and applications are appropriate to the students’ continued experience of mathematics as a coherent, interrelated, interesting and intrinsically valuable study that forms the basis for future learning. The concepts and techniques of differential and integral calculus form a strong basis of the courses, and are developed and used across the courses, through a range of applications and in increasing complexity.

The Mathematics Advanced course is focused on enabling students to appreciate that mathematics is a unique and powerful way of viewing the world to investigate order, relation, pattern, uncertainty and generality. The course provides students with the opportunity to develop ways of thinking in which problems are explored through observation, reflection and reasoning.

The Mathematics Advanced course provides a basis for further studies in disciplines in which mathematics and the skills that constitute thinking mathematically have an important role. It is designed for those students whose future pathways may involve mathematics and its applications in a range of disciplines at the tertiary level.

Exponential

The exponential function is one of the most important functions in mathematics (though it would have to admit that the linear function ranks even higher in importance). To form an exponential function, we let the independent variable be the exponent. A simple example is the function0

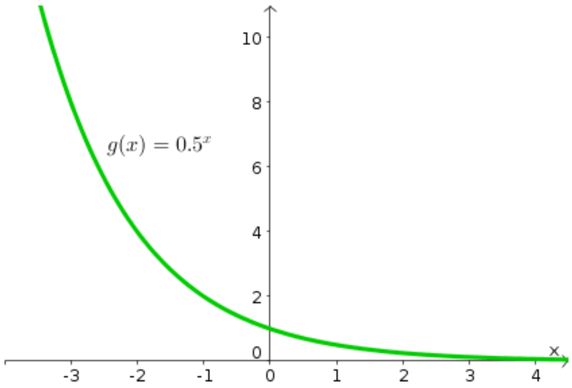


As illustrated in the above graph of ff, the exponential function increases rapidly. Exponential functions are solutions to the simplest types of dynamical systems. For example, an exponential function arises in simple models of bacteria growth

An exponential function can describe growth or decay. The function

g(x)=(12)xg(x)=(12)x

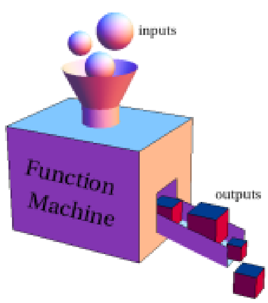
is an example of exponential decay. It gets rapidly smaller as xx increases, as illustrated by its graph.



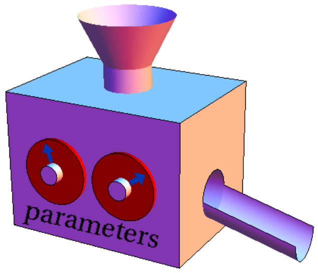
In the exponential growth of f(x)f(x), the function doubles every time you add one to its input xx. In the exponential decay of g(x)g(x), the function shrinks in half every time you add one to its input xx. The presence of this doubling time or half-life is characteristic of exponential functions, indicating how fast they grow or decay.

Parameters of the exponential function

As with any function, the action of an exponential function f(x)f(x) can be captured by the function machine metaphor that takes inputs xx and transforms them into the outputs f(x)f(x).



The function machine metaphor is useful for introducing parameters into a function. The above exponential functions f(x)f(x) and g(x)g(x) are two different functions, but they differ only by the change in the base of the exponentiation from 2 to 1/2. We could capture both functions using a single function machine but dials to represent parameters influencing how the machine works.



We could represent the base of the exponentiation by a parameter bb. Then, we could write ff as a function with a single parameter (a function machine with a single dial):

f(x)=bx.f(x)=bx.

When b=2b=2, we have our original exponential growth function f(x)f(x), and when b=12b=12, this same ff turns into our original exponential decay function g(x)g(x). We could think of a function with a parameter as representing a whole family of functions, with one function for each value of the parameter.

We can also change the exponential function by including a constant in the exponent. For example, the function

h(x)=23xh(x)=23x

is also an exponential function. It just grows faster than f(x)=2xf(x)=2x since h(x)h(x) doubles every time you add only 1/31/3 to its input xx. We can introduce another parameter kk into the definition of the exponential function, giving us two dials to play with. If we call this parameter kk, we can write our exponential function ff as

f(x)=bkx.f(x)=bkx.

You can explore the influence of both parameters bb and kk in the following applet.

It turns out that adding both parameters bb and kk to our definition of ff is really unnecessary. We can still get the full range of functions if we eliminate either bb or kk. You can see this fact through the above applet. For example, you can see that the function f(x)=32xf(x)=32x (k=2k=2, b=3b=3) is exactly the same as the function f(x)=9xf(x)=9x (k=1k=1, b=9b=9). In fact, for any change you make to kk, you can make a compensating change in bb to keep the function the same. To see this, check the “fix function” checkbox. Then, if you change either bb or kk, the applet will automatically make a compensatory change in the other parameter to keep the function the same. If you are curious why this is true, you can check out the calculation showing the two parameters are redundant.

Since it is silly to have both parameters bb and kk, we will typically eliminate one of them. The easiest thing to do is eliminate kk and go back to the function

f(x)=bx.f(x)=bx.

We will use this function a bit at first, changing the base bb to make the function grow or decay faster or slower.

However, once you start learning some calculus, you'll see that it is more natural to get rid of the base parameter bb and instead use the constant kk to make the function grow or decay faster or slower. Except, we can't exactly get rid of the base bb. If we set b=1b=1, we'd have the boring function f(x)=1f(x)=1, or, if we set b=0b=0, we'd have the even more boring function f(x)=0f(x)=0. We need to choose some other value of bb. If we didn't have calculus, we'd probably choose b=2b=2, writing our exponential function as f(x)=2kxf(x)=2kx. Or, since we like the decimal system so well, maybe we'd choose b=10b=10 and write our exponential function of f(x)=10kxf(x)=10kx. According to the above discussion, it shouldn't matter whether we use b=2b=2 or b=10b=10, as we can get the same functions either way (just with different values of kk).
But, it turns out that calculus tells us there is a natural choice for the base bb. Once you learn some calculus, you'll see why the most common base bb throughout the sciences is the irrational number

e=2.718281828459045….e=2.718281828459045….

Fixing b=eb=e, we can write the exponential functions as

f(x)=ekx.f(x)=ekx.

(The applet understands the value of ee, so you can type ee in the box for bb.)
Using ee for the base is so common, that exex (“e to the xx”) is often referred to simply as *the exponential function*.

To increase the possibilities for the exponential function, we can add one more parameter cc that scales the function:

f(x)=cbkx.f(x)=cbkx.

Since f(0)=cbk0=cf(0)=cbk0=c, we can see that the parameter cc does something completely different than the parameters bb and kk. We'll often use two parameters for the exponential function: cc and one of bb or kk. For example, we might set k=1k=1 and use

f(x)=cbxf(x)=cbx

or set b=eb=e and use

f(x)=cekx.f(x)=cekx.

You can add the parameter cc to the applet by checking the “scale function” checkbox.

logarithmic and trigonometric function

A logarithmic function has three main components. The first component is the base, b; the second component is the fixed value, y, which is what you input into the function; and the third component is the output of the logarithm function, x. The output of the logarithm function is the answer to the following question: to what exponent must I raise the base, b, in order to achieve the fixed value y? That is, the logarithm with base b of y is the solution to this equation: $b^x = y$.

The conventional notation is $\log_b(y) = x$, which is read aloud as “log base b of y is equal to x.”

Example: if b = 2 and x = 8, then $\log_2(8) = 3$ since 3 satisfies the equation $2^x = 8$.

Another example: $\log_{10}(100) = 2$ since $10^2 = 100$.

There the most commonly used bases for logarithms are base b=2, base b=10, and base b=e (where e is the constant approximately equal to 2.718). The logarithm base e is commonly referred to as the natural logarithm and has many applications in pure mathematics and calculus. The standard notation for $\log_e(y)$ is ln(y). The standard notation for $\log_{10}(y)$ is log(y). If there is no base given, you assume it is base 10.

Now that we’ve gotten through the basics of what a logarithm is, let’s look at a few logarithmic identities.

Product Identity

This identity comes from the exponent rule for products: $b^x b^w = b^{x+w}$

Suppose we have two fixed values y and z and suppose we know that $\log_b(y) = x$ and $\log_b(z) = w$. This means that $b^x = y$ and $b^w = z$. If we multiply y and z, we get $yz = b^x b^w = b^{x+w}$. This means that x+w is the exponent which we must raise b to in order to achieve the fixed value of yz. That is,

$\log_b(yz) = \log_b(y) + \log_b(z)$.

This is the product rule for logarithms. In words, $\log_b(yz) = \log_b(y) + \log_b(z)$ means that the logarithm of a product is equal to the sum of the logarithms of the factors.

Let’s compute $\log_7(7*49)$. We notice that $\log_7(7) = 1$ and $\log_7(49) = 2$.

Thus, $\log_7(7*49) = \log_7(7) + \log_7(49) = 1+2 = 3$.

Quotient Identity

This identity comes from the exponent rule for products: $(b^x)/(b^w)=b^{x-w}$

Just as before, if $\log_b(y) = x$ and $\log_b(z) = w$. This means that $b^x = y$ and $b^w = z$. If we divide y and z, we get $y/z = (b^x)/(b^w) = b^{x-w}$.

This means that x-w is the exponent which we must raise b to in order to achieve the fixed value of y/z. That is,

$\log_b(y/z) = \log_b(y) - \log_b(z)$.

This is the quotient rule for logarithms. In words, $\log_b(y/z) = \log_b(y) - \log_b(z)$ means that the logarithm of a quotient is equal to the difference of the logarithms of the factors.

Let's compute $\log_{10}(1/10{,}000)$. We notice that $\log_{10}(1) = 0$ and $\log_{10}(10{,}000) = 4$.

Thus, $\log_{10}(1/10{,}000) = \log_{10}(1) - \log_{10}(10{,}000) = 0 - 4 = -4$. You can check with a calculator that $10^{(-4)} = 1/10{,}000$.

Power Identity

Notice that, $b^y = a$ implies that $(b^y)^x = b^{yx} = a^x$. This means that the exponent on b which causes a^x is the same as x times the logarithm base b of a. This gives us the power identity for logarithmic functions is:

$\log_b(a^x) = x*\log_b(a)$.

When you are trying to simplify logarithmic functions, it is helpful to remember that when you see an exponent inside of a log, you can pull that exponent out to the front of the function.

Example: $\log_2(16^5) = 5*\log_2(16) = 5\log_2(2^4) = 5*4*\log_2(2) = 5*4*1 = 20$.

A couple important things to note about the power identity:

1. The exponents which you "pull out" do not have to be whole numbers. For example: $\log_{15}(225^{-4}) = -4*\log_{15}(225)$ and $\log_3(90.125) = 0.125\log_3(9)$.

2. You can only pull an exponent out if the entire fixed quantity has that exponent applied to it.
   --For example, if the fixed quantity is $(w+y+z)^a$, then $\log_b((w+y+z)^a) = a*\log_b(w+y+z)$
   --However, if the fixed quantity is $w^a + y^a$, then $\log_b(w^a + y^a)$ cannot be simplified.

Inverse functions

The power rule for the logarithm gives us that

$\log_b(b^x) = x\log_b(b) = x$.

This means that the inverse function to $\log_b(y)$ is b^x. On the other hand, $\log_b(y)$ is the exponent we apply to b to get y. That means that

$y = b^{\log_b y}$

This means that the inverse function to taking the x-th power of b is the logarithm function base b. When you are solving equations and you've got a $\log_b$ on one side, you can exponentiate the entire equation base b to "cancel" out that logarithm. Here's what I mean:

Start with $\log_{13}(x) = 25$. Exponentiate both sides with respect to 13 to get: $13^{\log_{13} x} = 13^{25}$. Canceling out the 13 and $\log_{13}$ we get: $x = 13^{25}$.

Similarly, if we start with $8^x = 64$ and then take the logarithm of both sides we get $\log_8(8^x) = \log_8(64)$. Canceling out the log8 and the 8 on the right, we get $x = \log_8(64) = 2$.

Changing Bases

There is a straightforward equation for computing the $\log_b(x)$ with respect to another basis, k:

$\log_b(x) = (\log_k(x))/(\log_k(b))$.

For example, $\log_{16}(32) = \log_2(32)/\log_2(16) = 5/4$.

The exponential and the logarithmic functions are perhaps the most important functions you'll encounter whenever dealing with a physical problem. They are the inverse of each other and can be used to represent a large range of numbers very conveniently.

They are continuous and differentiable over their entire domain, and the simplicity in notation, of their derivatives, would give you an idea about their huge significance in mathematics as well as other subjects. Let us now first understand these functions individually, before moving on to the connection between them.

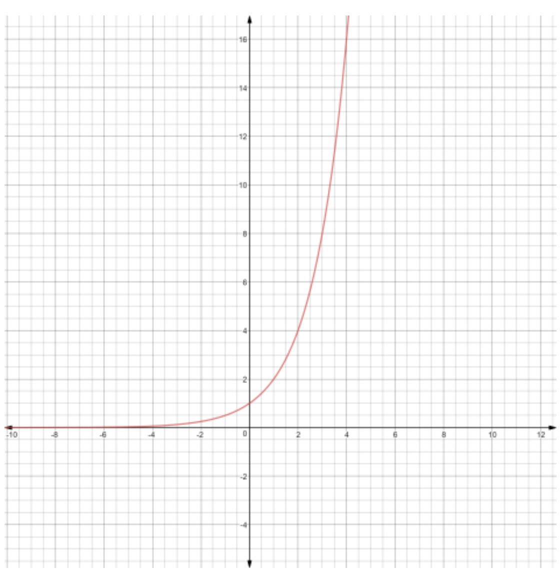
xponential Functions

The term ‘exponent’ implies the ‘power’ of a number. For eg – the exponent of 2 in the number 2^3 is equal to 3. Clearly then, the exponential functions are those where the variable occurs as a power. An exponential function is defined as-f(x)=axf(x)=axwhere a is a positive real number, not equal to 1.

If a = 1, then $f(x) = 1^x$, which is equal to 1, ∀ x. Hence the graph of the function would just be a straight line of constant y (= 1). Depending on the value of ‘a’, we can have two possible cases:

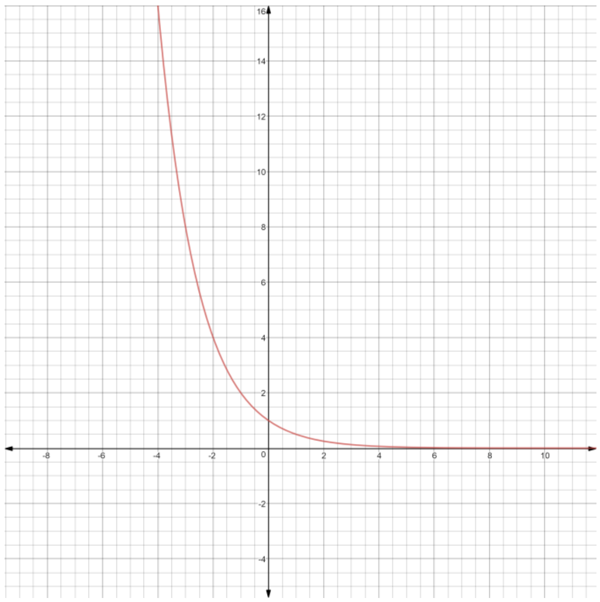
Case 1: a > 1

Here, the exponential function increases very rapidly with increasing x and tends to +∞ as x tends to +∞. When x = 0, $a^x = 1$; and when x tends to -∞, the function tends to 0. The general graph of the function looks like this: (where a = 2)



Case 2: a < 1

The function decreases very rapidly with increasing x and tends to 0 as x tends to +∞. When x = 0, a^x = 1 as usual; and when x tends to -∞, the function tends to +∞. The general graph of such a function looks like this – (where a = 2 again)



Properties of Exponential Functions

- The domain of the exponential function is (-∞,+∞) i.e. it is defined ∀ x.
- The range of the exponential function is (0,+∞). This property should be clear from the graph of the function a^x. Otherwise, also, it is logical that the power of any real number can't be a negative number. Only imaginary numbers can have such a behavior.
- The points (0,1) and (1, a) always lie on the graph of the function a^x.
- 'a' must necessarily be a positive number. If a is a negative number, then for any fractional values of x, we will get an imaginary number as a result which can't be plotted on the same graph. For eg- $(-2)^{1/2} = \sqrt{2}i$.

  - The Product Rule –ax.ay=ax+yax.ay=ax+y
  - The Quotient Rule –axay=ax–yaxay=ax–y
  - The exponential function is continuous and differentiable throughout its domain. The derivative is given asddx(ax)=axln(a)
  - where ln(a) or $\log_e$(a)is the natural logarithm of a. We'll define it formally in some time. The standard exponential function e^x is a unique function in mathematics with the property of being equal to its derivative. Thus, we have – ddx(ex)=exddx(ex)=ex

In fact, the calculation behind these derivatives forms one of the methods of defining the number 'e' which is equal to 2.71828… That's all about exponential functions for now.

Logarithmic Functions

Since we had already disclosed that the logarithm function and the exponential function are inverses of each other, it should be obvious then that the logarithm function does the opposite of 'taking the power of a number'. Let's look at it mathematically –

General Notation

- Exponential Form –by=xby=x
- Logarithmic Form –y=logbxy=logbxwhere 'b' is the base of the log.

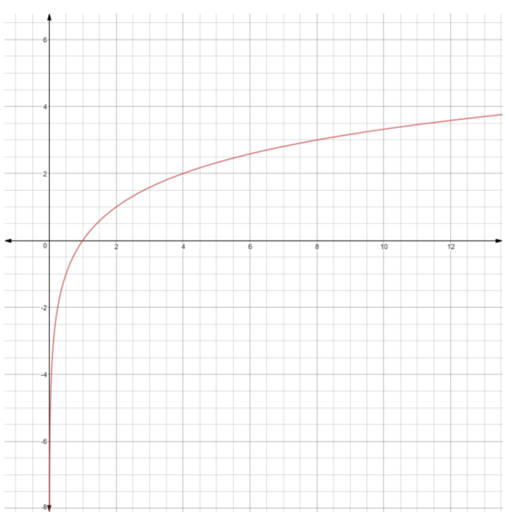
With these two forms, you can easily see that the value of the function f(x) = $\log_b x$ is the power to which 'b' must be raised to get 'x'. 'x' therefore, can't be negative since that would require 'b' to be imaginary, the conditions on the base 'b' –

- b > 0: It follows directly from the exponential representation of the logarithmic function.
  - b ≠ 1: Since 1 raised to any power would only give 1.

Depending on the value of 'b', we will have two possible cases –

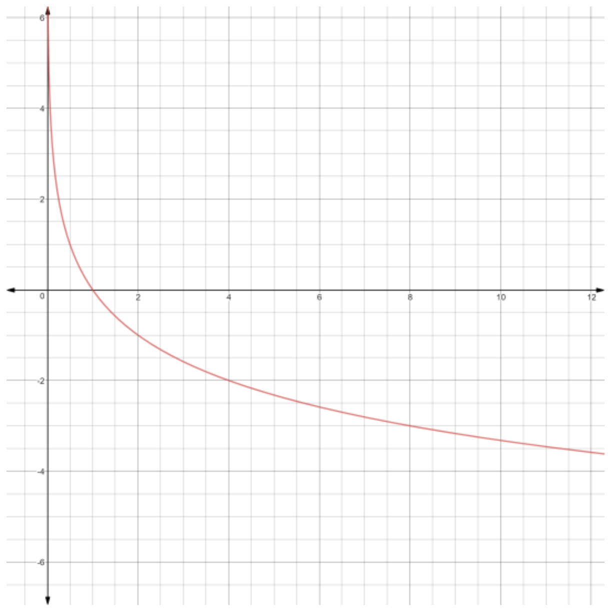
Case 1: b > 1

Here, the logarithmic function decreases very rapidly with decreasing x and tends to -∞ as x tends to 0. When x tends to +∞, the function also tends to +∞ with an ever-decreasing rate of increase. The general graph of the function looks like this – (where b = 2)



Case 2: 0 < b < 1

Here the function increases very rapidly to +∞ as x tends to 0, and falls at an ever decreasing rate to -∞ as x tends to +∞. The general graph is as shown – (where b = 0.5)



Properties of Logarithmic Functions

- The domain of the logarithmic functions is (0, +∞).
- The range of the logarithmic function is (-∞,+∞).
- The points (1,0) and (b,1) always lie on the graph of the function $\log_b x$.
- The Product Rule:logb(xy)=logbx+logbylogb(xy)=logbx+logby
- The Quotient Rule:logb(xy)=logbx–logbylogb(xy)=logbx–logby
- The Power Rule:logbax=xlogbalogbax=xlogbaGeneralization:logbaf(x)=f(x)logbalogbaf(x)=f(x)logba
- Change of Base Formula – To change the logarithm from a given base ‘b’ to base ‘a’logbx=logaxlogablogbx=logaxlogab
- The logarithm function is continuous and differentiable throughout its domain. The derivative is given asddx(logbx)=1xln(b)ddx(logbx)=1xln(b)where ln(b) or $\log_e b$ is the natural logarithm of b. This is a standard logarithm function. It has the base = e = 2.71828. Its derivative –ddx(ln(x))=1xddx(ln(x))=1xsince ln(e) = 1.

Relation between Exponential and Logarithmic Functions

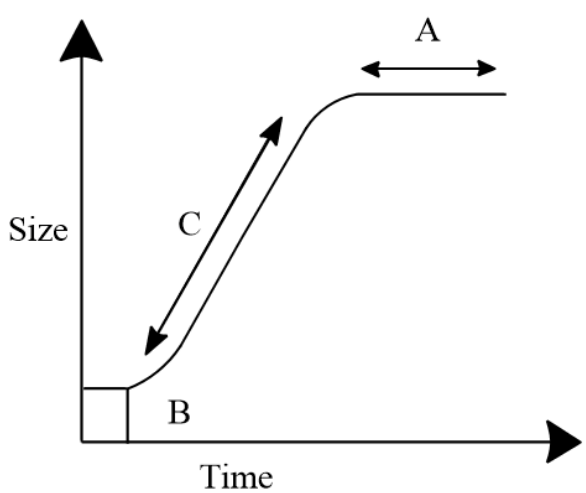
We have already told you that the logarithmic and the exponential functions are inverses of each other. You can now verify this from the properties as well.

- The range and the domain of the two functions are exchanged.
- The points (0,1) and (1, a) always lie on the exponential function's graph while (1,0) and (b,1) always lie on the logarithmic function's graph.
- Product and Quotient Rules of the exponential and the logarithm functions follow from each other.

Let us now put our statement in a mathematical form for the standard functions –
eln(x)=ln(ex)=xeln(x)=ln(ex)=x
General formula –
blogbx=logbbx=xblogbx=logbbx=x

Solved Examples for You

Question: Given below is a graph drawn on the parameters of growth versus time.A,B,C respectively represents



- Exponential phase, log phase, and steady-state phase
- Steady-state phase, lag phase, and log phase
- Slow growing phase, lag phase, and steady-state phase
- Lag phase, steady-state phase, and logarithmic phase
- Log phase, lag phase, and steady-state phase

Solution: B. The first stage in the growth phase is a lag phase, where there is minimal growth. The next stage in the growth phase is the log phase, which is also known as the exponential phase where the growth is manifold. The final stage is a steady state where the growth is zero and thus known as the steady state.

This concludes our discussion on this topic of the exponential and logarithmic functions.

Unit-V Integration

Integration as inverse of differentiation

In mathematics, we usually need to find the derivative of some mathematical functions. It gives the rate of change of one variable with respect to others. Integration is the opposite process of differentiation. The fundamental use of integration is to get back the function whose derivatives are known. So, it is like an anti-derivative procedure. Thus, integrals are computed by viewing an integration as an inverse operation to differentiation. In this topic, the student will learn the Integration concepts as well as some integration formula with examples. Let us learn it!

Integration is the algebraic method to find the integral for a function at any point on the graph. Finding the integral of some function with respect to some variable x means finding the area to the x-axis from the curve. Therefore, the integral is also called the anti-derivative because integrating is the reverse process of differentiating.

The integral comes from not only to determine the inverse process of taking the derivative. But also for solving the area problem as well. Similar to the process of differentiation for finding the slope at any point on the graph, this process of integration will be used to find the area of the curve up to any point on the graph.

The integral of the function of x from range a to b will be the sum of the rectangles to the curve at each interval of change in x as the number of rectangles goes to infinity.

The notation, which we have stuck with for historical reasons, is as peculiar as the notation for derivatives:

The integral of a function f(x) with respect to x is written as:

$\int f(x)dx\int f(x)dx$

Also, integration is considered as almost an inverse to the operation of differentiation means that if,

$ddxf(x)=g(x)ddxf(x)=g(x)$

then

$\int g(x)dx=f(x)+C\int g(x)dx=f(x)+C$

The extra C called the constant of integration, which is really necessary. This is because that after all differentiation kills off constants, which is why integration and differentiation are not exactly inverse operations of each other.

Since integration is almost the inverse operation of differentiation, the recollection of formulas and processes for differentiation is possible. So, many differentiation formulae will be used to provide the corresponding formula for the integration.

Definite integrals are the special kind of integration, where both endpoints are fixed. So, it always represents some bounded region, for computation.

Some properties of Integration:

And since the derivative of a sum is the sum of the derivatives, the integral of a sum is the sum of the integrals:

$\int f(x)+g(x)dx=\int f(x)dx+\int g(x)dx\int f(x)+g(x)dx=\int f(x)dx+\int g(x)dx$

And, likewise, constants ‘go through’ the integral sign:

$\int c\cdot f(x)dx=c\cdot\int f(x)dx\int c\cdot f(x)dx=c\cdot\int f(x)dx$

Formula for Integration:

1. $\int xndx=1n+1xn+1+C\int xndx=1n+1xn+1+C$
2. unless $n=-1$ unless $n=-1$
3. $\int exdx=ex+C\int exdx=ex+C$
4. $\int 1xdx=\ln x+C\int 1xdx=\ln x+C$
5. $\int \sin xdx=-\cos x+C\int \sin xdx=-\cos x+C$
6. $\int \cos xdx=\sin x+C\int \cos xdx=\sin x+C$
7. $\int \sec 2xdx=\tan x+C\int \sec 2xdx=\tan x+C$
8. $\int 11+x2dx=\arctan x+C\int 11+x2dx=\arctan x+C$
9. $\int axdx=ax\ln a+C\int axdx=ax\ln a+C$
10. $\int \log axdx=1\ln a\cdot 1x+C\int \log axdx=1\ln a\cdot 1x+C$
11. $\int 1\sqrt{}1-x2dx=\arcsin x+C\int 11-x2dx=\arcsin x+C$
12. $\int 1x\sqrt{}x2-1dx=$ arcsecx+C

integration of simple Functions

Integration is an important concept in mathematics and—together with its inverse, differentiation—is one of the two main operations in calculus. Given a function ff of a real variable xx, and an interval [a,b][a,b] of the real line, the definite integral ∫baf(x)dx∫abf(x)dx is defined informally to be the area of the region in the xyxy-plane bounded by the graph of ff, the xx-axis, and the vertical lines x=ax=a and x=bx=b, such that area above the xx-axis adds to the total, and that below the xx-axis subtracts from the total. The term integral may also refer to the notion of the anti-derivative, a function FF whose derivative is the given function ff.

More rigorously, once an anti-derivative FF of ff is known for a continuous real-valued function ff defined on a closed interval [a,b][a,b], the definite integral of ff over that interval is given by

∫baf(x)dx=F(b)−F(a)∫abf(x)dx=F(b)−F(a)

If FF is one anti-derivative of ff, then all other anti-derivatives will have the form F(x)+CF(x)+C for some constant CC. The collection of all anti-derivatives is called the indefinite integral of ff and is written as

∫fdx=F(x)+C∫fdx=F(x)+C

Integration proceeds by adding up an infinite number of infinitely small areas. This sum can be computed by using the anti-derivative.

The integral of a linear combination is the linear combination of the integrals.

∫ba(αf+βg)(x)dx=α∫baf(x)dx+β∫bag(x)dx

If f(x)≤g(x)f(x)≤g(x) for each xx in [a,b][a,b], then each of the upper and lower sums of ff is bounded above by the upper and lower sums, respectively, of gg:

∫baf(x)dx≤∫bag(x)dx

Introduction

In calculus, integration by parts is a theorem that relates the integral of a product of functions to the integral of their derivative and anti-derivative. It is frequently used to find the anti-derivative of a product of functions into an ideally simpler anti-derivative. The rule can be derived in one line by simply integrating the product rule of differentiation.

Theorem of integration by parts

Let’s take the functions u=u(x)u=u(x) and v=v(x)v=v(x). When taking their derivatives, we are left with du=u‘(x)du=u‘(x) and dxdv=v′(x)dxdxdv=v′(x)dx. Now, let’s take a look at the principle of integration by parts:

∫u(x)v′(x)dx=u(x)v(x)−∫u′(x)v(x) dx∫u(x)v′(x)dx=u(x)v(x)−∫u′(x)v(x) dx

or, more compactly,

∫udv=uv−∫vdu∫udv=uv−∫vdu

Proof

Suppose u(x)u(x) and v(x)v(x) are two continuously differentiable functions. The product rule states:

ddx(u(x)v(x))=u(x)ddx(v(x))+ddx(u(x))v(x)ddx(u(x)v(x))=u(x)ddx(v(x))+ddx(u(x))v(x)

Integrating both sides with respect to xx, over an interval a≤x≤ba≤x≤b,

∫baddx(u(x)v(x))dx=∫bau′(x)v(x)dx+∫bau(x)v′(x)dx∫abddx(u(x)v(x))dx=∫abu′(x)v(x)dx+∫abu(x)v′(x)dx

then applying the fundamental theorem of calculus,

∫baddx(u(x)v(x))dx=[u(x)v(x)]ba∫abddx(u(x)v(x))dx=[u(x)v(x)]ab

gives the formula for “integration by parts”:

[u(x)v(x)]ba=∫bau′(x)v(x)dx+∫bau(x)v′(x)dx[u(x)v(x)]ab=∫abu′(x)v(x)dx+∫abu(x)v′(x)dx.

Visulization

Let’s define a parametric curve by (x,y)=(f(t),g(t))(x,y)=(f(t),g(t)).

integration by parts

Integration by parts is another technique for simplifying integrands. As we saw in previous posts, each differentiation rule has a corresponding integration rule. In the case of integration by parts, the corresponding differentiation rule is the Product Rule. The technique of integration by parts allows us to simplify integrands of the form:

$$\int f(x)g(x)dx\int f(x)g(x)dx$$

Examples of this form include:

$$\int x\cos x\ dx,\int e^x\cos x\ dx,\int x2ex\ dx\int x\cos⁡x\ dx,\int e^x\cos⁡x\ dx,\int x2ex\ dx$$

As integration by parts is the product rule applied to integrals, it helps to state the Product Rule again. The Product Rule is defined as:

$$ddx[f(x)g(x)]=f'(x)g(x)+f(x)g'(x)ddx[f(x)g(x)]=f'(x)g(x)+f(x)g'(x)$$

When we apply the product rule to indefinite integrals, we can restate the rule as:

$$\int ddx[f(x)g(x)]\ dx=\int [f'g(x)+f(x)g'(x)]\ dx\int ddx[f(x)g(x)]\ dx=\int [f'g(x)+f(x)g'(x)]\ dx$$

Then, rearranging so we get $f(x)g'(x)\ dxf(x)g'(x)\ dx$ on the left side of the equation:

$$\int f(x)g'(x)\ dx=\int ddx[f(x)g(x)]\ dx-\int f'(x)g(x)\ dx\int f(x)g'(x)\ dx=\int ddx[f(x)g(x)]\ dx-\int f'(x)g(x)\ dx$$

Which gives us the integration by parts formula! The formula is typically written in differential form:

$$\int u\ dv=uv-\int v\ du$$

Examples

The following examples walkthrough several problems that can be solved using integration by parts. We also employ the wonderful SymPy package for symbolic computation to confirm our answers. To use SymPy later to verify our answers, we load the modules we will require and initialize several variables for use with the SymPy library.

Example 1: Evaluate the integrand ∫xsinx2 dx∫xsinx2 dx

Recalling the differential form of the integration by parts

formula, ∫u dv=uv−∫v du∫u dv=uv−∫v du, we set u=xu=x and dv=sinx2dv=sinx2

Solving for the derivative of uu, we arrive at du=1 dx=dxdu=1 dx=dx. Next, we find the antiderivative of dvdv. To find this antiderivative, we employ the Substitution Rule.

u=12x,du=12 dx,dudx=2u=12x,du=12 dx,dudx=2

y=sinu,dy=−cosu du,dydu=−cosuy=sinu,dy=−cosu du,dydu=−cosu

Therefore, v=−2cosx2v=−2cosx2

Entering these into the integration by parts formula:

−2xcosx2−(−2)∫cosx2−2xcosx2−(−2)∫cosx2

Then, solving for the integrand ∫cosx2∫cosx2, we employ the Substitution Rule again as before to arrive at 2sinx22sinx2 (the steps in solving this integrand are the same as before when we solved for ∫sinx2∫sinx2). Thus, the integral is evaluated as:

−2xcosx2+4sinx2+C−2xcosx2+4sinx2+C

Using SymPy's integrate, we can verify our answer is correct (SymPy does not include the constant of integration C).

Integration is a very important computation of calculus mathematics. Many rules and formulas are used to get integration of some functions. A special rule, which is integration by parts, is available for integrating the products of two functions. This topic will derive and illustrate this rule which is Integration by parts formula. Also, some examples will help the students to get their concept. Let us start!

What is integration by parts method?

This method is very useful in order to master the technique of integrations. Many times we have to integrate the product of two functions. Functions often arise as the products of other functions, and so we have to integrate these products. For example, we may be asked to determine

∫x∫x; cosxcosx; dx

Here, the integrand is the product of the two functions x and cos x. A rule exists for integrating the products of functions which is required for getting the solution.

integration by substitution

Integration by Substitution.

A key strategy in mathematical problem-solving is substitution or changing the variable: that is, replacing one variable with another, related one. A problem that starts out difficult can sometimes become very easy with an appropriate change of variable. Integration problems are no exception.

Several variants of this technique are used in integration, but they all depend on the following key fact:

$$\int ydxdudx=\int ydu.$$

Indefinite Integration

Here's a simple example. Suppose we have to calculate

$$\int xex2dx.$$

Now, as it stands, this problem is hard, because we have to integrate a product, one half of which is the composite function ex2, and neither a product rule nor a chain rule exist for integration

However, if we set u=x2, then dxdu=2x and thus x=21dxdu. Thus

$$\int xex2dx = = = \int ex2 \times xdx \int eu \times 21dxdudx\ 21\int eudu.$$

This is now an easy integral; the answer is 21eu+c, or (in terms of our original variable x) 21ex2+c.

The reason we were able to use this technique is that the integrand, xex2, has the form ``function of a function, times derivative of the inner function''. When an integrand takes this form, the substitution ``u= (inner function)'' will generally make the integration simpler.

Summary of Technique

Given an integral:

1. Check that the integrand has the form ``function of a function, times derivative of the inner function'' (possibly times a constant).
2. Set u equal to the inner function.
3. Using

$$\int ydxdudx=\int ydu,$$

create a simpler integral in terms of u.

4. Perform this integral to calculate the answer in terms of u.
5. Express this answer in terms of the original variable x.

0 **Definite Integration**

When performing an indefinite integral by substitution, the last step is always to convert back to the variable you started with: to convert an expression in u to an expression in x. With definite integration, however, there's an alternative: you can change your x-limits to u-limits, and then (in effect) forget about x.

Here's an example. To calculate

$$\int 0\pi/2cosxesinxdx,$$

we first set u=sinx, then dxdu=cosx.

Now, if x=0 then u=sinx=0, and if x=$\pi/2$ then u=sin($\pi/2$)=1.

Thus

$$\int 0\pi/2cosxesinxdx = = = = = \int 0\pi/2esinx \times cosxdx \int 0\pi/2eu \times dxdudx \int 01eudu \left[eu\right]10\ e-1.$$

Summary of Technique

Given an integral:

1. Check that the integrand has the form ``function of a function, times derivative of the inner function'' (possibly times a constant).
2. Set u equal to the inner function.
3. Using

$$\int x0x1ydxdudx=\int u0u1ydu,$$

create a simpler integral in terms of u, with fresh limits, representing u-values rather than x-values.

4. Perform this integral.

Substitution "The Other Way Round"

You've already seen how to perform integration using substitutions of the form

$$u=f(x),$$

where x is the old variable and u the new. Notice that the new variable is expressed as a function of the old one.

There are classes of integrals that can instead be attacked using substitutions expressed ``the other way round'': substitutions in which the old variable is expressed as a function of the new one.

An example is

$$\int \sqrt{} 9-x2dx.$$

The substitution that works here is $x=3\sin\theta$. Using the fact that

$$\int ydx=\int ydx d\theta d\theta,$$

we obtain

$$\int \sqrt{} 9-x2dx = = = = = \int \sqrt{} 9-x2dxd\theta d\theta \int \sqrt{} 9-9\sin2\theta \times 3\cos\theta d\theta \int 3\sqrt{} 1-\sin2$$

$$\theta \times 3\cos \theta d\theta \int 3\cos \theta \times 3\cos \theta d\theta 9\int \cos2 \theta d\theta.$$

At this point, you need to recall that

$$\cos2\theta \equiv 21(1+\cos2\theta),$$

and therefore

$$9\int \cos2\theta d\theta = = 29\int (1+\cos2\theta)d\theta \; 29\left(\theta+21\sin2\theta\right)+c.$$

The problem that faces us now is how to express this answer in terms of our original variable x. This becomes easier when we recall that

$$\sin2\theta \equiv 2\sin\theta\cos\theta,$$

and deduce that our answer can be expressed as

$$29(\theta+\sin\theta\cos\theta)+c.$$

Now, $x=3\sin\theta$ and thus $\sin\theta=x/3$ and $\cos\theta=\sqrt{1-x2/9}$. We can therefore write our answer as

$$29\left(\sin-1x3+x3\sqrt{1-9x2}\right)+c,$$

which simplifies to give

$$21\left(9\sin-1x3+x\sqrt{9-x2}\right)+c.$$

It's probably not obvious when this technique is applicable. The answer is that when the integrand contains an expression of the form $\sqrt{a2-x2}$, you should try the substitution $x=a\sin\theta$, and when it contains an expression of the form a2+x2, you should try $x=a\tan\theta$.

A set of related substitutions involve the hyperbolic functions, which you meet later in the course.

Summary of Technique

Given an integral:

1. Check that the integrand contains an expression of the form $\sqrt{a2-x2}$ or a2+x2.
2. Perform, respectively, the substitution $x=a\sin\theta$ or $x=a\tan\theta$.
3. Using

$$\int ydx=\int ydxd\theta d\theta,$$

create an integral in terms of θ.

4. Perform this integral to calculate the answer in terms of θ.
5. Express this answer in terms of the original variable x (this can require some thought).

Definite Integration "The Other Way Round"

With substitutions of the form $x=f(\theta)$, it can be rather difficult, as you've seen, performing the final step of converting your expression in the new variable θ to one in the old variable x. If the integral is definite, however, you can completely avoid having to do this, by changing the limits on the integral instead.

For example, consider

$$\int_0^{3/2}\sqrt{9-x^2}\,dx.$$

As in the indefinite case, we use the substitution $x=3\sin\theta$, but this time we note that if $x=0$, $\theta=\sin^{-1}0=0$ and if $x=3/2$, $\theta=\sin^{-1}(1/2)=\pi/6$. Using the fact that

$$\int_{x_0}^{x_1} y\,dx=\int_{\theta_0}^{\theta_1} y\frac{dx}{d\theta}\,d\theta,$$

we obtain

$$\int_0^{3/2}\sqrt{9-x^2}\,dx = \int_0^{\pi/6}\sqrt{9-x^2}\frac{dx}{d\theta}\,d\theta = \int_0^{\pi/6}\sqrt{9-9\sin^2\theta}\times 3\cos\theta\,d\theta = \int_0^{\pi/6}3\sqrt{1-\sin^2\theta}\times 3\cos\theta\,d\theta = \int_0^{\pi/6}3\cos\theta\times 3\cos\theta\,d\theta = 9\int_0^{\pi/6}\cos^2\theta\,d\theta = \frac{9}{2}\int_0^{\pi/6}(1+\cos 2\theta)\,d\theta = \frac{9}{2}\left[\theta+\frac{1}{2}\sin 2\theta\right]_0^{\pi/6} = \frac{9}{2}\left(\frac{\pi}{6}+\frac{\sqrt{3}}{4}\right) = \frac{3}{8}\left(2\pi+3\sqrt{3}\right).$$

Summary of Technique

Given an integral:

1. Check that the integrand contains an expression of the form $\sqrt{a^2-x^2}$ or a^2+x^2.
2. Perform, respectively, the substitution $x=a\sin\theta$ or $x=a\tan\theta$.
3. Using

$$\int x0x1ydx=\int \theta 0\theta 1ydxd\theta d\theta,$$

create an integral in terms of θ, with fresh limits, representing \theta-values rather than x-values.

4. Perform this integral.

definite integrals

A definite integral is an integral

$$\int_a^b f(x)\,dx$$

with upper and lower limits. If x is restricted to lie on the real line, the definite integral is known as a Riemann integral (which is the usual definition encountered in elementary textbooks). However, a general definite integral is taken in the complex plane, resulting in the contour integral

$$\int_a^b f(z)\,dz,$$

with a, b, and z in general being complex numbers and the path of integration from a to b known as a contour.

The first fundamental theorem of calculus allows definite integrals to be computed in terms of indefinite integrals, since if F is the indefinite integral for a continuous function $f(x)$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

This result, while taught early in elementary calculus courses, is actually a very deep result connecting the purely algebraic indefinite integral and the purely analytic (or geometric) definite integral. Definite integrals may be evaluated in the Wolfram Language using Integrate[f, x, a, b].

The question of which definite integrals can be expressed in terms of elementary functions is not susceptible to any established theory. In fact, the problem belongs to transcendence theory, which appears to be "infinitely hard." For example, there are definite integrals that are equal to the Euler-Mascheroni constant γ. However, the problem of deciding whether γ can be expressed in terms of the values at rational values of elementary functions involves the decision as to whether γ is rational or algebraic, which is not known.

Integration rules of definite integration include

$$\int_a^a f(x)\,dx = 0$$

and

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

For $c \in (a, b)$,

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

If g' is continuous on $[a, b]$ and f is continuous and has an antiderivative on an interval containing the values of $g(x)$ for $a \le x \le b$, then

$$\int_a^b f(g(x))\,g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

Watson's triple integrals are examples of (very) challenging multiple integrals. Other challenging integrals include Ahmed's integral and Abel's integral.

Definite integration for general input is a tricky problem for computer mathematics packages, and some care is needed in their application to definite integrals. Consider the definite integral of the form

$$I(a) = \int_0^{\pi/2} \frac{dx}{1 + (\tan x)^a},$$

which can be done trivially by taking advantage of the trigonometric identity

$$\tan\left(\tfrac{1}{2}\pi - x\right) = \cot x.$$

Letting $z \equiv (\tan x)^a$,

$$\begin{aligned} I(a) &= \int_0^{\pi/4} \frac{dx}{1+z} + \int_{\pi/4}^{\pi/2} \frac{dx}{1+z} \\ &= \int_0^{\pi/4} \frac{dx}{1+z} + \int_0^{\pi/4} \frac{dx}{1+\frac{1}{z}} \\ &= \int_0^{\pi/4} \left(\frac{1}{1+z} + \frac{1}{1+\frac{1}{z}} \right) dx \\ &= \int_0^{\pi/4} dx \\ &= \tfrac{1}{4}\pi. \end{aligned}$$

Many computer mathematics packages, however, are able to compute this integral only for specific values of a, or not at all. Another example that is difficult for computer software packages is

$$\int_{-\pi}^{\pi} \ln\left[2\cos\left(\tfrac{1}{2}x\right)\right] dx = 0,$$

which is nontrivially equal to 0.

Some definite integrals, the first two of which are due to Bailey and Plouffe (1997) and the third of which is due to Guénard and Lemberg (2001), which were identified by Borwein and Bailey (2003, p. 61) and Bailey et al. (2007, p. 62) to be "technically correct" but "not useful" as computed by Mathematica Version 4.2 are reproduced below. More recent versions of Wolfram Language return them directly in the same simple form given by Borwein and Bailey without even the need for additional simplification:

$$\begin{aligned} \int_0^1 \frac{t^2 \ln t \, dt}{(t^2-1)(t^4+1)} &= \frac{\pi^2\left(2-\sqrt{2}\right)}{32} \\ &= 0.18067 \ldots \\ \int_0^{\pi/4} \frac{t^2\, dt}{\sin^2 t} &= K + \tfrac{1}{4}\pi \ln 2 - \tfrac{1}{16}\pi^2 \\ &= 0.84351 \ldots \\ \int_0^{\pi} \frac{x \sin x \, dx}{1+\cos^2 x} &= \tfrac{1}{4}\pi^2 \end{aligned}$$

$$= 2.467401 \ldots$$

(OEIS A091474, A091475, and A091476), where K is Catalan's constant. A fourth integral proposed by a challenge is also trivially computable in modern versions of the Wolfram Language,

$$\int_0^{\pi/4} \frac{t^3\, d\, t}{\sin^2 t} = \frac{3}{4} \pi K - \frac{1}{64} \pi^3 + \frac{3}{32} \pi^2 \ln 2 - \frac{105}{64} \zeta(3)$$

$$= 0.3429474 \ldots$$

(OEIS A091477), where $\zeta(3)$ is Apéry's constant.

A pretty definite integral due to L. Glasser and O. Oloa (L. Glasser, pers. comm., Jan. 6, 2007) is given by

$$\int_0^{\pi/2} \frac{x^2}{x^2 + \ln^2(2 \cos x)}\, d\, x = \frac{1}{8} \pi [1 - \gamma + \ln(2 \pi)]$$

$$= 0.887759656 \ldots$$

(OEIS A127196), where γ is the Euler-Mascheroni constant. This integral (in the form considered originally by Oloa) is the $n = 1$ case of the class of integrals

$$\oint \frac{\ln^n z\, d\, z}{z \ln(1 + z)}$$

previously studied by Glasser. The closed form given above was independently found by Glasser and Oloa (L. Glasser, pers. comm., Feb. 2, 2010; O. Oloa, pers. comm., Feb. 2, 2010), and proofs of the result were subsequently published by Glasser and Manna (2008) and Oloa (2008). Generalizations of this integral have subsequently been studied by Oloa and others; see also Bailey and Borwein (2008).

An interesting class of integrals is

$$C(a) = \int_0^1 \frac{\tan^{-1}\left(\sqrt{x^2 + a^2}\right)}{\sqrt{x^2 + a^2}\, \left(x^2 + 1\right)}\, d\, x,$$

which have the special values

$$C(0) = \frac{1}{8} \pi \ln 2 + \frac{1}{2} K$$

$$C(1) = \frac{1}{4}\pi - \frac{1}{2}\pi\sqrt{2} + \frac{3}{2}\sqrt{2}\tan^{-1}\left(\sqrt{2}\right)$$

$$C\left(\sqrt{2}\right) = \frac{5}{96}\pi^2$$

(Bailey et al. 2007, pp. 42 and 60).

An amazing integral determined empirically is

$$\frac{2}{\sqrt{3}}\int_0^1 \frac{\ln^6 x \tan^{-1}\left(\frac{x\sqrt{2}}{x-2}\right)}{x+1}\,dx = \frac{1}{81\,648}\left[-229\,635\,L_3(8) + 29\,852\,550\,L_3(7)\ln 3 - 1\,632\,960\,L_3(6)\pi^2 + 27\,760\,320\,L_3(5)\zeta(3) - 275\,184\,L_3(4)\pi^4 + 36\,288\,000\,L_3(3)\zeta(5) - 30\,008\,L_3(2)\pi^6 - 57\,030\,120\,L_3(1)\zeta(7)\right],$$

where

$$L_3(s) = \sum_{n=1}^{\infty} \frac{1}{(3n-2)^s} - \frac{1}{(3n-1)^s}$$

$$= \frac{1}{3^s}\left[\zeta\left(s, \tfrac{1}{3}\right) - \zeta\left(s, \tfrac{2}{3}\right)\right]$$

(Bailey et al. 2007, p. 61).

A complicated-looking definite integral of a rational function with a simple solution is given by

$$\int_0^\infty \frac{x^8 - 4x^6 + 9x^4 - 5x^2 + 1}{x^{12} - 10x^{10} + 37x^8 - 42x^6 + 26x^4 - 8x^2 + 1}\,dx = \frac{1}{2}\pi$$

(Bailey et al. 2007, p. 258).

Another challenging integral is that for the volume of the Reuleaux tetrahedron,

$$V = \int_0^1 \left[\frac{8\sqrt{3}}{1+3t^2} - \frac{16\sqrt{2}\,(3t+1)\left(4t^2+t+1\right)^{3/2}}{\left(3t^2+1\right)\left(11t^2+2t+3\right)^2} - \frac{\sqrt{2}\,\left(249t^2+54t+65\right)}{\left(11t^2+2t+3\right)^2}\right]dt,$$

$$= \frac{8}{3}\pi - \frac{27}{4}\cos^{-1}\left(\frac{1}{3}\right) + \frac{1}{4}\sqrt{2}$$

$$= 0.4221577 \ldots$$

(OEIS A102888; Weisstein).

Integrands that look alike could provide very different results, as illustrated by the beautiful pair

$$\int_{-\infty}^{\infty} \frac{dx}{(e^x - x + 1)^2 + \pi^2} = \frac{1}{2}$$

$$\int_{-\infty}^{\infty} \frac{dx}{(e^x - x)^2 + \pi^2} = \frac{1}{1 + W(1)}$$

$$= 0.638103743 \ldots$$

due to V. Adamchik (OEIS A115287; Moll 2006; typo corrected), where $W(1)$ is the omega constant and $W(z)$ is the Lambert W-function. These can be computed using contour integration.

Computer mathematics packages also often return results much more complicated than necessary. An example of this type is provided by the integral

$$\phi(\alpha) = \int_0^{\pi} \ln\left(1 - 2\alpha\cos x + \alpha^2\right) dx = 2\pi \ln|\alpha|$$

for $\alpha \in \mathbb{R}$ and $|\alpha| > 1$ which follows from a simple application of the Leibniz integral rule (Woods 1926, pp. 143-144).

There are a wide range of methods available for numerical integration. Good sources for such techniques include Press et al. (1992) and Hildebrand (1956). The most straightforward numerical integration technique uses the Newton-Cotes formulas (also called quadrature formulas), which approximate a function tabulated at a sequence of regularly spaced intervals by various degree polynomials. If the endpoints are tabulated, then the 2- and 3-point formulas are called the trapezoidal rule and Simpson's rule, respectively. The 5-point formula is called Boole's rule. A generalization of the trapezoidal rule is romberg integration, which can yield accurate results for many fewer function evaluations.

If the analytic form of a function is known (instead of its values merely being tabulated at a fixed number of points), the best numerical method of integration is called Gaussian quadrature. By picking the optimal abscissas at which to compute the function, Gaussian quadrature produces the most accurate approximations possible. However, given the speed of modern computers, the additional complication of the Gaussian quadrature formalism often makes it less desirable than the brute-force method of simply repeatedly calculating twice as many points on a regular grid until convergence is obtained. An excellent reference for Gaussian quadrature is Hildebrand (1956).

The June 2, 1996 comic strip FoxTrot by Bill Amend (Amend 1998, p. 19; Mitchell 2006/2007) featured the following definite integral as a "hard" exam problem intended for a remedial math class but accidentally handed out to the normal class: