IIT Madras

ONLINE DEGREE

Natural numbers and integers

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

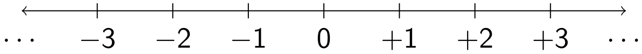
Mathematics for Data Science 1
Week 1

Natural numbers

- Numbers keep a count of objects
  - 7 represents "seven"-ness
- $1, 2, 3, 4, \ldots$
- 0 to represent no objects at all
- Natural numbers: $\mathbb{N} = \{0, 1, 2, \ldots\}$
  - Sometimes $\mathbb{N}_0$ to emphasize 0 is included
- Addition, subtraction, multiplication, division
  - Which of these always produce a natural number as the answer?

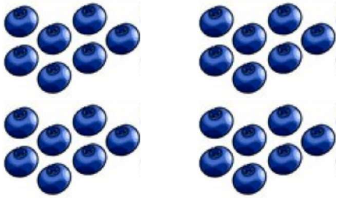
Integers

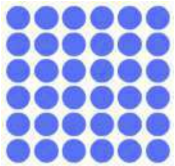
- $5 - 6$ is not a natural number
- Extend the natural numbers with negative numbers
- $-1, -2, -3, \ldots$
- Integers: $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
- Number line



Multiplication and exponentiation

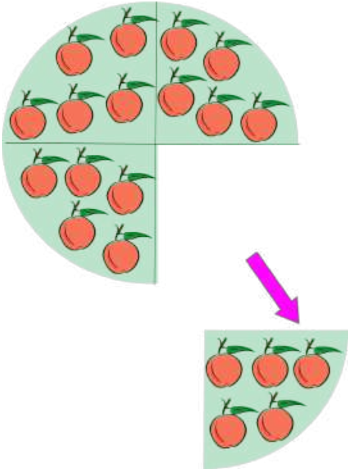
- 7×4 — make 4 groups of 7
- $m \times n = \underbrace{m + m + \cdots + m}_{n \ times}$
  - Notation: $m \times n$, $m \cdot n$, mn
  - Sign rule for multiplying negative numbers
    - $-m \times n = -(m \cdot n)$, $-m \times -n = m \cdot n$
- $m \times m = m^2$ — m squared
- $m \times m \times m = m^3$ — m cubed
- $m^k = \underbrace{m \times m \times \cdots \times m}_{k \ times}$ — m to the power k
- Multiplication is repeated addition
  Exponentiation is repeated multiplication





Division

- You have 20 mangos to distribute to 5 friends. How many do you give to each of them?
  - Give them 1 each. You have $20 - 5 = 15$ left.
  - Another round. You have $15 - 5 = 10$ left.
  - Third round. You have $10 - 5 = 5$ left.
  - Fourth round. You have $5 - 5 = 0$ left.
  - $20 \div 5 = 4$
- Division is repeated subtraction
- What if you had only 19 mangos to start with?
  - After distributing 3 to each, you have 4 left
  - Cannot distribute another round
  - The quotient of $19 \div 5$ is 3
  - The remainder of $19 \div 5$ is 4
    - $19 \bmod 5 = 4$



Factors

- a divides b if $b \bmod a = 0$
  - $a \mid b$
  - b is a multiple of a
- $4 \mid 20$, $7 \mid 63$, $32 \mid 1024$, ...
- $4 \nmid 19$, $9 \nmid 100$, ...
- a is a factor of b if $a \mid b$
- Factors occur in pairs — factors of 12 are $\{1, 12\}$, $\{2, 6\}$, $\{3, 4\}$
- ... unless the number is a perfect square — factors of 36 : $\{1, 36\}, \{2, 18\}, \{3, 12\}, \{4, 9\}, \{6\}$

Prime numbers

- p is prime if it has only two factors $\{1, p\}$
  - 1 is not a prime — only one factor
- Prime numbers are $2, 3, 5, 7, 11, 13, \ldots$
  - Sieve of Eratosthenes — remove multiples of p
- Every number can be decomposed into prime factors
  - $12 = 2 \cdot 2 \cdot 3 = 2^2 \cdot 3$
  - $126 = 2 \cdot 3 \cdot 3 \cdot 7 = 2 \cdot 3^2 \cdot 7$
- This decomposition is unique — prime factorization

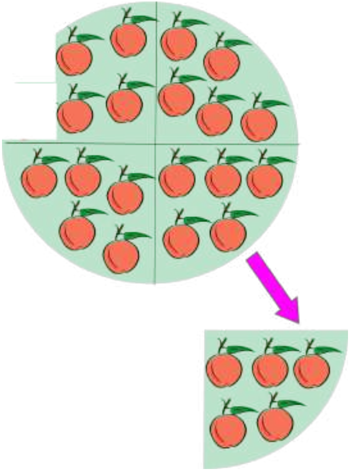
	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Summary

- $\mathbb{N}$: natural numbers $\{0, 1, 2, \ldots\}$
- $\mathbb{Z}$: integers $= \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
- Arithmetic operations: $+, -, \times, \div, m^n$
- Quotient, remainder, $a \bmod b$
- Divisibility, $a \mid b$
- Factors
- Prime numbers
- Prime factorization

Rational numbers

- Cannot represent $19 \div 5$ as an integer
- Fractions : $3\ \frac{4}{5}$
- Rational number: $\dfrac{p}{q}$, p and q are integers
  - Numerator p, denominator q
  - Use $\mathbb{Q}$ to denote rational numbers
- The same number can be written in many ways
  - $\dfrac{3}{5} = \dfrac{6}{10} = \dfrac{30}{50} = \cdots$
- Useful to add, subtract, compare rationals
  - $\dfrac{3}{5} + \dfrac{3}{4} = \dfrac{12}{20} + \dfrac{15}{20} = \dfrac{27}{20}$
  - $\dfrac{3}{5} < \dfrac{3}{4}$ because $\dfrac{12}{20} < \dfrac{15}{20}$

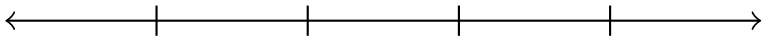


Reduced form

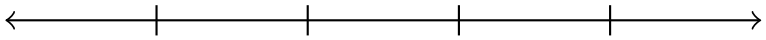
- Representation is not unique
  - $\frac{3}{5} = \frac{6}{10} = \frac{30}{50} = \cdots$
- Reduced form : $\frac{p}{q}$,
  where p,q have no common factors
  - Reduced form of $\frac{18}{60}$ is $\frac{3}{10}$
- Greatest Common Divisor: $gcd(18, 60) = 6$
  - Recall prime factorization
  - $18 = 2 \cdot 3 \cdot 3$, $60 = 2 \cdot 2 \cdot 3 \cdot 5$
  - Common prime factors are $2 \cdot 3$
  - Can find $gcd(m, n)$ more efficiently

Density

- For each integer, we have a next integer and a previous integer
  - For m, next is $m+1$, previous is $m-1$



- Next: No integer between m and $m+1$
  Previous: No integer between $m-1$ and m
- Not possible for rationals
  - Between any two rationals we can find another one
  - Suppose $\frac{m}{n} < \frac{p}{q}$
    Their average $\left(\frac{m}{n} + \frac{p}{q}\right)/2$ lies between them



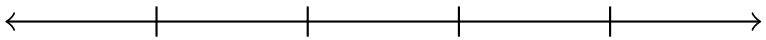
- Rationals are dense, integers are discrete

Summary

- $\mathbb{Q}$: rational numbers
- $\frac{p}{q}$, where p, q are integers
- Representation is not unique $\frac{p}{q} = \frac{n \cdot p}{n \cdot q}$
- Reduced form, $gcd(p, q) = 1$
- Rationals are dense — cannot talk of next or previous

Beyond rationals

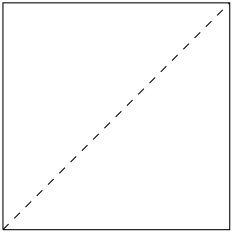
- Rational numbers are dense
  - Between any two rationals we can find another one
- Is every point on the number line a rational number?



- For an integer m, its square is $m^2 = m \cdot m$
- Square root of m, $\sqrt{m}$, is r such that $r \cdot r = m$
- Perfect squares — $1, 4, 9, 16, 25, \ldots, 256, \ldots$
- Square roots — $1, 2, 3, 4, 5, \ldots, 16, \ldots$
- What about integers that are not perfect squares?

Beyond rationals . . .

- $\sqrt{2}$ cannot be written as $\frac{p}{q}$
- Yet we can draw a line of length $\sqrt{2}$
  - Diagonal of a square whose sides have length 1
- $\sqrt{2}$ is irrational
- Real numbers: $\mathbb{R}$ — all rational and irrational numbers
- Like rationals, real numbers are dense
  - If $r < r'$, then $\frac{(r + r')}{2}$ lies between r and r'



Beyond reals

- Some well known irrational numbers
  - $\pi = 3.1415927\ldots$
  - $e = 2.7182818\ldots$
- Can we stop with real numbers?
  - What about $\sqrt{-1}$
  - For any real number r, r^2 must be positive — law of signs for multiplication
- $\sqrt{-1}$ is a complex number
- Fortunately we don't need to worry about them!

Summary

- Real numbers extend rational numbers
- Typical irrational numbers — square roots of integers that are not perfect squares
- Real numbers are dense, like rationals
- Every natural number is an integer
- Every integer is a rational number
- Every rational number is a real number
- Complex numbers extend real numbers, but we won't discuss them

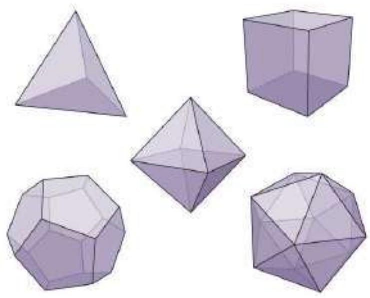
Sets

- A set is a collection of items
  - Days of the week: {Sun,Mon,Tue,Wed,Thu,Fri,Sat}
  - Factors of 24: {1,2,3,4,6,8,12,24}
  - Primes below 15: {2,3,5,7,11,13}
- Sets may be infinite
  - Different types of numbers: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$
- No requirement that members of a set have uniform type
  - Set of objects in a painting
  - Spot the dog!

Three Musicians, Pablo Picasso
MOMA, New York

Order, duplicates, cardinality

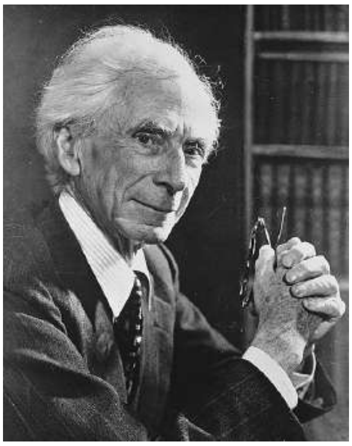
- Sets are unordered
  - {Kohli, Dhoni, Pujara}
  - {Pujara, Kohli, Dhoni}
- Duplicates don't matter (unfortunately?)
  - {Kohli, Dhoni, Pujara, Kohli}
- Cardinality: number of items in a set
  - For finite sets, count the items
    - {1,2,3,4,6,8,12,24} has cardinality 8
    - May not be obvious that a set is finite
  - What about infinite sets?
    - Is $\mathbb{Q}$ bigger than $\mathbb{Z}$?
    - Is $\mathbb{R}$ bigger than $\mathbb{Q}$?
    - Separate discussion



The Platonic solids
Set of cardinality 5
Wikimedia

Describing sets, membership

- Finite sets can be listed out explicitly
  - {Kohli, Dhoni, Pujara}
  - {1,2,3,4,6,8,12,24}
- Infinite sets cannot be listed out
  - $\mathbb{N} = \{0, 1, 2, \ldots\}$ is not formal notation
- Not every collection of items is a set
  - Collection of all sets is not a set
  - Russell's Paradox: Separate discussion
- Items in a set are called elements
  - Membership: $x \in X$, x is an element of X
  - $5 \in \mathbb{Z}$, $\sqrt{2} \notin \mathbb{Q}$

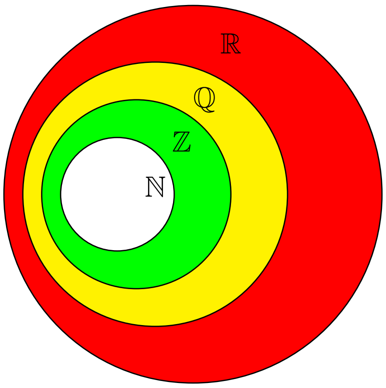


Bertrand Russell
© *Dutch National Archives*

Subsets

- X is a subset of Y
  Every element of X is also an element of Y
- Notation: $X \subseteq Y$
- Examples
  - {Kolhi,Pujara} $\subseteq$ {Kohli,Dhoni,Pujara}
  - Primes $\subseteq \mathbb{N}$, $\mathbb{N} \subseteq \mathbb{Z}$, $\mathbb{Z} \subseteq \mathbb{Q}$, $\mathbb{Q} \subseteq \mathbb{R}$
- Every set is a subset of itself: $X \subseteq X$
  - $X = Y$ if and only if $X \subseteq Y$ and $Y \subseteq X$
- Proper subset: $X \subseteq Y$ but $X \neq Y$
  - Notation: $X \subset Y$, $X \subsetneq Y$
  - $\mathbb{N} \subsetneq \mathbb{Z}$, $\mathbb{Z} \subsetneq \mathbb{Q}$, $\mathbb{Q} \subsetneq \mathbb{R}$,

Venn Diagram



The empty set and the powerset

- The empty set has no elements — $\emptyset$
- $\emptyset \subseteq X$ for every set X
  - Every element of $\emptyset$ is also in X
- A set can contain other sets
- Powerset — set of subsets of a set
  - $X = \{a, b\}$
  - Powerset is $\{\emptyset, \{a\}, \{b\}, \{a, b\}\}$
- Set with n elements has 2^n subsets
  - $X = \{x_1, x_2, \ldots, x_n\}$
  - In a subset, either include or exclude each x_i
  - 2 choices per element, $\underbrace{2 \cdot 2 \cdots 2}_{n \text{ times}} = 2^n$ subsets

Subsets and binary numbers

- $X = \{x_1, x_2, \ldots, x_n\}$
- n bit binary numbers
  - 3 bits: 000, 001, 010, 011, 100, 101, 110, 111
- Digit i represents whether x_i is included in a subset
  - $X = \{a, b, c, d\}$
  - 0101 is $\{b, d\}$
  - 0000 is $\emptyset$, 1111 is X
- 2^n n bit numbers

Constructing subsets

Set comprehension

- The subset of even integers

  $\{x \mid x \in \mathbb{Z}, x \bmod 2 = 0\}$

  - Begin with an existing set, $\mathbb{Z}$
  - Apply a condition to each element in that set
    - $x \in \mathbb{Z}$ such that $x \bmod 2 = 0$
  - Collect all the elements that match the condition

- Examples
  - The set of perfect squares

    $\{m \mid m \in \mathbb{N}, \sqrt{m} \in \mathbb{N}\}$
  - The set of rationals in reduced form

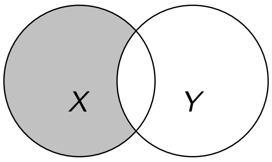
    $\{p/q \mid p, q \in \mathbb{Z}, gcd(p, q) = 1\}$

Intervals

- Integers from -6 to $+6$
  $\{z \mid z \in \mathbb{Z}, -6 \leq z \leq 6\}$
- Real numbers between 0 and 1
- Closed interval $[0, 1]$
  – include endpoints
  $\{r \mid r \in \mathbb{R}, 0 \leq r \leq 1\}$
- Open interval $(0, 1)$
  – exclude endpoints
  $\{r \mid r \in \mathbb{R}, 0 < r < 1\}$
- Left open $(0, 1]$
  $\{r \mid r \in \mathbb{R}, 0 < r \leq 1\}$

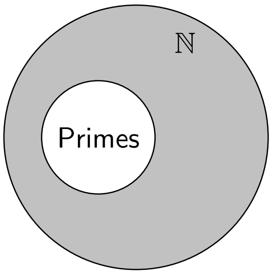
Union, intersection, complement

- Union — combine X and Y, $X \cup Y$
  $\{a, b, c\} \cup \{c, d, e\} = \{a, b, c, d, e\}$
- Intersection — elements common to X and Y, $X \cap Y$
  $\{a, b, c, d\} \cap \{a, d, e, f\} = \{a, d\}$
- Set difference — elements in X that are not in Y, $X \setminus Y$ or $X - Y$
  $\{a, b, c, d\} \setminus \{a, d, e, f\} = \{b, c\}$



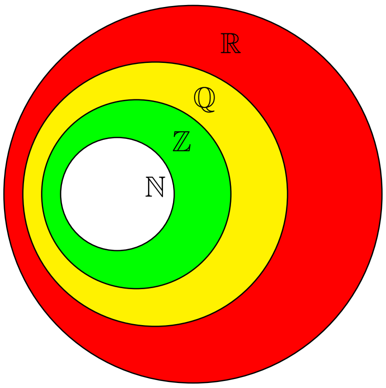
Union, intersection, complement

- Union — combine X and Y, $X \cup Y$
  $\{a, b, c\} \cup \{c, d, e\} = \{a, b, c, d, e\}$
- Intersection — elements common to X and Y, $X \cap Y$
  $\{a, b, c, d\} \cap \{a, d, e, f\} = \{a, d\}$
- Set difference — elements in X that are not in Y, $X \setminus Y$ or $X - Y$
  $\{a, b, c, d\} \setminus \{a, d, e, f\} = \{b, c\}$
- Complement — elements not in X, $\overline{X}$ or X^c
  - Define complement relative to larger set, *universe*
  - Complement of prime numbers in $\mathbb{N}$ are composite numbers



Summary

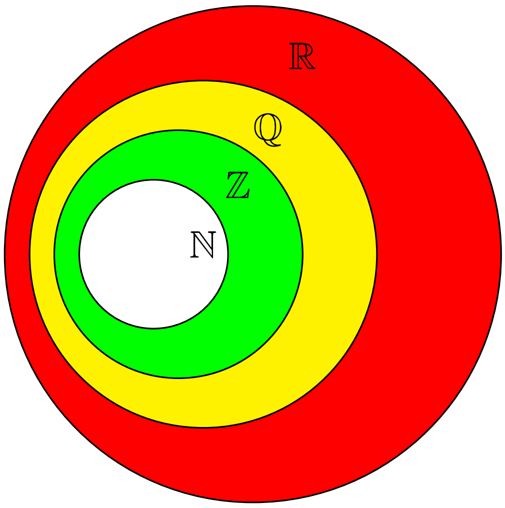
- Sets are a standard way to represent collections of mathematical objects
- Sets may be finite or infinite
- Can carve out interesting subsets of sets
- Set operations: union, intersection, difference, complement



Sets

- A set is a collection of items
- Finite sets can be listed out explicitly
  - {Kohli, Dhoni, Pujara}, {1,4,9,16,25}
- Infinite sets cannot be listed out
  - $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ not formal
- Membership $x \in X$, Subset $X \subseteq Y$
  - $5 \in \mathbb{Z}$, $\sqrt{2} \notin \mathbb{Q}$
  - Primes $\subseteq \mathbb{N}$, $\mathbb{N} \subseteq \mathbb{Z}$, $\mathbb{Z} \subseteq \mathbb{Q}$, $\mathbb{Q} \subseteq \mathbb{R}$
- Powerset — set of subsets of a set
  - $X = \{a, b\}$, powerset $\{\emptyset, \{a\}, \{b\}, \{a, b\}\}$
  - Set with n elements has 2^n subsets

Venn Diagram



Set Comprehension

- Squares of the even integers
  $\{x^2 \mid x \in \mathbb{Z}, x \bmod 2 = 0\}$
  - Generate Elements drawn from existing set
  - Filter Select elements that satisfy a constraint
  - Transform Modify selected elements
- More filters
  - Rationals in reduced form
    $\{p/q \mid p/q \in \mathbb{Q}, gcd(p,q) = 1\}$
  - Reals in interval $[-1, 2)$
    $\{r \mid r \in \mathbb{R}, -1 \leq r < 2\}$

$\{0, 4, 16, 36, 64, 100, 144, 196, 256, \ldots\}$

$\cdots$	-2	-1	0	1	2	3	4	5	$\cdots$
$\cdots$	-2		0		2		4		$\cdots$
$\cdots$	4		0		4		16		$\cdots$

Set Comprehension . . .

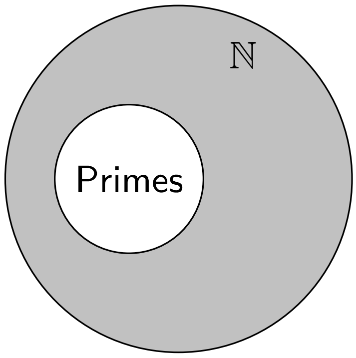
- Cubes of first 5 natural numbers
  $Y = \{n^3 \mid n \in \{0, 1, 2, 3, 4\}\}$
- Cubes of first 500 natural numbers?
  $Y = \{n^3 \mid n \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, \dots, 498, 499\}\}$
- Use set comprehension to define first 500 natural numbers
  $X = \{n \mid n \in \mathbb{N}, n < 500\}$
- Now, a more readable version
  $X = \{n \mid n \in \mathbb{N}, n < 500\}$
  $Y = \{n^3 \mid n \in X\}$

Perfect squares

- Integers whose square root is also an integer
  $\{z \mid z \in \mathbb{Z}, \sqrt{z} \in \mathbb{Z}\}$
- All squares are positive, so this is the same as
  $\{n \mid n \in \mathbb{N}, \sqrt{n} \in \mathbb{N}\}$
- Alternatively, generate all the perfect squares
  $\{n^2 \mid n \in \mathbb{N}\}$
- Extend the definition to rationals
  - $\frac{9}{16} = \left(\frac{3}{4}\right)^2$ is a square, $\frac{1}{2} \neq \left(\frac{p}{q}\right)^2$ for any p, q is not
  - $\{q \mid q \in \mathbb{Q}, \sqrt{q} \in \mathbb{Q}\}$, or $\{q^2 \mid q \in \mathbb{Q}\}$
- Choose the generator as required

Union, intersection, complement

- Union — combine X and Y, $X \cup Y$
  $\{a, b, c\} \cup \{c, d, e\} = \{a, b, c, d, e\}$
- Intersection — elements common to X and Y, $X \cap Y$
  $\{a, b, c, d\} \cap \{a, d, e, f\} = \{a, d\}$
- Set difference — elements in X that are not in Y, $X \setminus Y$ or $X - Y$
  $\{a, b, c, d\} \setminus \{a, d, e, f\} = \{b, c\}$
- Complement — elements not in X, $\overline{X}$ or X^c
  - Define complement relative to larger set, *universe*
  - Complement of prime numbers in $\mathbb{N}$ are composite numbers



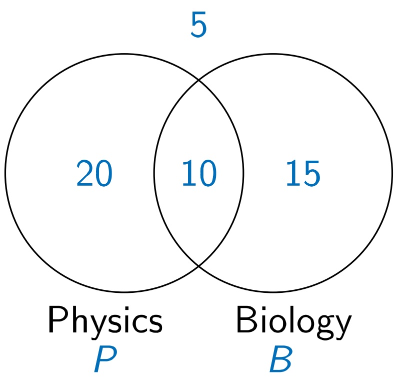
Counting problems

- In a class, 30 students took Physics, 25 took Biology and 10 took both, and 5 took neither. How many students are there in the class?
  - Draw sets for Physics (P) and Biology (Q)
  - 10 students are in $P \cap B$
  - This leaves 20 students in $P \setminus B$

    Took Physics, but did not take Biology
  - Likewise 15 students in $B \setminus P$

    Took Biology, but did not take Physics
  - 5 students in $\overline{P \cup B}$

    In the class, but took neither Physics nor Biology
- Class strength: $5 + 20 + 10 + 15 = 50$

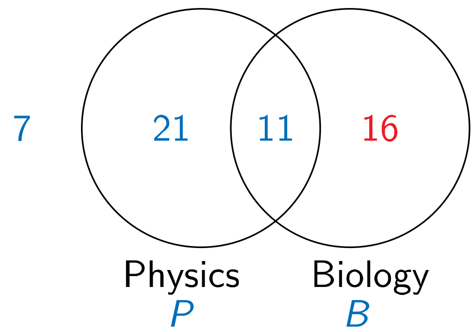


Counting problems

- In a class of 55 students, 32 students took Physics, 11 took both Physics and Biology, and 7 took neither.

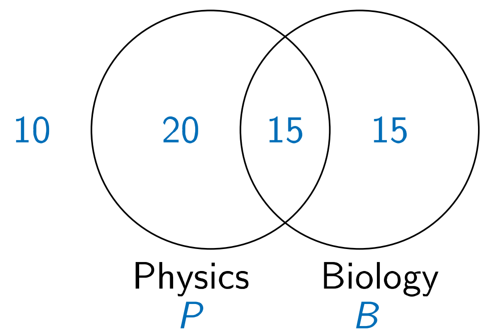
  How many students took Biology but not Physics?

  - $7 + 21 + 11 + x = 55$
  - $x = 55 - 39 = 16$



- In a class of 60 students, 35 students took Physics, 30 took Biology, and 10 took neither. How many took both Physics and Biology?

  - $|Y|$: Cardinality of Y (number of elements)
  - $|P| + |B| = 35 + 30 = 65$
  - $|P \cup B| = 60 - 10 = 50$
  - So $65 - 50 = 15$ must have taken both



Summary

- Set notation is useful way to concisely describe collections of objects
- Set comprehension combines generators, filters and tranformations to produce new sets from old
- Venn diagrams can be useful to work out problems involving sets

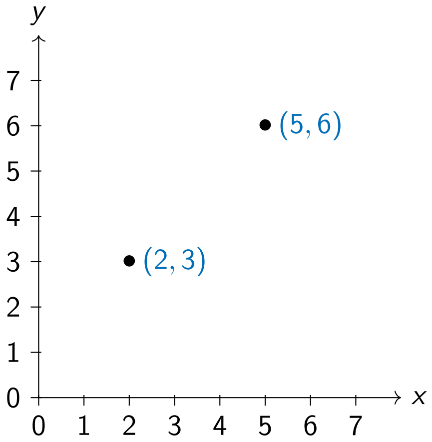
New sets from old

- A set is a collection of items
- We can combine sets to form new ones
  - $X \cup Y$, $X \cap Y$, $X \setminus Y$
  - $\overline{X}$ with respect to Y
- Define subsets using set comprehension
  - Odd integers
    $\{z \mid z \in \mathbb{Z}, z \bmod 2 = 1\}$
  - Rationals not in reduced form
    $\{p/q \mid p, q \in \mathbb{Z}, gcd(p, q) > 1\}$
  - Reals in $[3, 17)$
    $\{r \mid r \in \mathbb{R}, 3 \leq r < 17\}$

"New lamps for old"
Aladdin's Picture Book
Walter Crane (1876)

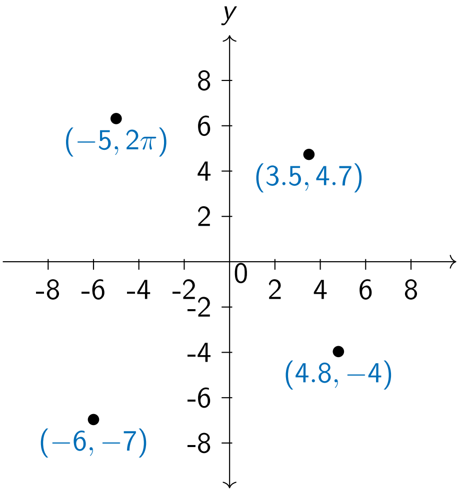
Cartesian product

- $A \times B = \{(a, b) \mid a \in A, b \in B\}$
  - Pair up elements from A and B
  - $A = \{0, 1\}$, $B = \{2, 3\}$
  - $A \times B = \{(0, 2), (0, 3), (1, 2), (1, 3)\}$
- In a pair, the order is important
  - $(0, 1) \neq (1, 0)$
- For sets of numbers, visualize product as two dimensional space
  - $\mathbb{N} \times \mathbb{N}$
  - $\mathbb{R} \times \mathbb{R}$



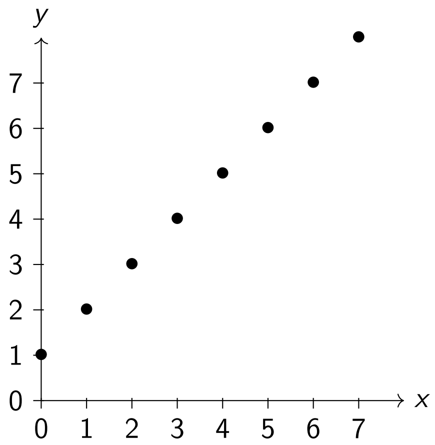
Cartesian product

- $A \times B = \{(a,b) \mid a \in A, b \in B\}$
  - Pair up elements from A and B
  - $A = \{0,1\}$, $B = \{2,3\}$
  - $A \times B = \{(0,2),(0,3),(1,2),(1,3)\}$
- In a pair, the order is important
  - $(0,1) \neq (1,0)$
- For sets of numbers, visualize product as two dimensional space
  - $\mathbb{N} \times \mathbb{N}$
  - $\mathbb{R} \times \mathbb{R}$



Binary relations

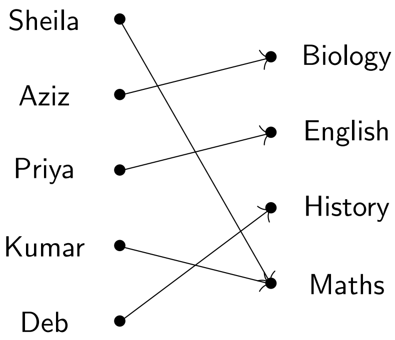
- Select some pairs from the Cartesian product
- Combine Cartesian product with set comprehension
- $\{(m,n) \mid (m,n) \in \mathbb{N} \times \mathbb{N}, n = m+1\}$
  - $\{(0,1),(1,2),(2,3),\dots,(17,18),\dots\}$
- Pairs (d,n) where d is a factor of n
  - $\{(d,n) \mid (d,n) \in \mathbb{N} \times \mathbb{N}, d|n\}$
  - $\{(1,1),\dots,(2,82),\dots,(14,56),\dots\}$
- Binary relation $R \subseteq A \times B$
- Notation: $(a,b) \in R$, $a \; R \; b$



More relations

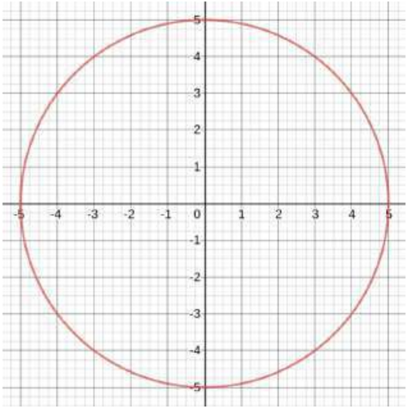
- Teachers and courses
  - T, set of teachers in a college
    C, set of courses being offered
  - $A \subseteq T \times C$ describes the allocation of teachers to courses
  - $A = \{(t, c) \mid (t, c) \in T \times C, t \text{ teaches } c\}$
- Mother and child
  - P, set of people in a country
  - $M \subseteq P \times P$ relates mothers to children
  - $M = \{(m, c) \mid \ (m, c) \in P \times P, \ m \text{ is the mother of } c\}$

A relation as a graph



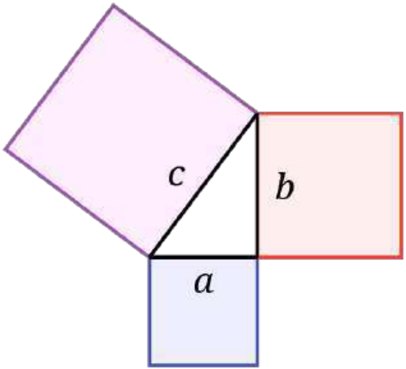
More relations

- Points at distance 5 from $(0,0)$
  - Distance from $(0,0)$ to (a,b) is $\sqrt{a^2+b^2}$
  - $\{(a,b) \mid (a,b) \in \mathbb{R}\times\mathbb{R}, \sqrt{a^2+b^2} = 5\}$
  - $\{(0,5),(5,0),(3,4),(-3,-4),\ldots\}$
  - A circle with centre at $(0,0)$
- Rationals in reduced form
  - A subset of Q
    - $\{p/q \mid (p,q) \in \mathbb{Z}\times\mathbb{Z}, gcd(p,q) = 1\}$
  - . . . but also a relation on $\mathbb{Z}\times\mathbb{Z}$
    - $\{(p,q) \mid (p,q) \in \mathbb{Z}\times\mathbb{Z}, gcd(p,q) = 1\}$



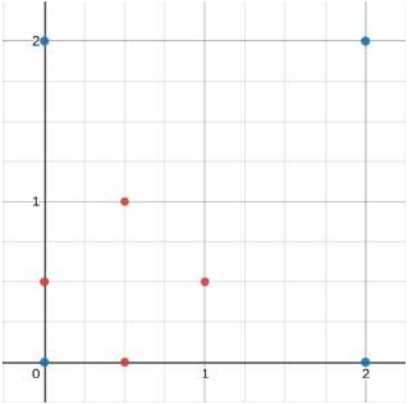
Beyond binary relations

- Cartesian products of more than two sets
- Pythagorean triples
  - Square on the hypotenuse is the sum of the squares on the opposite sides
  - $\{(a, b, c) \mid (a, b, c) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}, a, b, c > 0, a^2 + b^2 = c^2\}$



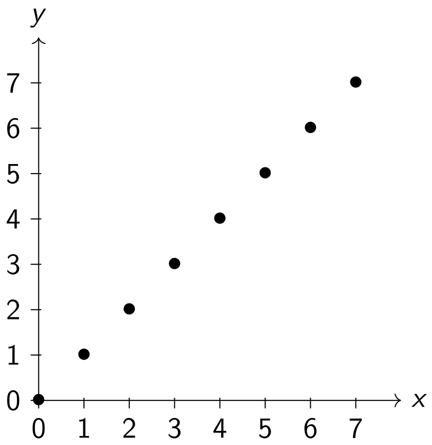
Beyond binary relations

- Cartesian products of more than two sets
- Pythagorean triples
  - Square on the hypotenuse is the sum of the squares on the opposite sides
  - $\{(a, b, c) \mid \ (a, b, c) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}, \ a, b, c > 0, a^2 + b^2 = c^2\}$
- Corners of squares
  - A corner is a point $(x, y) \in \mathbb{R} \times \mathbb{R}$
  - $((x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4))$ are related if they are four corners of a square
  - For instance:
    - $((0, 0), (0, 2), (2, 2), (2, 0))$
    - $((0, 5, 0), (0, 0.5), (0.5, 1), (1, 0.5))$
  - $Sq \subset \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$



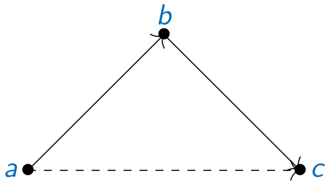
Back to binary relations

- Identity relation $I \subseteq A \times A$
  - $I = \{(a,b) \mid (a,b) \in A \times A, a = b\}$
  - $I = \{(a,a) \mid (a,a) \in A \times A\}$
  - $I = \{(a,a) \mid a \in A\}$
- Reflexive relations
  - $R \subseteq A \times A$, $I \subseteq R$
  - $\{(a,b) \mid (a,b) \in \mathbb{N} \times \mathbb{N}, a, b > 0, a|b\}$
    - $a|a$ for all $a > 0$
- Symmetric relations
  - $(a,b) \in R$ if and only if $(b,a) \in R$
  - $\{(a,b) \mid (a,b) \in \mathbb{N} \times \mathbb{N}, gcd(a,b) = 1\}$
  - $\{(a,b) \mid (a,b) \in \mathbb{N} \times \mathbb{N}, |a-b| = 2\}$



Back to binary relations ...

- Transitive relations
  - If $(a,b) \in R$ and $(b,c) \in R$ then $(a,c) \in R$
  - $\{(a,b) \mid (a,b) \in \mathbb{N} \times \mathbb{N}, a|b\}$
    - If $a|b$ and $b|c$ then $a|c$
  - $\{(a,b) \mid (a,b) \in \mathbb{R} \times \mathbb{R}, a < b\}$
    - If $a < b$ and $b < c$ then $a < c$



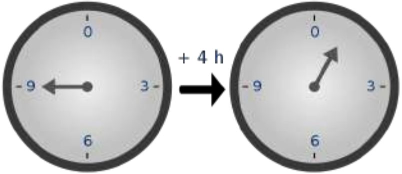
- Antisymmetric relations
  - If $(a,b) \in R$ and $a \neq b$, then $(b,a) \notin R$
  - $\{(a,b) \mid (a,b) \in \mathbb{R} \times \mathbb{R}, a < b\}$
    - If $a < b$ then $b \not< a$
  - $M \subseteq P \times P$ relates mothers to children
    - If $(p,c) \in M$ then $(c,p) \notin M$

Equivalence relations

- Reflexive, symmetric and transitive
- Same remainder modulo 5
  - $7 \bmod 5 = 2$, $22 \bmod 5 = 2$
  - If $a \bmod 5 = b \bmod 5$ then $(b - a)$ is a multiple of 5
  - $$\mathbb{Z}Mod5 = \{(a, b) \mid \ a, b \in \mathbb{Z}, (b - a) \bmod 5 = 0\}$$
  - Divides integers into 5 groups based on remainder when divided by 5
- An equivalence relation partitions a set
- Groups of equivalent elements are called equivalence classes

Measuring time

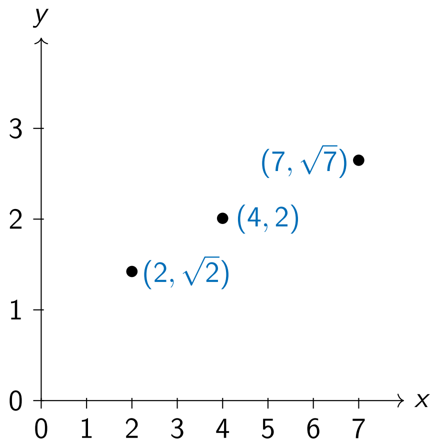
Clock displays hours modulo 12



2:00 am is equivalent to 2:00 pm

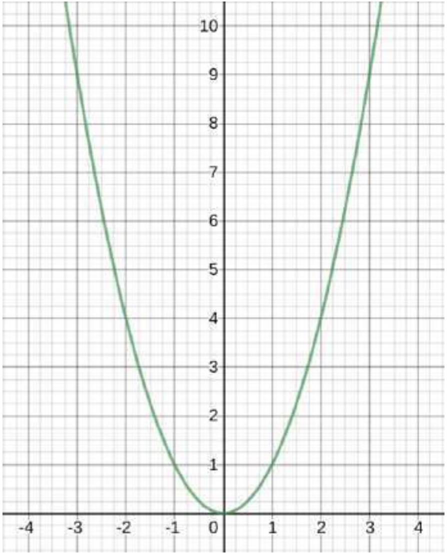
Summary

- Cartesian products generate n-tuples from n sets
  - $(x_1, x_2, \ldots, x_n) \in X_1 \times X_2 \times \cdots \times X_n$
- A relation picks out a subset of a Cartesian product
  - $\{(m, r) \mid (m, r) \in \mathbb{N} \times \mathbb{R}, r = \sqrt{m}\}$
- Properties of relations
  - Reflexive, symmetric, transitive, antisymmetric
- Equivalence relations partition a set



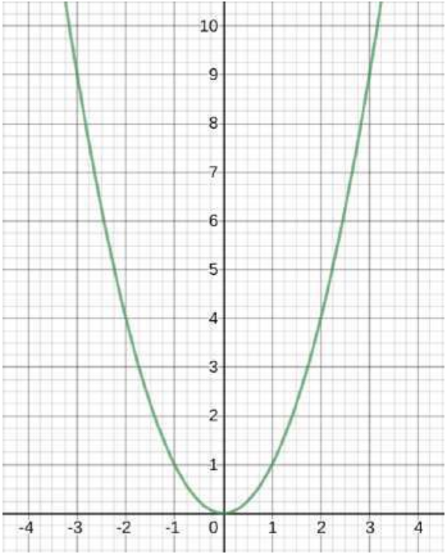
Functions

- A rule to map inputs to outputs
- Convert x to x^2
  - The rule: $x \mapsto x^2$
  - Give it a name: $sq(x) = x^2$
    - Input is a parameter
- Need to specify the input and output sets
  - Domain: Input set
    - $domain(sq) = \mathbb{R}$
  - Codomain: Output set of possible values
    - $codomain(sq) = \mathbb{R}$
  - Range: Actual values that the output can take
    - $range(sq) = \mathbb{R}_{\geq 0} = \{r \mid r \in \mathbb{R}, r \geq 0\}$
- $f : X \rightarrow Y$, domain of f is X, codomain is Y



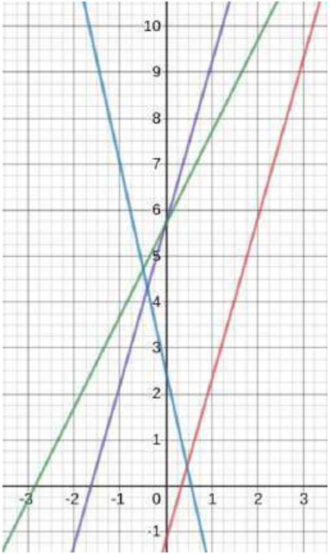
Functions and relations

- Associate a relation R_f with each function f
- $R_{sq} = \{(x, y) \mid x, y \in \mathbb{R}, y = x^2\}$
  - Additional notation: $y = x^2$
- $R_f \subset domain(f) \times range(f)$
- Properties of R_f
  - Defined on the entire domain
    - For each $x \in domain(f)$, there is a pair $(x, y) \in R_f$
  - Single-valued
    - For each $x \in domain(f)$, there is exactly one $y \in codomain(f)$ such that $(x, y) \in R_f$
- Drawing f as a graph is plotting R_f



Lines

- $f(x) = 3.5x + 5.7$
  - 3.5 is the slope
  - 5.7 is intercept where the line crosses the y=axis, whe $x = 0$
- Changing the slope and intercept produce different lines
  - $f(x) = 3.5x - 1.2$
  - $f(x) = 2x + 5.7$
  - $f(x) = -4.5x + 2.5$
- In all these cases
  - Domain $= \mathbb{R}$
  - Codomain $=$ Range $= \mathbb{R}$

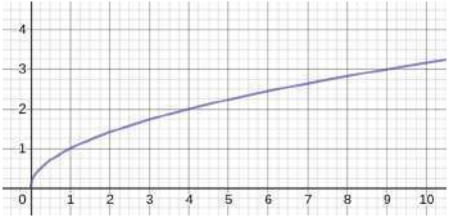


More functions

- $x \mapsto \sqrt{x}$

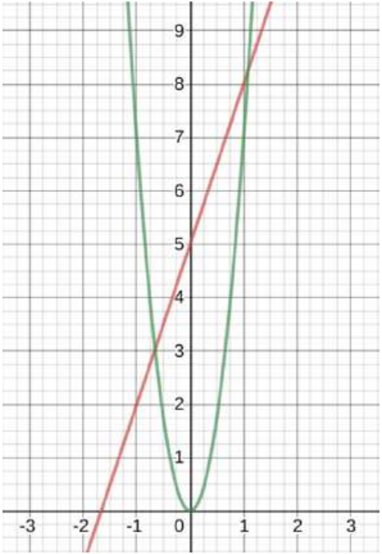
- Is this a function?
  - $5^2 = (-5)^2 = 25$
  - $\sqrt{25}$ gives two options
  - By convention, take positive square root

- What is the domain?
  - Depends on codomain
  - Negative numbers do not have real square roots
  - If codomain is $\mathbb{R}$, domain is $\mathbb{R}_{\geq 0}$
  - If codomain is the set $\mathbb{C}$ of complex numbers, domain is $\mathbb{R}$



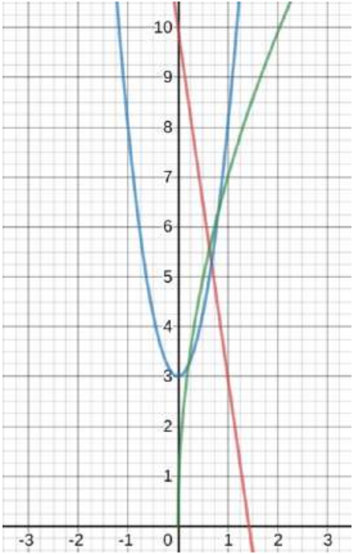
Types of functions

- Injective: Different inputs produces different outputs — one-to-one
  - If $x_1 \neq x_2$, $f(x_1) \neq f(x_2)$
  - $f(x) = 3x + 5$ is injective
  - $f(x) = 7x^2$ is not: for any a, $f(a) = f(-a)$



Types of functions

- Injective: Different inputs produces different outputs — one-to-one
  - If $x_1 \neq x_2$, $f(x_1) \neq f(x_2)$
  - $f(x) = 3x + 5$ is injective
  - $f(x) = 7x^2$ is not: for any a, $f(a) = f(-a)$
- Surjective: Range is the codomain — onto
  - For every $y \in codomain(f)$, there is an $x \in domain(f)$ such that $f(x) = y$
  - $f(x) = -7x + 10$ is surjective
  - $f(x) = 5x^2 + 3$ is not surjective for codomain $\mathbb{R}$
  - $f(x) = 7\sqrt{x}$ is not surjective for codomain $\mathbb{R}$



- Bijective: $1-1$ correspondence between domain and codomain
  - Every $x \in domain(f)$ maps to a distinct $y \in codomain(f)$
  - Every $y \in codomain(f)$ has a unique pre-image $x \in domain(f)$ such that $y = f(x)$

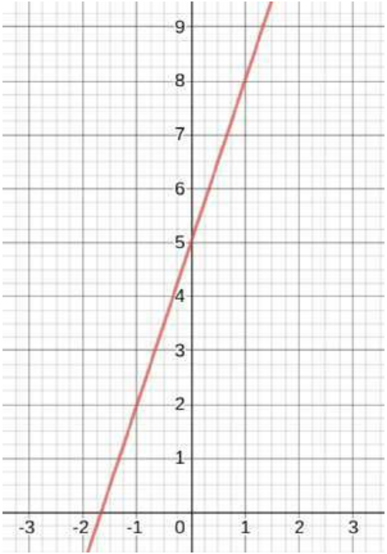
Theorem

A function is bijective if and only if it is injective and surjective

- From the definition, if a function is bijective it is injective and surjective
- Suppose a function f is injective and surjective
  - Injectivity guarantees that f satisfies the first condition of a bijection.
  - Surjectivity says every $y \in codomain(f)$ has a pre-image. Injectivity guarantees this pre-image is unique.

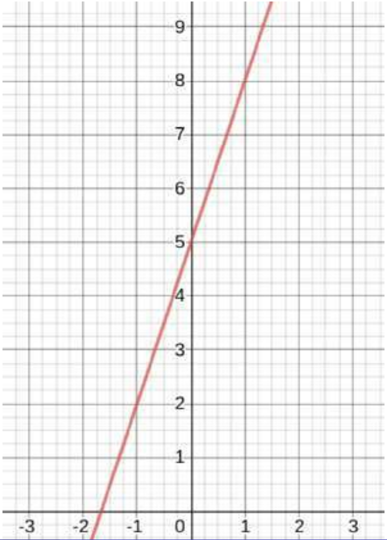
Bijections and cardinality

- For finite sets we can count the items
- What if we have two large sacks filled with marbles?
  - Do we need to count the marbles in each sack?
  - Pull out marbles in pairs, one from each sack
  - Do both sacks become empty simultaneously?
  - Bijection between the marbles in the sacks
- For infinite sets
  - Number of lines is the same as $\mathbb{R} \times \mathbb{R}$
  - Every line $y = mx + c$ is determined uniquely by (m, c) and vice versa



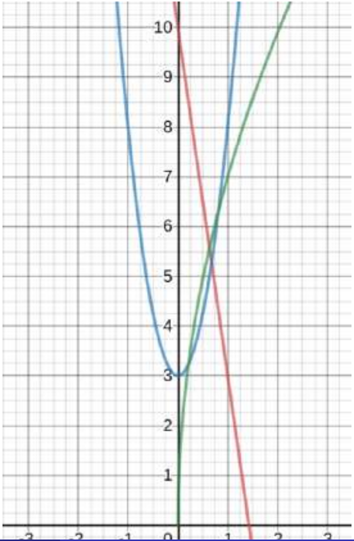
Bijections and cardinality . . .

- For every pair of points (x_1, y_1) and $(x_2.y_2)$, there is a unique line passing through both points
- Number of lines is same as cardinality of $\mathbb{R} \times \mathbb{R}$
- Does this show that $(\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) = \mathbb{R}^2 \times \mathbb{R}^2$ has the same cardinality as $\mathbb{R} \times \mathbb{R}$?
- The correspondence is not a bijection — many pairs of points describe the same line
- Be careful to establish that a function is a bijection



Summary

- A function is given by a rule mapping inputs to outputs
- Define the domain, codomain and range
- Associate a relation R_f with each function f
- Properties of functions: injective (one-to-one), surjective (onto)
- Bijections: injective and surjective (one-to-one and onto)
- A bijection establishes that domain and codomain have same cardinality



Relations

- $A \times B$ — Cartesian product, all pairs (a, b), $a \in A$ and $b \in B$
- $A = \{1, 4, 7\}$, $B = \{1, 16, 49\}$
  - $A \times B = \{(1,1),(1,16),(1,49),(4,1),(4,16),(4,49),(7,1),(7,16),(7,49)\}$
  - $B \times A = \{(1,1),(16,1),(49,1),(1,4),(16,4),(49,4),(1,7),(16,7),(49,7)\}$
  - $B \times B = \{(1,1),(1,16),(1,49),(16,1),(16,16),(16,49),(49,1),(49,16),(49,49)\}$
- Can take Cartesian product of more than two sets
  - $A \times B \times A =$ $\{(1,1,1),(1,1,4),(1,1,7),(1,16,1),(1,16,7),\ldots,(7,49,1),(7,49,16),(7,49,49)\}$
- A relation picks out certain tuples in the Cartesian product
  - $S \subseteq A \times B = \{(1,1),(4,16),(7,49)\}$
  - $S = \{(a,b) \mid (a,b) \in A \times B, b = a^2\}$

Examples of relations

- Divisibility
  - Pairs of natural numbers (d, n) such that $d|n$
  - Pairs such as $(7, 63), (17, 85), (3, 9), \ldots$
  - $D = \{(d, n) \mid (d, n) \in \mathbb{N} \times \mathbb{N}, d|n\}$
  - Can also extend to integer divisors
  - $E = \{(d, n) \mid (d, n) \in \mathbb{Z} \times \mathbb{N}, d|n\}$
  - Now $(-7, 63), (-17, 85), (-3, 9), \ldots$ are also in E
- Prime powers
  - Pairs of natural numbers (p, n) such that p is prime and $n = p^m$ for some natural number m
  - Examples: $(3, 1), (5, 625), (7, 343), \ldots$
  - First define primes: $P = \{p \mid p \in \mathbb{N}, \text{factors}(p) = \{1, p\}, p \neq 1\}$
  - Prime powers: $PP = \{(p, n) \mid (p, n) \in P \times \mathbb{N}, n = p^m \text{ for some } m \in \mathbb{N}\}$

Beyond numbers

Airline routes

- An airline flies to set of cities — e.g. Bangalore, Chennai, Delhi, Kolkata, ...
- Let C denote the set of cities served by the airline
- Some cities are connected by direct flights
- $D \subseteq C \times C$
- Is D reflexive, irreflexive?
  - Hopefully irreflexive!
- Is D symmetric?
  - If there is a direct flight from Bangalore to Delhi, is there always a direct flight back from Delhi to Bangalore
  - For bigger cities, yes
  - For smaller cities, may have a triangular route
    Chennai $\rightarrow$ Madurai $\rightarrow$ Salem $\rightarrow$ Chennai

Tables as relations

- Flying distances between cities

Source	**Destination**	**Distance (km)**
Bangalore	Chennai	290
Chennai	Delhi	1752
Delhi	Bangalore	1735
Delhi	Chennai	1752
...	...	...

- Table is a relation: $Dist \subseteq C \times C \times \mathbb{N}$
- Some entries are useless: (Delhi,Delhi,0)
- Restrict to cities served by direct direct flights
  $Dist = \{(a, b, d) \mid (a, b) \in D, d \text{ is distance from } a \text{ to } b\}$
- Distances are symmetric, even if D is not
- Save space by representing only one direction in the table

Tables as relations . . .

Roll no	Name	Date of birth)
A71396	Abhay Shah	03-07-2001
B82976	Payal Ghosh	18-06-1999
F98989	Jeremy Pinto	22-02-2003
C93986	Payal Ghosh	14-05-2000
. . .	. . .	. . .

- Some columns are special — each student has a unique roll number
  - Such a column is called a key
  - Name is not a key, in general
- Given the roll number, can retrieve the data for a student
  - Function from Roll Numbers to (Name, Date of Birth)
  - (key,value) pairs

Operations on relations

Roll No	Name	Date of birth)
A71396	Abhay Shah	03-07-2001
B82976	Payal Ghosh	18-06-1999
F98989	Jeremy Pinto	22-02-2003
C93986	Payal Ghosh	14-05-2000
...	...	...

Roll no	Subject	Grade)
A71396	English	B
B82976	Mathematics	A
C93986	Physics	B
B82976	Chemistry	A
...	...	...

- Generate a table with roll numbers, names and grades
- Join the relations on Roll No
- $\{(r, n, s, g) \mid (r, n, d) \in \text{Students}, (r', s, g) \in \text{Grades}, r = r'\}$

Roll No	Name	Subject	Grade)
A71396	Abhay Shah	English	B
B82976	Payal Ghosh	Mathematics	A
B82976	Payal Ghosh	Chemistry	A
C93986	Payal Ghosh	Physics	B
...	...	...	...

Summary

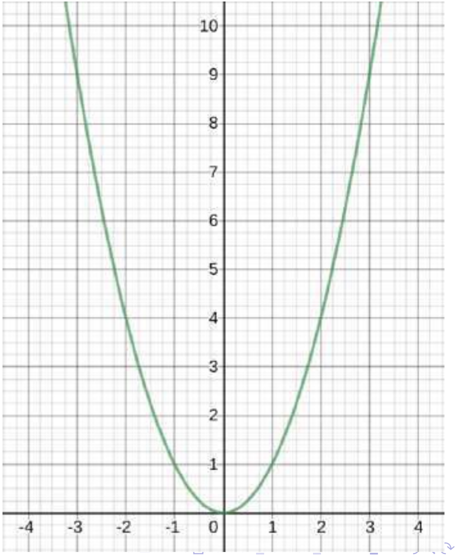
- A relation describes special tuples in a Cartesian product
- Data tables are essentially relations
- Combining information on tables can be described in terms of operations on relations

Functions

- A rule to map inputs to outputs
  - $x \mapsto x^2$, $g(x) = x^2$
- Domain, codomain, range
- Associated relation

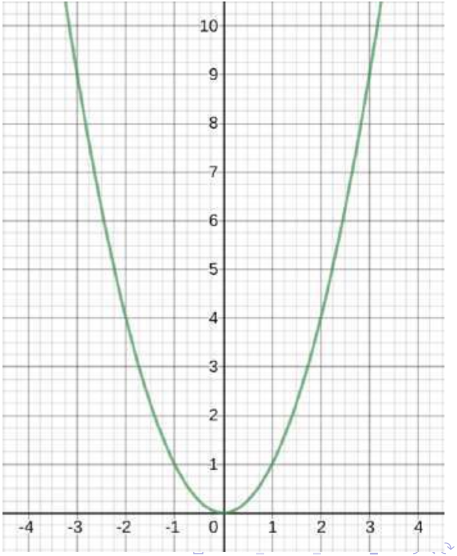
  $R_{sq} = \{(x,y) \mid x, y \in \mathbb{R}, y = x^2\}$
- Can have functions on other sets:

  Mother: People $\rightarrow$ *People*
- Will focus more on functions on numbers
- What questions are we interested in?



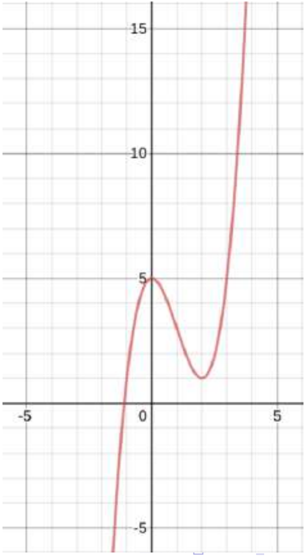
Range of values

- What range of values does the output span
- $f(x) = x^2$ is always positive, range is 0 to $+\infty$



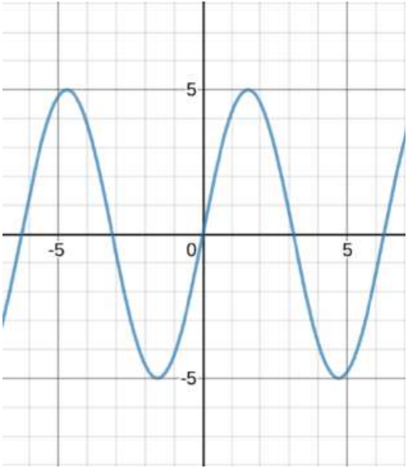
Range of values

- What range of values does the output span
- $f(x) = x^2$ is always positive, range is 0 to $+\infty$
- $f(x) = x^3 - 3x^2 + 5$ ranges from $-\infty$ to $+\infty$



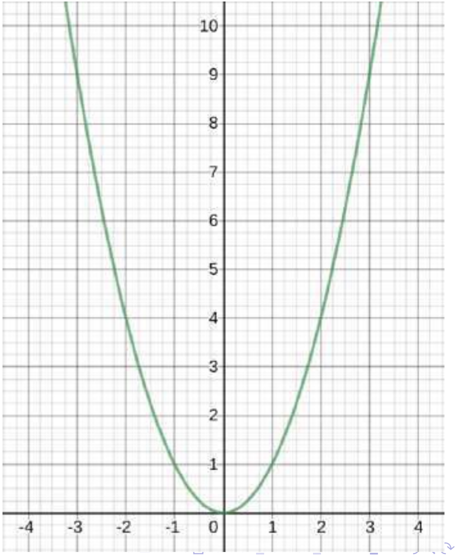
Range of values

- What range of values does the output span
- $f(x) = x^2$ is always positive, range is 0 to $+\infty$
- $f(x) = x^3 - 3x^2 + 5$ ranges from $-\infty$ to $+\infty$
- $f(x) = 5\sin(x)$ has a bounded range, from -5 to $+5$



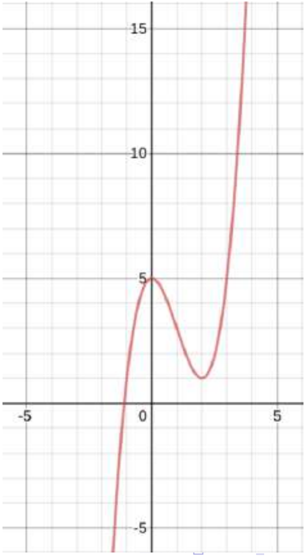
Maxima and minima

- $f(x) = x^2$ attains a minimum value at $x = 0$, no maximum value



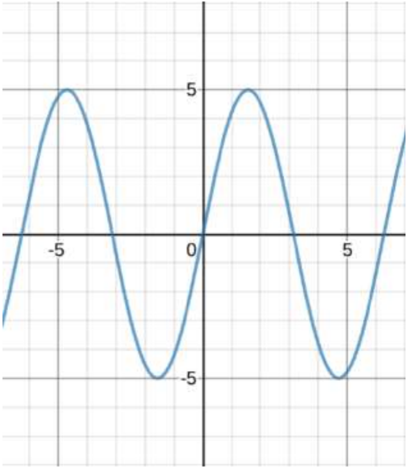
Maxima and minima

- $f(x) = x^2$ attains a minimum value at $x = 0$, no maximum value
- $f(x) = x^3 - 3x^2 + 5$ has no global minimum or maximum, but a local maximum at $x = 0$ and local minimum at $x = 2$



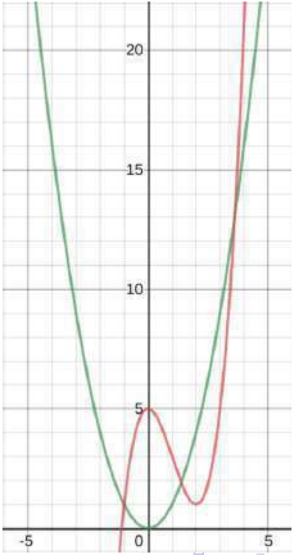
Maxima and minima

- $f(x) = x^2$ attains a minimum value at $x = 0$, no maximum value
- $f(x) = x^3 - 3x^2 + 5$ has no global minimum or maximum, but a local maximum at $x = 0$ and local minimum at $x = 2$
- $f(x) = 5\sin(x)$ periodically attains minimum value -5 and maximum value $+5$, infinitely often



Comparing functions

- Does one function grow faster than another?
- $f(x) = x^3 - 3x^2 + 5$ grows faster than $g(x) = x^2$
- Let $G(y)$ be the number of Data Science graduates in year y
- Let $J(y)$ be the number of new Data Science jobs in year y
- Ideally, $G(y)$ and $J(y)$ should grow at similar rates
- If $J(y)$ grows faster than $G(y)$, more students will opt to study Data Science



Summary

- We will typically study functions over numbers
- Many properties of functions are interesting
  - Range of outputs
  - Inputs for which function attains (local) maximum, minimum value
  - Relative growth rates of functions
  - ...

How many primes are there?

- A prime number p has exactly two factors, 1 and p
- The first few prime numbers are $2, 3, 5, 7, \ldots$
- Is the set of prime numbers finite?
- Equivalently, is there a largest prime?
- Euclid proved, around 300 BCE, that there cannot be a largest prime
- Hence there must be infinitely many primes

Euclid of Alexandria

A fact about divisibility

Observation

If $n|(a+b)$ and $n|a$, then $n|b$

- Since $n|(a+b)$, $a+b = u \cdot n$
- Since $n|a$, $a = v \cdot n$
- Therefore $a+b = vn+b = un$
- Hence $b = (u-v)n$

Euclid of Alexandria

There is no largest prime number

- Suppose the list of primes is finite, say $\{p_1, p_2, \ldots, p_k\}$
- Consider $n = p_1 \cdot p_2 \cdots p_k + 1$.
- If n is a composite number, at least one prime p_j is a factor, so $p_j | n$.
- Since p_j appears in the product $p_1 \cdot p_2 \cdots p_k$, we have $p_j | p_1 \cdot p_2 \cdots p_k$
- From our observation about divisibility, if $p_j | n$ and $p_j | p_1 \cdot p_2 \cdots p_k$, we must also have $p_j | 1$, which is not possible
- So n must also be a prime, which is clearly bigger than p_k

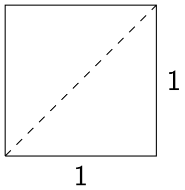
Euclid of Alexandria

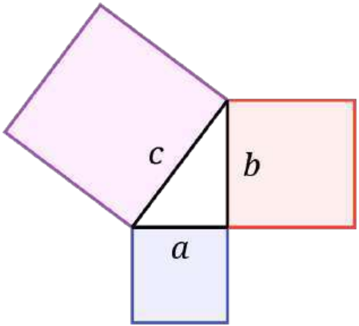
More about primes

- Prime numbers have been extensively studied in mathematics
- Let $\pi(x)$ denote the number of primes smaller than x
- The Prime Number Theorem says that $\pi(x)$ is approximately $\dfrac{x}{\log(x)}$ for large values of x
- Checking whether a number is a prime can be done efficiently — [Agrawal, Kayal, Saxena 2002]
- No known efficient way to find factors of non-prime numbers
- Large prime numbers are used in modern cryptography
- Essential for electronic commerce

Irrational numbers

- The discovery of irrational numbers is attributed to the ancient Greeks
- Since Pythagoras, it was known that the diagonal of a unit square has length $\sqrt{2}$
- His followers spent many years trying to prove it was rational
- Hippasus is attributed with proving that $\sqrt{2}$ is irrational, around 500 BCE
- The followers of Pythagoras were shocked by the discovery
- Allegedly, they drowned Hippasus at sea to suppress this fact from the public





The proof of Hippasus that $\sqrt{2}$ is not a rational number

- If $\sqrt{2}$ is rational, it can be written as a reduced fraction p/q, where $gcd(p, q) = 1$
- From $\sqrt{2} = p/q$, squaring both sides, $2 = p^2/q^2$
- Cross multiplying, $p^2 = 2q^2$, so $p^2 = p \cdot p$ is even
- The product of two odd numbers is odd and the product of two even numbers is even, so p is even, say $p = 2a$
- So $p^2 = (2a)^2 = 4a^2 = 2q^2$
- Therefore $q^2 = 2a^2$, so q^2 is also even
- By the same reasoning, q is even, say $q = 2b$.
- So $p = 2a$ and $q = 2b$, which means $gcd(p, q) \geq 2$, which contradicts our assumption that p/q was in reduced form.

Hippasus
Engraving by Girolamo Olgiati, 1580

Summary

- The proof of Hippasus follows a pattern commonly used in mathematical reasoning
- To show that a fact P holds, assume $not(P)$ and derive a contradiction
- Using a similar strategy, can show that for any natural number n that is not a perfect square, $\sqrt{n}$ is irrational

Hippasus
Engraving by Girolamo Olgiati, 1580

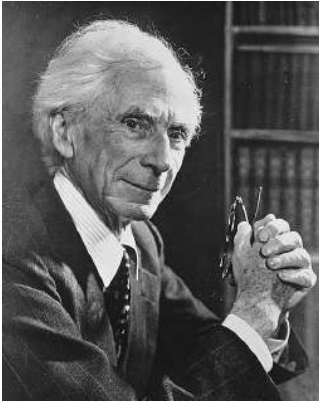
Set theory as a foundattion for mathematics

- A set is a collection of items
- Use set theory to build up all of mathematics
  - Georg Cantor, Richard Dedekind 1870s
- Natural numbers can be "defined" as follows
  - 0 corresponds to the empty set $\emptyset$
  - 1 is the set $\{0, \{0\}\} = \{\emptyset, \{\emptyset\}\}$
  - 2 is the set $\{1, \{1\}\}$
  - ...
  - $j+1$ is the set $\{j, \{j\}\}$
- Define arithmetic operations in terms of set building

Georg Cantor

Russell's Paradox

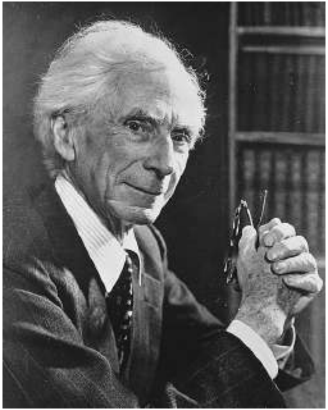
- Set theory assumes the emptyset $\emptyset$ and basic set building operations
  - Union $\cup$, Intersection $\cap$, Cartesian product $\times$, ...
  - Set comprehension — subset that satisfies a condition
- Is every collection a set? Is there a set of all sets?
- Consider S, all sets that do not contain themselves
  - S is a set, by set comprehension
  - Does S belong to S?
  - Yes? But elements of S do not contain themselves
  - No? Any set that does not contain itself should be in S
- Russell's Paradox — also discovered by Ernst Zermelo
- Cannot have "set of all sets"



Bertrand Russell
©*Dutch National Archives*

Sets and collections

- Russell's Paradox tells us that not every collection can be called a set
- Collection that is not a set is sometimes called a class
- The paradox had a major impact on set theory as a logical foundation of mathematics
- For us, just be sure that we always build new sets from existing sets
- Don't manufacture sets "out of thin air" — "set" of all sets



Bertrand Russell

Are there degrees of infinity?

- Cardinality of a set is the number of elements
- For finite sets, count the elements
- What about infinite sets?
  - Is $\mathbb{N}$ smaller than $\mathbb{Z}$?
  - Is $\mathbb{Z}$ smaller than $\mathbb{Q}$?
  - Is $\mathbb{Q}$ smaller than $\mathbb{R}$?
- First systematically studied by Georg Cantor
- To compare cardinalities of infinite sets, use bijections
  - One-to-one and onto function
  - Pairs elements from the sets so that none are left out

Georg Cantor

Countable sets

- Starting point of infinite sets is $\mathbb{N}$
- Suppose we have a bijection f between $\mathbb{N}$ and a set X
  - Enumerate X as $\{f(0), f(1), \ldots, \}$
  - X can be "counted" via f
  - Such a set is called countable

Georg Cantor

$\mathbb{Z}$ is countable

- $\mathbb{Z}$ extends $\mathbb{N}$ with negative integers
- Intuitively, $\mathbb{Z}$ is twice as large as $\mathbb{N}$
- Can we set up a bijection between $\mathbb{N}$ and $\mathbb{Z}$?

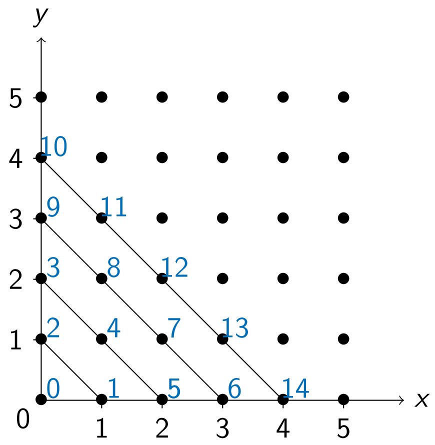
$$
\begin{array}{ccccccccccc}
\cdots, & -4, & -3, & -2, & -1, & 0, & 1, & 2, & 3, & 4, & \cdots \\
 & & & & 2, & 0, & 1, & & & & \\
 & & & 4, & 2, & 0, & 1, & 3, & & & \\
\cdots, & 8, & 6, & 4, & 2, & 0, & 1, & 3, & 5, & 7, & \cdots
\end{array}
$$

- The enumeration is effective
  - $f(0) = 0$
  - For i odd, $f(i) = (i+1)/2$
  - For i even, $f(i) = -(i/2)$
- $\mathbb{Z}$ is countable

Georg Cantor

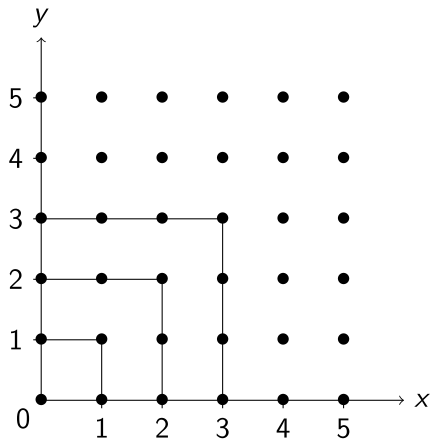
Is $\mathbb{Q}$ countable?

- $\mathbb{Q}$ is dense, $\mathbb{Z}$ is discrete
- Are there more rationals than integers?
- There is an obvious bijection between $\mathbb{Z} \times \mathbb{Z}$ and $\mathbb{Q}$
  - $(p, q) \mapsto \frac{p}{q}$
- Sufficient to check cardinality of $\mathbb{Z} \times \mathbb{Z}$
  - For simplicity, we restrict to $\mathbb{N} \times \mathbb{N}$
- Enumerate $\mathbb{N} \times \mathbb{N}$ diagonally



Is $\mathbb{Q}$ countable?

- $\mathbb{Q}$ is dense, $\mathbb{Z}$ is discrete
- Are there more rationals than integers?
- There is an obvious bijection between $\mathbb{Z} \times \mathbb{Z}$ and $\mathbb{Q}$
  - $(p, q) \mapsto \frac{p}{q}$
- Sufficient to check cardinality of $\mathbb{Z} \times \mathbb{Z}$
  - For simplicity, we restrict to $\mathbb{N} \times \mathbb{N}$
- Enumerate $\mathbb{N} \times \mathbb{N}$ diagonally
- Other enumeration strategies are also possible
- Can easily extend these to $\mathbb{Z} \times \mathbb{Z}$
- Hence $\mathbb{Q}$ is countable



Is $\mathbb{R}$ countable?

- $\mathbb{R}$ extends $\mathbb{Q}$ by irrational numbers
- Cantor showed that $\mathbb{R}$ is not countable
- First, a different set
  - Infinite sequences over $\{0, 1\}$
    $0\ 1\ 0\ 1\ 1\ 0\ \cdots$
- Suppose there is some enumeration

	b_0	b_1	b_2	b_3	b_4	$\cdots$
s_0	0	1	1	1	0	$\cdots$
s_1	1	0	1	0	0	$\cdots$
s_2	1	1	1	1	1	$\cdots$
s_3	0	1	1	0	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Is $\mathbb{R}$ countable?

- $\mathbb{R}$ extends $\mathbb{Q}$ by irrational numbers
- Cantor showed that $\mathbb{R}$ is not countable
- First, a different set
  - Infinite sequences over $\{0, 1\}$
    $0\ 1\ 0\ 1\ 1\ 0\ \cdots$
- Suppose there is some enumeration
- Flip b_i in s_i
- Read off the diagonal sequence
- Diagonal sequence differs from each s_i at b_i
- New sequence that it not part of the enumeration

	b_0	b_1	b_2	b_3	b_4	$\cdots$
s_0	1	1	1	1	0	$\cdots$
s_1	1	1	1	0	0	$\cdots$
s_2	1	1	0	1	1	$\cdots$
s_3	0	1	1	1	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Is $\mathbb{R}$ countable?

- Infinite sequences over $\{0,1\}$ cannot be enumerated
- Each sequence can be read as a decimal fraction

  0.011101110011
- Injective function from $\{0,1\}$ sequences to interval $[0,1) \subseteq \mathbb{R}$
- Hence $[0,1) \subseteq \mathbb{R}$ cannot be enumerated
- So $\mathbb{R}$ is not countable

	b_0	b_1	b_2	b_3	b_4	$\cdots$
s_0	1	1	1	1	0	$\cdots$
s_1	1	1	1	0	0	$\cdots$
s_2	1	1	0	1	1	$\cdots$
s_3	0	1	1	1	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Summary

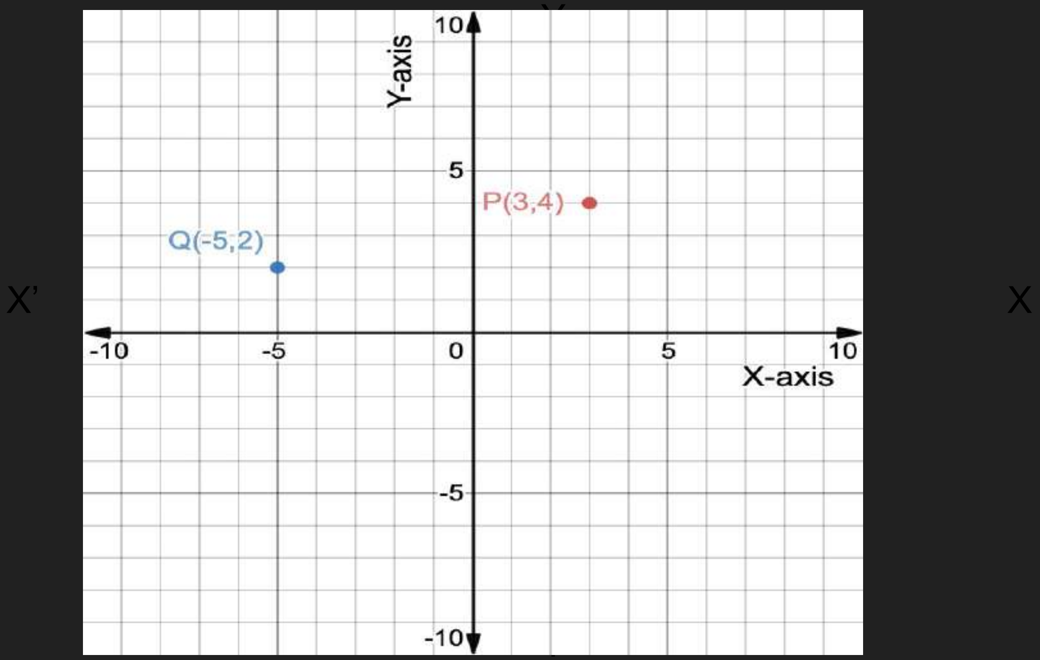
- Any set that has a bijection from $\mathbb{N}$ is countable
- $\mathbb{Z}$ and $\mathbb{Q}$ are countable
- $\mathbb{R}$ is not countable — diagonalization
- Is there a set whose size is between $\mathbb{N}$ and $\mathbb{R}$?
- Continuum Hypothesis — one of the major questions in set theory
- Paul Cohen showed that you can neither prove nor disprove this hypothesis within set theory

Georg Cantor

Elements of Coordinate Geometry

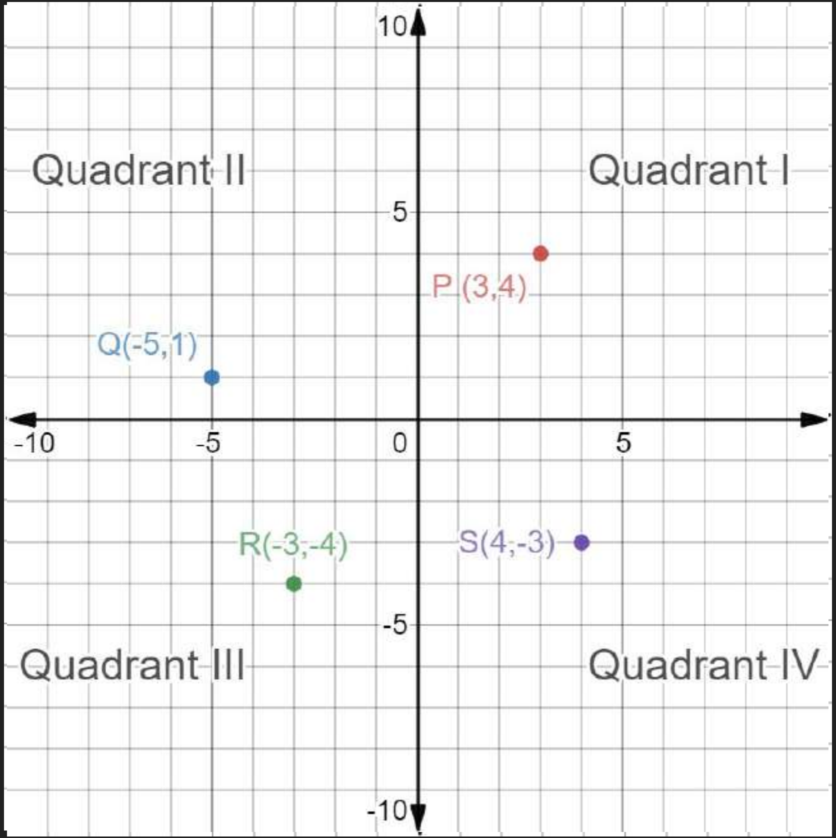
Axes, Points and Lines

Rectangular Coordinate System



Coordinate Plane

- The horizontal line is called X-axis.
- The Vertical line is called Y-axis.
- The point of intersection of these two lines is called origin.
- Any point on the coordinate plane can be represented by an ordered pair (x,y).
- For example, P=(3,4), Q=(-5,2).

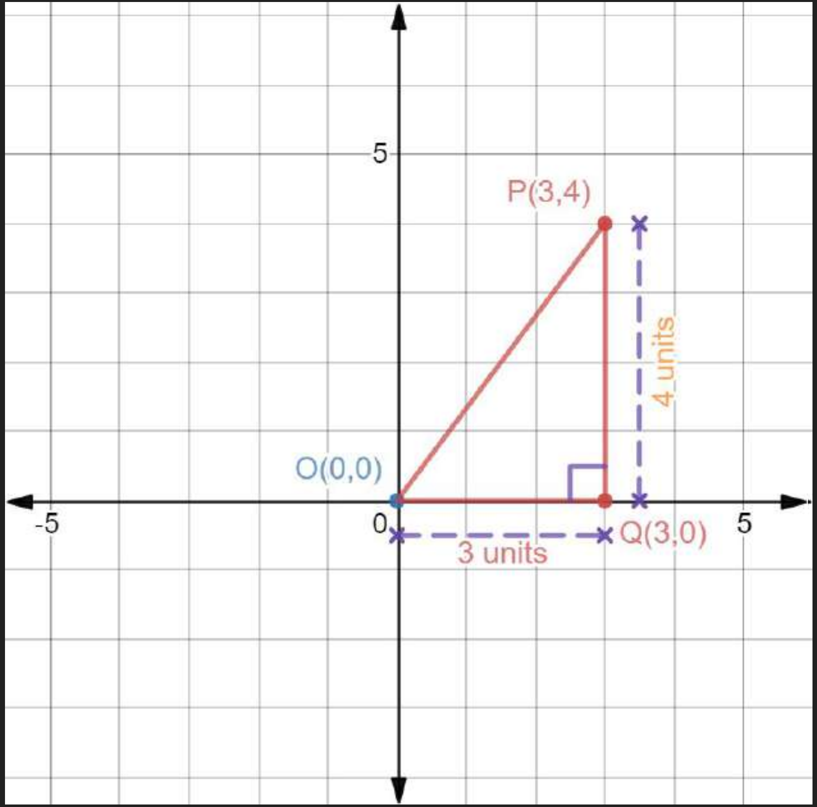


Quadrants in the coordinate system

The coordinate axes split the coordinate plane into **four quadrants and two axes.**

- Quadrant I: (+,+)
- Quadrant II: (-,+)
- Quadrant III: (-,-)
- Quadrant IV:(+,-)
- X-axis: (±, 0)
- Y-axis: (0, ±)
- Origin (0,0)

Distance of a Point from Origin



Goal: To find the distance of Point P (3,4) from the origin.

1. Drop a perpendicular on X-axis which intersects the X-axis at Q (3,0).
2. By Pythagorean Theorem,

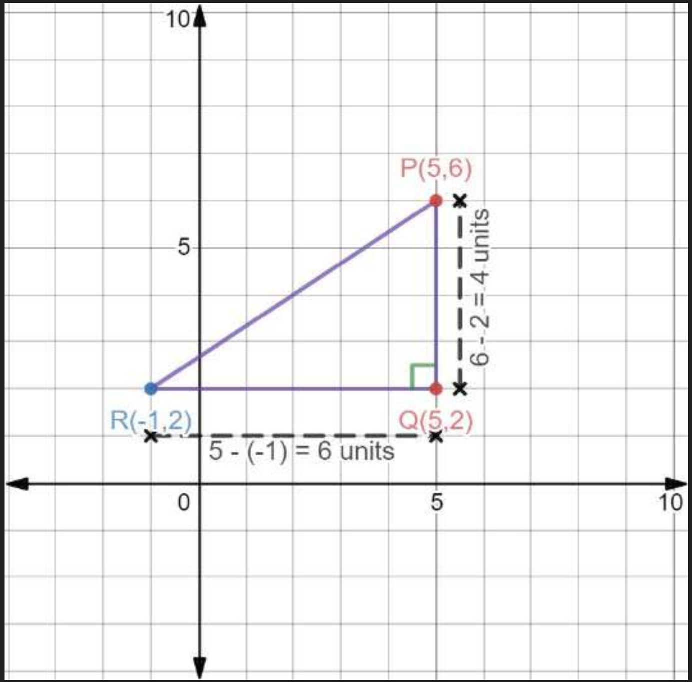
$$OP^2 = OQ^2 + QP^2$$

Hence,

$$OP = \sqrt{OQ^2 + QP^2} = \sqrt{3^2 + 4^2} = 5.$$

Distance Between Any Two Points

Goal: To find the distance between any two Points P (x_1, y_1) and R (x_2, y_2).



- Construct a right-angled triangle with right angle at Point Q (x_1, y_2).
- By Pythagorean Theorem,

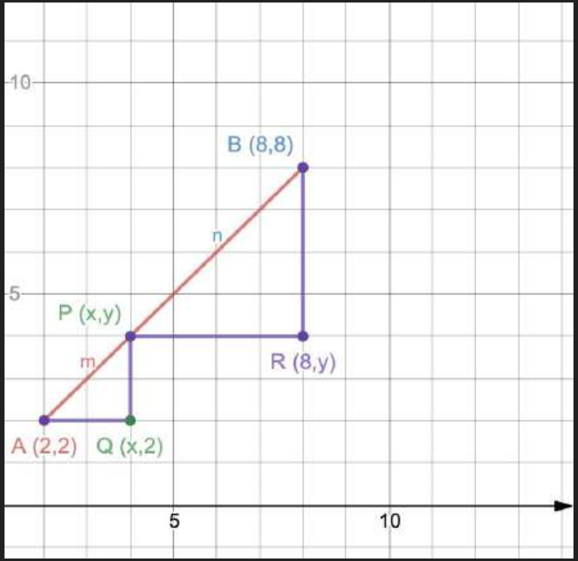
$$PR^2 = QR^2 + PQ^2.$$

$$PR = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$= \sqrt{6^2 + 4^2} = \sqrt{52} = 2\sqrt{13}.$$

Section Formula

Given that, the Point P cuts the line segment AB in the *m:n* ratio. Our goal is to find the coordinates of P.



Let the Coordinates of A and B are (x_1,y_1) and (x_2,y_2), respectively. Assume that P has the coordinates (x,y).

Observe that $\triangle AQP \sim \triangle PRB$. Hence,

$$\frac{m}{n} = \frac{AP}{PB} = \frac{AQ}{PR} = \frac{PQ}{BR}$$

$$\frac{m}{n} = \frac{x-x_1}{x_2-x} = \frac{y-y_1}{y_2-y}$$

$$x = \frac{mx_2+nx_1}{m+n}, y = \frac{my_2+ny_1}{m+n}$$

Area of a Triangle using coordinates

Goal: To find area of $\triangle ABC$ with known coordinates.

Let the coordinates of the vertices be A (x_1, y_1), B(x_2,y_2) and C(x_3, y_3).

A($\triangle$ABC) = A(□ADFC) - A(□ADEB) - A(□BEFC)

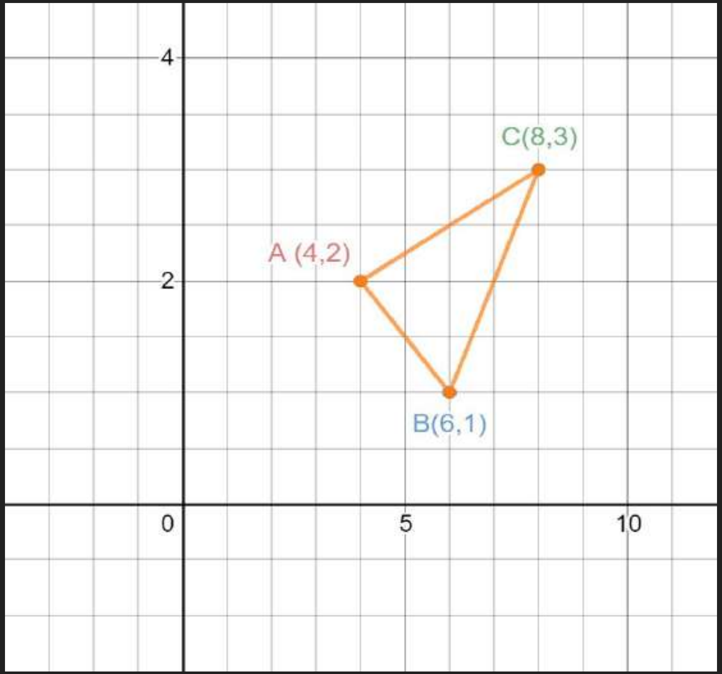
A(□ADFC) = ½ (AD+CF)xDF = ½ $(y_1 + y_3)(x_3 - x_1)$

A(□ADEB) = ½ (AD+EB)xDE = ½ $(y_1 + y_2)\ (x_2 - x_1)$

A(□BEFC) = ½ (BE+CF)xEF = ½$(y_2 + y_3)(x_3 - x_2)$

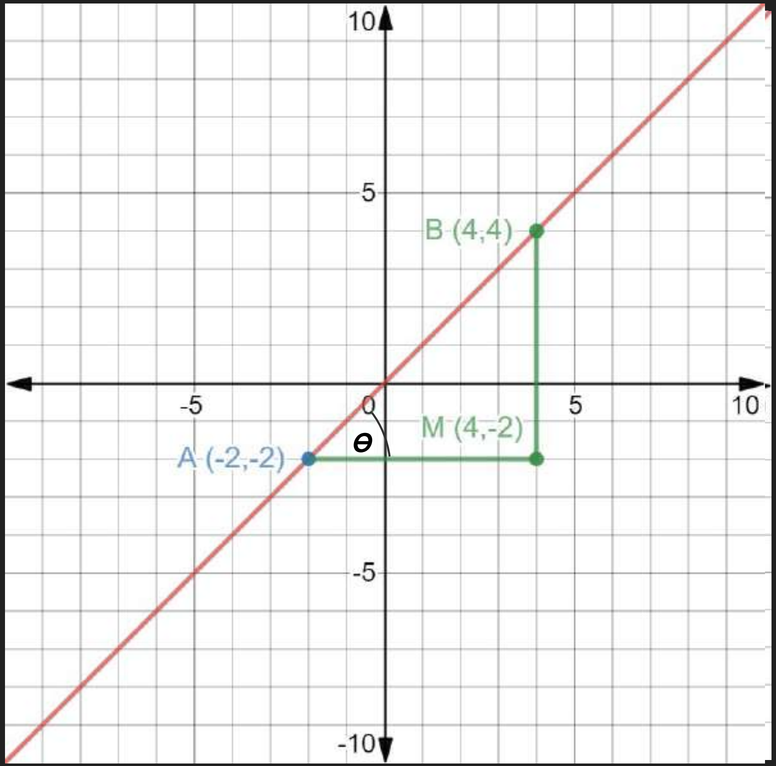
A($\triangle$ABC) = ½(2+3)x4 - ½(2+1)x2 - ½(1+3)x2

= 10 - 3 - 4 = 3 square units.



A($\triangle$ABC) = ½ $|x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)|$.

Slope of a Line



Goal: To find the slope of a line, given on a coordinate plane.

- Identify two points on the line, say, A (x_1, y_1) and B (x_2, y_2).
- Construct a right angled triangle with a right angle at the Point M (x_2, y_1).

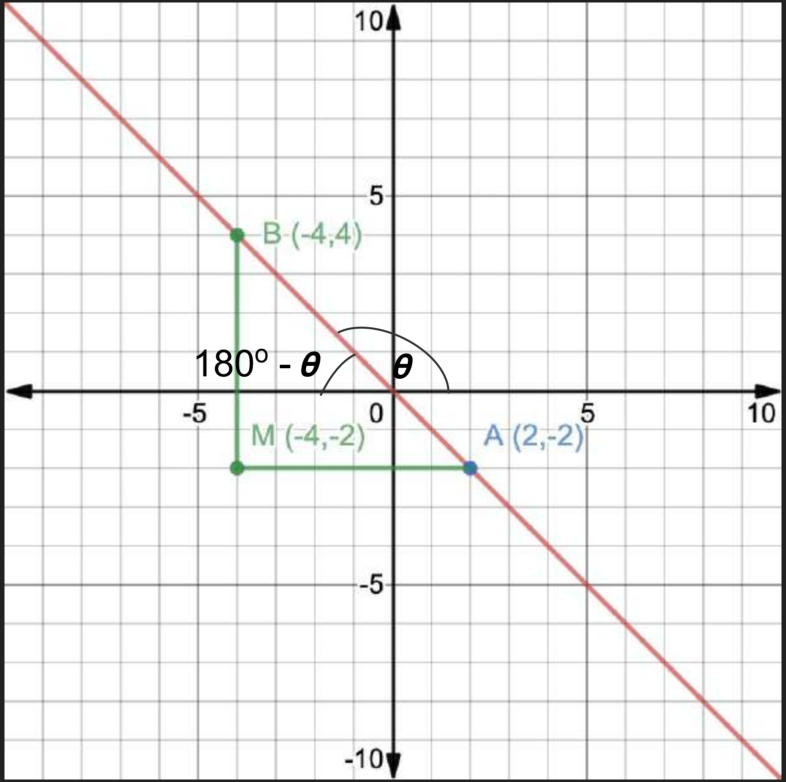
- Define

$$m = \frac{MB}{AM} = \frac{y_1 - y_2}{x_1 - x_2} = \tan\theta.$$

- The *m* is called slope of a line.
- $\boldsymbol{\theta}$ is called the inclination of the line with positive X-axis, measured in anticlockwise direction.

- $0^\circ \leq \boldsymbol{\theta} \leq 180^\circ$

Slope of a line (Continued)



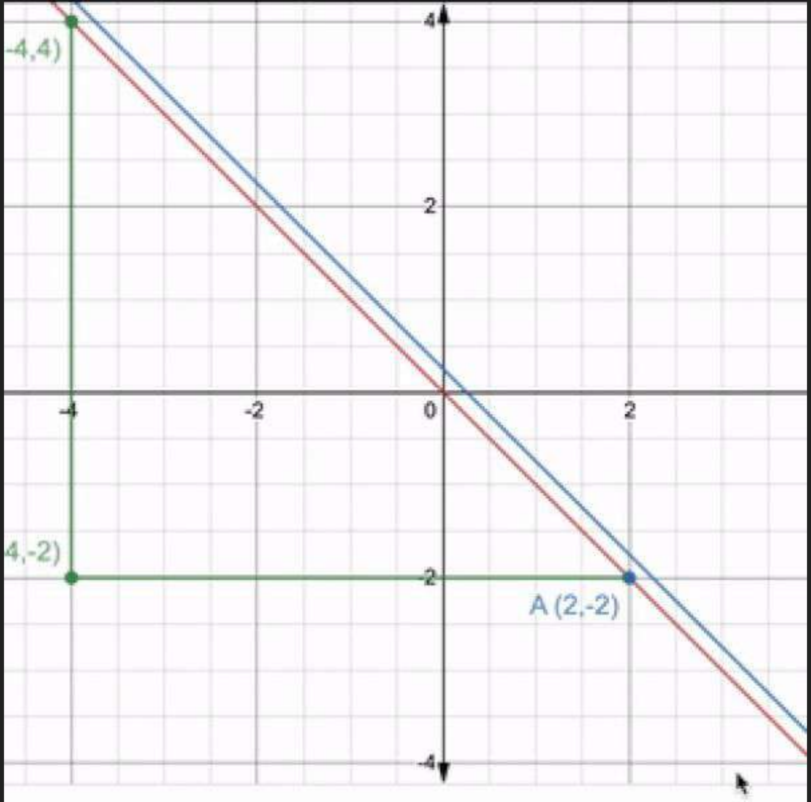
- Observe that the lines parallel to X-axis have inclination of 0º. Hence the slope *m* = tan 0 = 0.
- The inclination of a vertical line is 90º. Hence, the slope *m* is undefined.

Definition: If ***θ*** is the inclination of a line ***l***, then tan***θ*** is called the slope or gradient of line ***l***.

If ***θ*** ≠ 90º, then *m* = tan***θ***.

$$m = \tan(180 - \theta) = -\tan\theta = \frac{y_1 - y_2}{x_1 - x_2}.$$

Can slope of a line uniquely determine a line?



Answer: No, it can not uniquely determine the line.

How is the slope useful?

To explore:

- Condition for parallel lines
- Condition for perpendicular lines

Characterization of Parallel Lines via slope

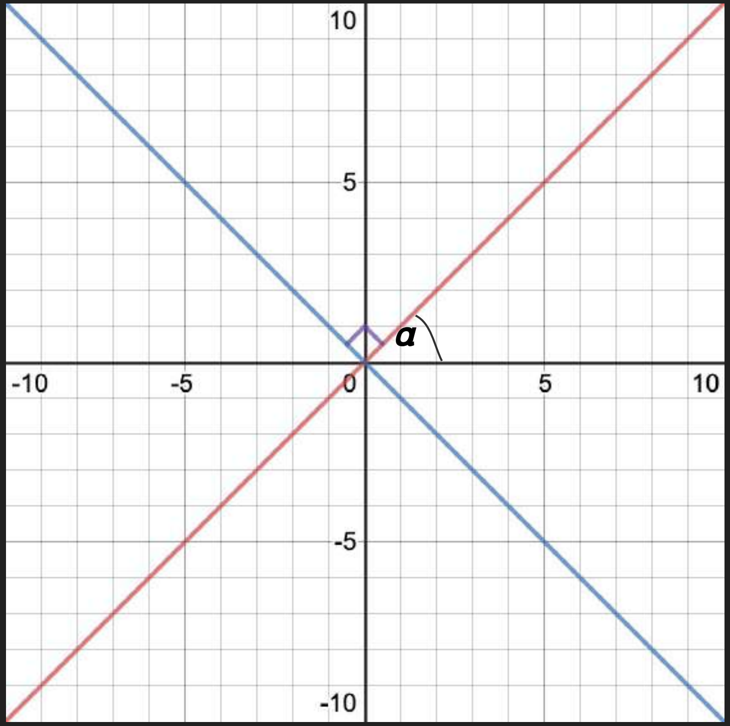
Let $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ be two non-vertical lines with slopes m_1 and m_2 with inclinations $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively.

- If $\boldsymbol{l}_1$ is parallel to $\boldsymbol{l}_2$, then $\boldsymbol{\alpha} = \boldsymbol{\beta}$.
- It is clear that $\tan\boldsymbol{\alpha} = \tan\boldsymbol{\beta}$.
- Hence, $m_1 = m_{2.}$

- Assume $m_1 = m_2$. Then $\tan\boldsymbol{\alpha} = \tan\boldsymbol{\beta}$.
- Since, $0^o \leq \boldsymbol{\alpha}, \boldsymbol{\beta} \leq 180^{o,}$, $\boldsymbol{\alpha} = \boldsymbol{\beta}$.
- Therefore, $\boldsymbol{l}_1$ is parallel to $\boldsymbol{l}_2$

Two non-vertical lines $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ are parallel if and only if their slopes are equal.

Characterization of Perpendicular Lines via Slope



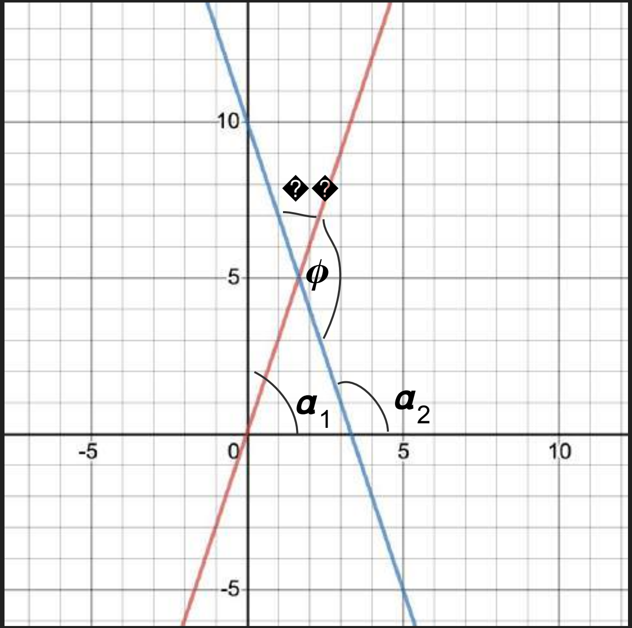
Let l_1 and l_2 be two non-vertical lines with slopes m_1 and m_2 with inclinations α and β respectively.

- If l_1 is perpendicular to l_2, then $90 + \alpha = \beta$.
- Now, $\tan\beta = \tan(90 + \alpha) = -\cot\alpha = -1/\tan\alpha$.
- Hence, $m_2 = -1/m_1$ or $m_1 m_2 = -1$.

- Assume $m_1 m_2 = -1$. Then $\tan\alpha \tan\beta. = -1$.
- $\tan\alpha = -\cot\beta = \tan(90+\beta)$ or $\tan(90-\beta.)$.
- Hence, α and β differ by 90^o which proves l_1 is perpendicular to l_2

Two non-vertical lines l_1 and l_2 are perpendicular if and only if $m_1 m_2 = -1$

Relation of Angles between the Two lines and their slopes



Let l_1 and l_2 be two non-vertical lines with slopes m_1 and m_2 with inclinations α_1 and α_2 respectively.

Suppose l_1 and l_2 intersect and let φ and θ be the adjacent angles formed by l_1 and l_2.

Now, $\theta = \alpha_2 - \alpha_1$, for $\alpha_1, \alpha_2 \neq 90^o$

Then,

$$\tan\theta = \tan(\alpha_2 - \alpha_1) = \frac{\tan\alpha_2 - \tan\alpha_1}{1+\tan\alpha_1\tan\alpha_2} = \frac{m_2-m_1}{1+m_1m_2}, m_1m_2 \neq -1.$$

$$\tan\phi = \tan(180 - \theta) = -\tan\theta = \frac{m_1-m_2}{1+m_1m_2}$$

Representation of a Line

- How to represent a line uniquely?
- Given a point, how to decide whether the point lies on a line?

In other words, for a given line ***l***, we should have a definite expression that describes the line in terms of coordinate plane.

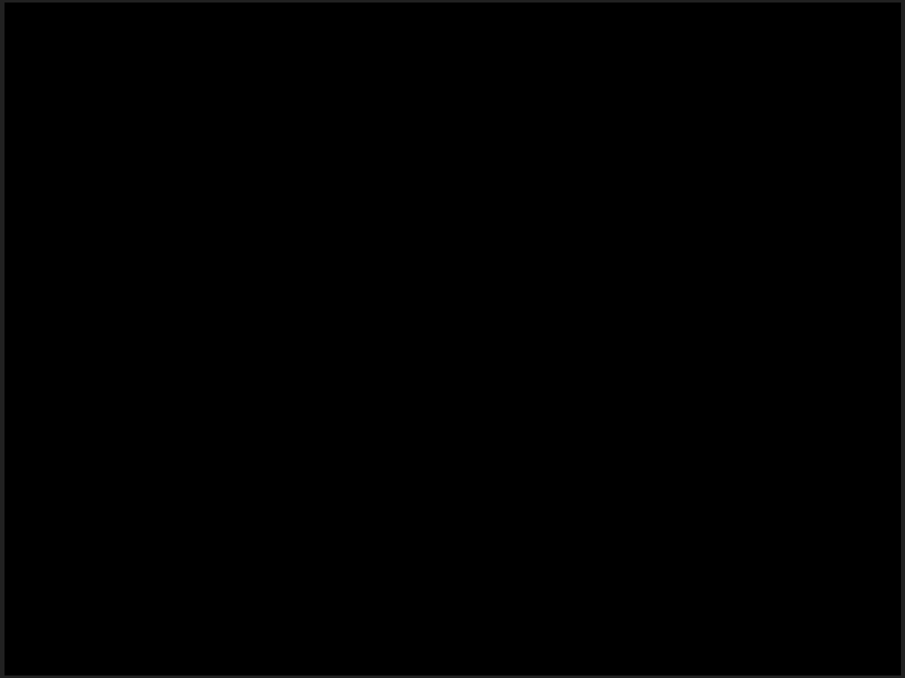
If the coordinates a given point P, satisfy the expression for the line ***l***, then the point P lies on the line ***l***.

Horizontal and Vertical Lines

Horizontal Lines: A line is a horizontal line only if it is paralllel to X-axis

- To locate such a line, we need to specify the value it takes on Y-axis.
- That is, the expression for such a line is of the form $y = a$.
- Then all points that lie on this line are of the form (x, a).

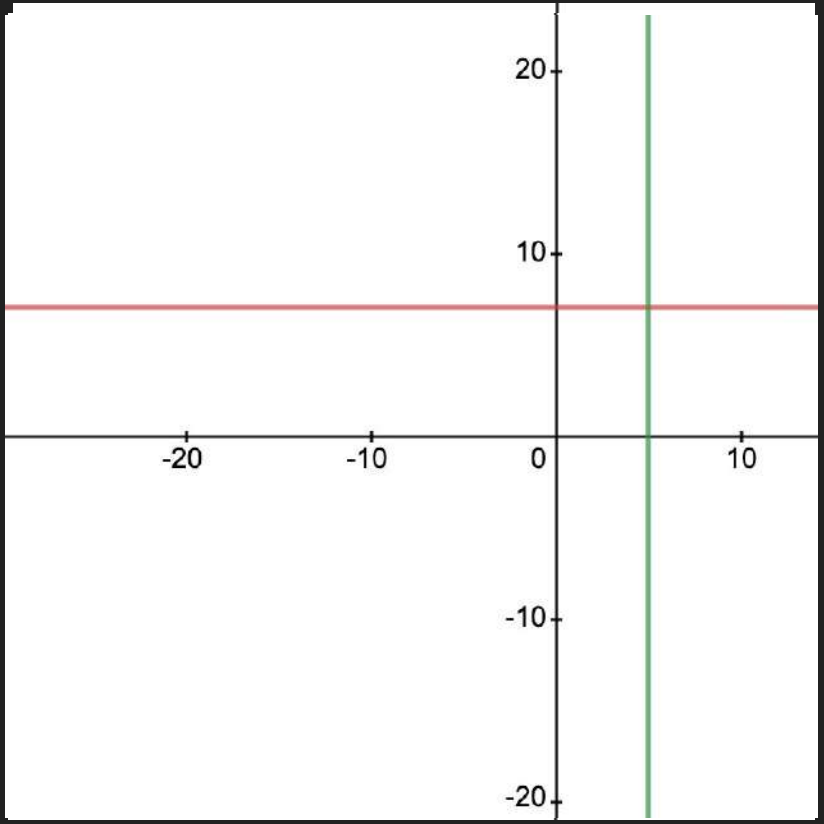
Horizontal and Vertical Lines



Vertical Lines: A line is a vertical line only if it is parallel to Y-axis

- To locate such a line, we need to specify the value it takes on X-axis.
- That is, the expression for such a line is of the form $x = b$.
- Then, all points that lie on this line are of the form (b, y).

Example



Question: Find the equation of the lines parallel to the axes and passing through (5,7).

The horizontal line is $y = 7$.

The vertical line is $x = 5$.

Equation of a Line: Point-Slope Form

For a non-vertical line ***l***, with slope m and a fixed point P (x_0, y_0) on the line, can we find the equation (algebraic representation) of the line?

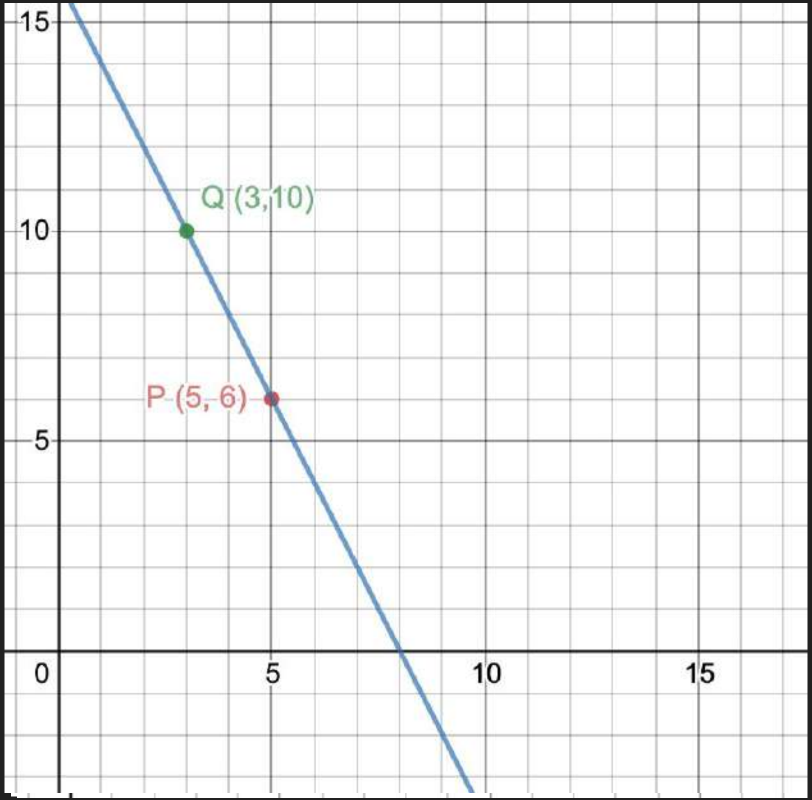
- Let Q(x, y) be an arbitrary point on line ***l***. Then, the slope of the line is given by

$$m = \frac{y - y_0}{x - x_0}$$

$$(y - y_0) = m(x - x_0)$$ (Point-Slope form)

Any point P (x,y) is on line ***l***, if and only if the coordinates of P satisfy the above equation.

Example



Q. Find the equation of a line through the point P(5,6) with slope -2.

Let Q(x,y) be an arbitrary point on this line. Then, using Point-Slope form, we get

$$-2 = \frac{y-6}{x-5}$$

$$(y-6) = 2(5-x) \text{ or } y = 16 - 2x.$$

Equation of a Line: Two-Point Form

Let the line ***l*** pass through the points P(x_1, y_1) and Q(x_2, y_2).

Assume that R(x,y) is an arbitrary point on the line ***l***.

Then, the points P, Q, and R are collinear.

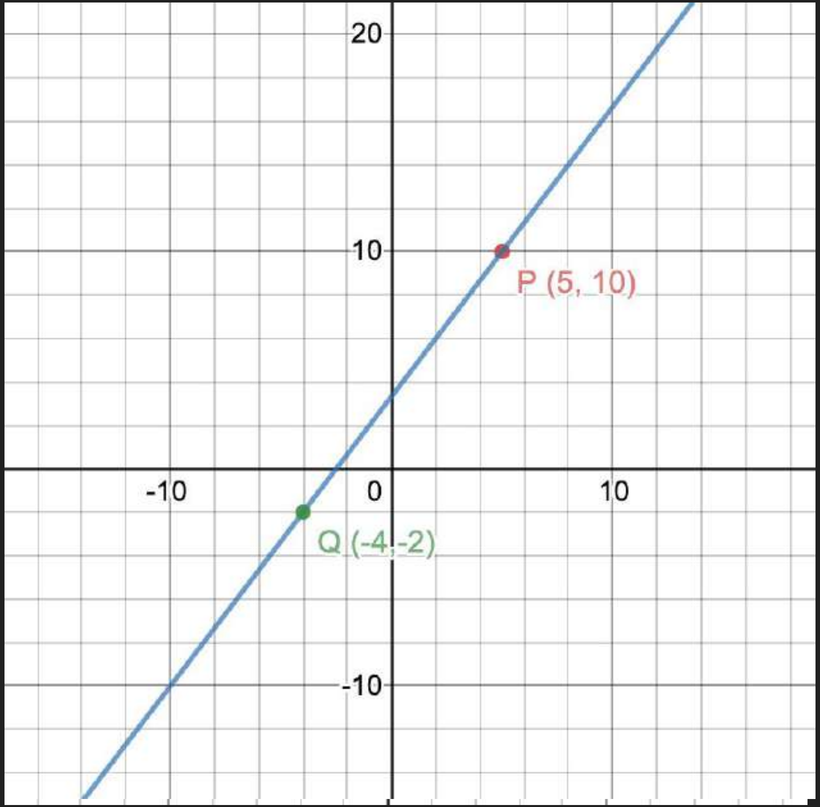
Hence, Slope of PR = Slope of PQ. Therefore, $\frac{y-y_1}{x-x_1} = \frac{y_2-y_1}{x_2-x_1}$

$$(y - y_1) = \frac{y_2-y_1}{x_2-x_1}(x - x_1).$$

(Two-Point form)

Any point R (x,y) is on line ***l***, if and only if, the coordinates of R satisfy the above equation.

Example



Q. Find the equation of a line passing through (5,10) and (-4, -2).

Let (x,y) be an arbitrary point on this line. Then by two-point form, we get

$$(y-10)=\frac{-2-10}{-4-5}(x-5)$$

$$3y=4x+10.$$

Equation of Line: Slope-Intercept Form

Let a line l with slope m cut Y-axis at c. Then c is called the y-intercept of the line l.

That is, the point $(0,c)$ lies on the line l.

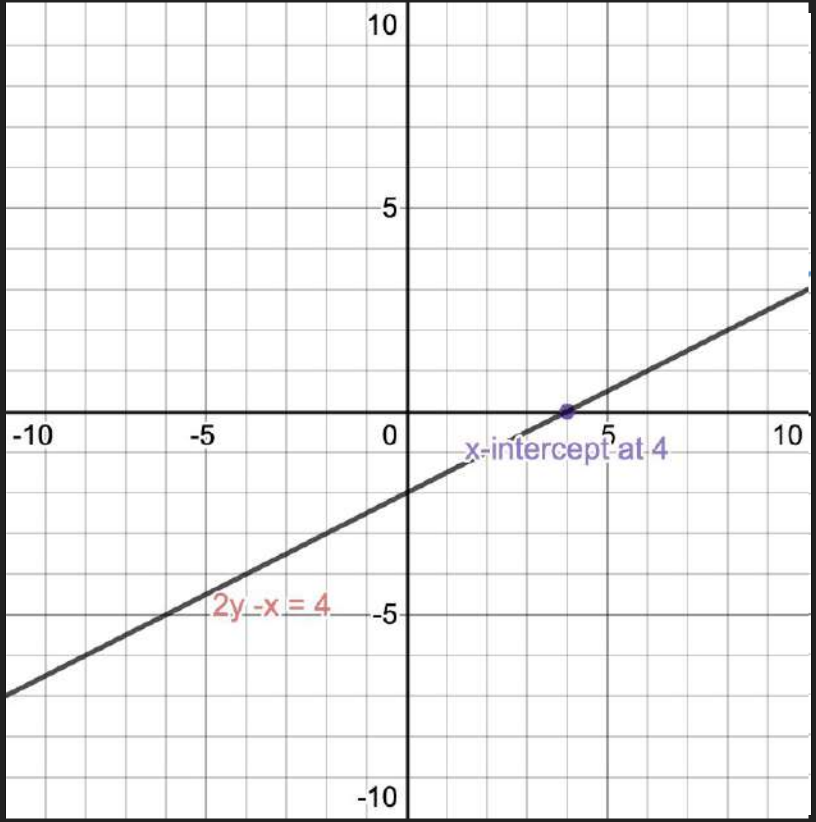
Therefore, by Point-Slope form, we get $y - c = mx$, or $y = mx + c$.

Let a line l with slope m cut X-axis at d. Then d is called the x-intercept of the line l.

That is, the point $(d,0)$ lies on the line l.

Therefore, by Point-Slope form, we get $y = m(x - d)$.

Examples



Q. Find the equation of a line with slope ½ and y-intercept -3/2.

The equation of the line is $y = \frac{1}{2}x - 3/2$

Q. Find the equation of a line with slope ½ and x-intercept 4.

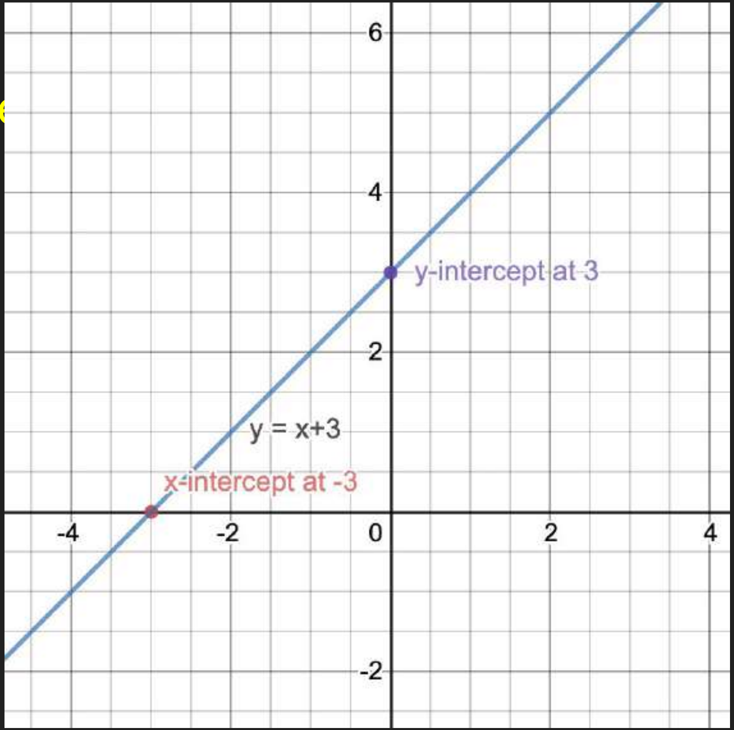
The equation of the line is $y = \frac{1}{2}(x - 4)$ or $2y - x + 4 = 0$.

Equation of a Line: Intercept Form

Suppose a line makes x-intercept at *a* and y-intercept at *b*. Then the two points on the line are (*a*,*0*) and (*0*,*b*).

Using two-point form,

$$(y-0)=\frac{b-0}{0-a}(x-a) \text{ or } \frac{x}{a}+\frac{y}{b}=1$$



Example

Q. Find the equation of a line having x-intercept at -3 and y-intercept at 3.

$$\frac{x}{-3}+\frac{y}{3}=1 \text{ or } y=x+3.$$

General Equation of a Line

Different forms of Equation of Line	Representation	General Form Ax + By +C = 0
Slope-Point Form	$(y - y_0) = m(x - x_0)$	$m = -\frac{A}{B}, y_0 - mx_0 = -\frac{C}{B}$
Slope-Intercept Form	$y = mx + c \text{ or } y = m(x - d)$	$m = -\frac{A}{B}, c = -\frac{C}{B} \text{ or } d = -\frac{C}{A}$
Two-Point Form	$(y - y_1) = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1).$	$\frac{y_2 - y_1}{x_2 - x_1} = -\frac{A}{B}, y_1 + \frac{A}{B}x_1 = -\frac{C}{B}$
Intercept Form	$\frac{x}{a} + \frac{y}{b} = 1$	$a = -\frac{C}{A}, b = -\frac{C}{B}$

Any equation of the form Ax+By+C = 0, where A, B ≠ 0 simultaneously, is called *general linear equation* or *general equation of a line*.

Example

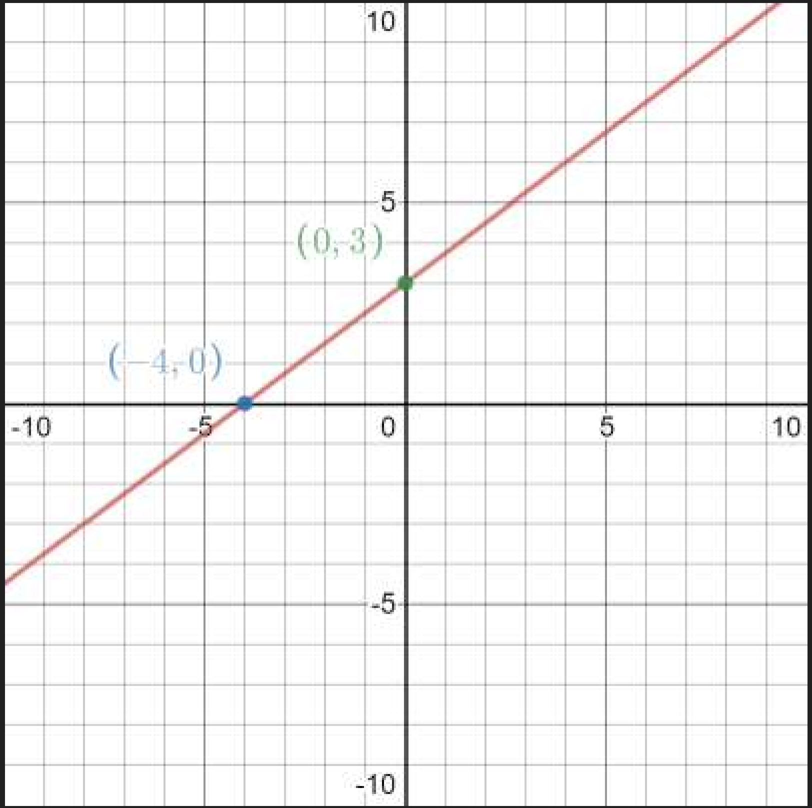
Question. The equation of a line is 3x -4y +12 = 0. Find the slope, x-intercept and y-intercept of the line.

Identify A = 3, B = -4 and C = 12.

Using Intercept form, $a = -C/A = -4$ and $b = -C/B = 3$.

Using Slope-intercept form, $y = \frac{3}{4}x + 3$. Hence, $m = \frac{3}{4}$.

Slope = ¾, x-intercept= -4 and y-intercept = 3.



Examples

Question. Show that the two lines $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$, $b_1, b_2 \neq 0$ are

- parallel if $a_1b_2 = a_2b_1$, and
- perpendicular if $a_1a_2 + b_1b_2 = 0$.

Using Slope-intercept form,

$$m_1 = -\frac{a_1}{b_1} \text{ and } m_2 = -\frac{a_2}{b_2}$$

If the lines are parallel, then $a_1b_2 = a_2b_1$.

If the lines are perpendicular, then $a_1a_2 + b_1b_2 = 0$.

Two non-vertical lines l_1 and l_2 are parallel if and only if their slopes are equal.

Two non-vertical lines l_1 and l_2 are perpendicular if and only if $m_1m_2 = -1$

Examples

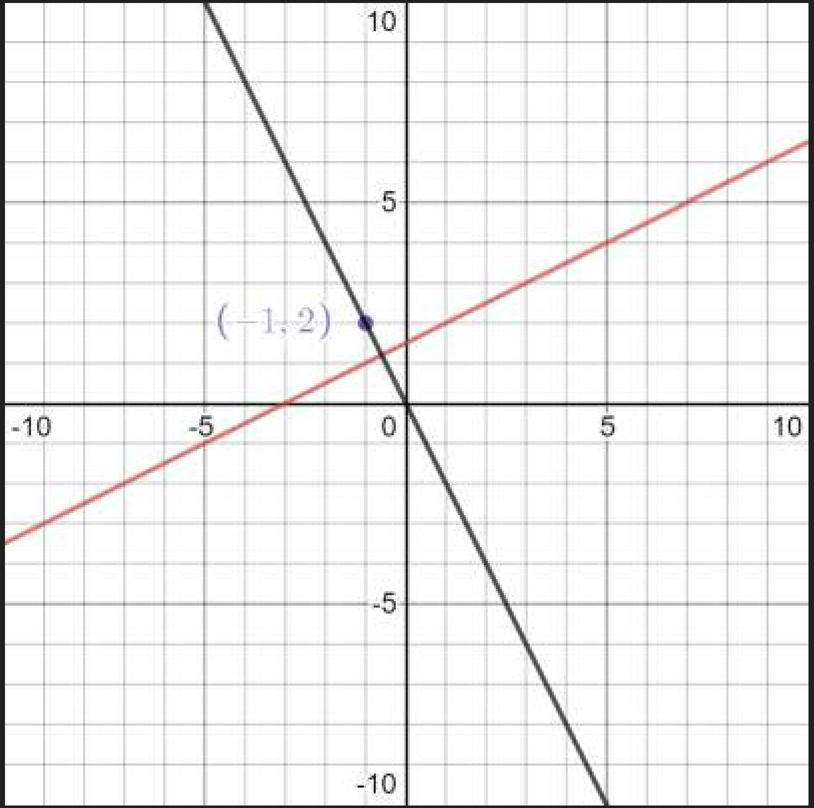
Find the equation of a line perpendicular to the line $x - 2y + 3 = 0$ and passing through the point (-1,2).

The slope of the given line is $m_1 = ½$.

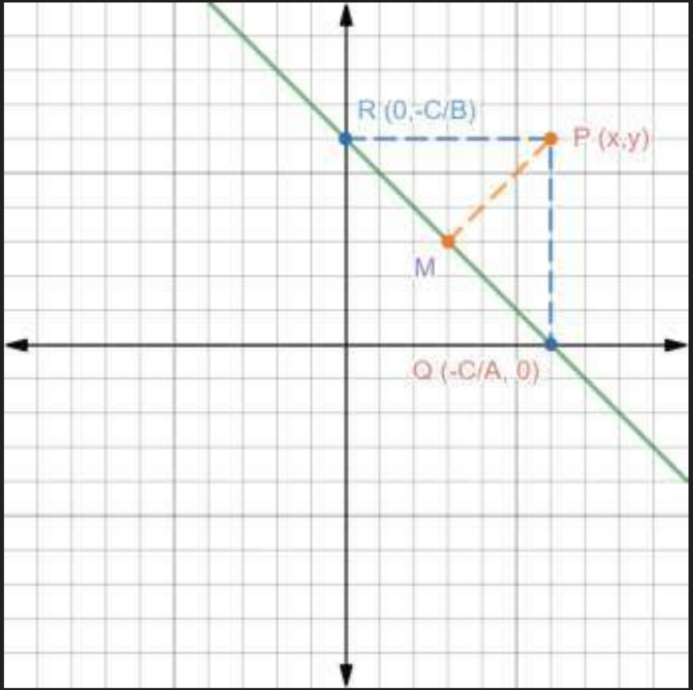
The slope of a line perpendicular to the given line is $m_2 = -1/m_1 = -2$.

To find the equation of the line passing through the point (-1,2) and slope -2.

$(y - 2) = -2 (x +1)$ or $y = -2x$.



Distance of a Point from a Line



Goal. To find the distance of the point P(x_1,y_1) from the line l having equation Ax + By + C = 0.

For A, B ≠ 0, Using Intercept form,

x-intercept = - C/A and y-intercept = -C/B

A(△PQR) = ½ QRxPM. Hence, PM = 2 A(△PQR)/QR

$$A(\Delta PQR) = \frac{1}{2}\left|x_1\left(\frac{-C}{B}\right) - \frac{C}{A}\left(y_1 + \frac{C}{B}\right)\right| = \frac{1}{2}\frac{|C|}{|AB|}|Ax_1 + By_1 + C|$$

$$QR = \sqrt{\frac{C^2}{A^2} + \frac{C^2}{B^2}} = \frac{|C|}{|AB|}\sqrt{A^2 + B^2}.$$

$$PM = \frac{2A(\Delta PQR)}{QR} = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}.$$

A(△PQR) = ½ $|x_1(y_2 - y_3) + x_2(y_3 - y_2) + x_3(y_1 - y_2)|$.

Distance between two Parallel Lines

Let l_1 and l_2 be two parallel lines with slopes *m*.

l_1: $y = mx + c_1$. Comparing with general form, we get x-intercept at $(-c_1/m)$.

l_2: $y = mx + c_2$ Comparing with general form, we get $A = -m$, $B = 1$ and $C = -c_2$.

By using Distance of a point from a line formula, where point is $(-c_1/m, 0)$, we get

$$\frac{|A(-c_1/m)+B(0)+C|}{\sqrt{A^2+B^2}} = \frac{|c_1-c_2|}{\sqrt{1+m^2}}.$$

For general form, $m = -A/B$, $c_1 = -C_1/B$ and $c_2 = -C_2/B$, then

$$d = \frac{|C_1-C_2|}{\sqrt{A^2+B^2}}$$

Examples

Question. Find the distance of the point (3, -5) from the line 3x - 4y -26 =0.

Ax+By+C = 0 implies A = 3, B = -4 and C = -26.
Also (x_1, y_1) = (3, -5). Then

$$d = \frac{|3(3)-4(-5)-26|}{\sqrt{3^2+(-4)^2}} = \frac{3}{5}.$$

Question. Find the distance between parallel lines 3x -4y + 7 = 0 and 3x-4y+5=0.

$$d = \frac{|C_1 - C_2|}{\sqrt{A^2+B^2}}$$

Observe that A = 3, B = -4 and C_1 = 7, C_2= 5. Then

$$d = \frac{|7-5|}{\sqrt{3^2+(-4)^2}} = \frac{2}{5}.$$

Example:Real-World

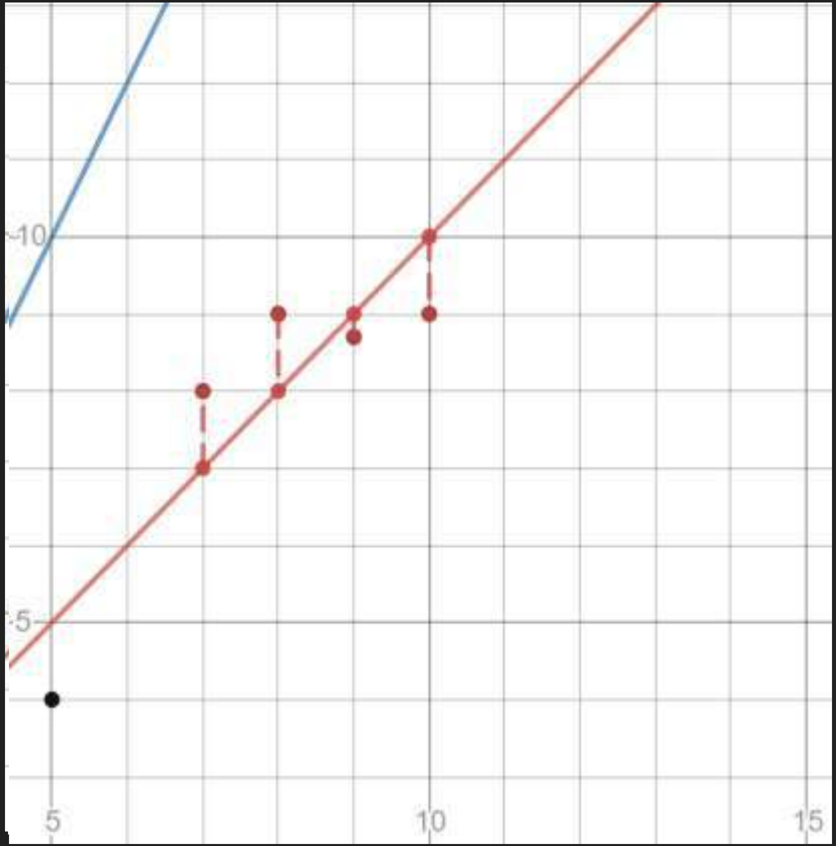
It is known that V=IR.

That is, Voltage= current x Resistance

In our context, this represents a line passing through the origin.

That is, y = m x, where y is the voltage, x is the current and m is the resistance.

You have been asked to perform an experiment to verify this phenomenon.



Example: Real-World (Contd)

How to say mathematically which line is better?

Let the equation of two lines be y = x and y = 2x.

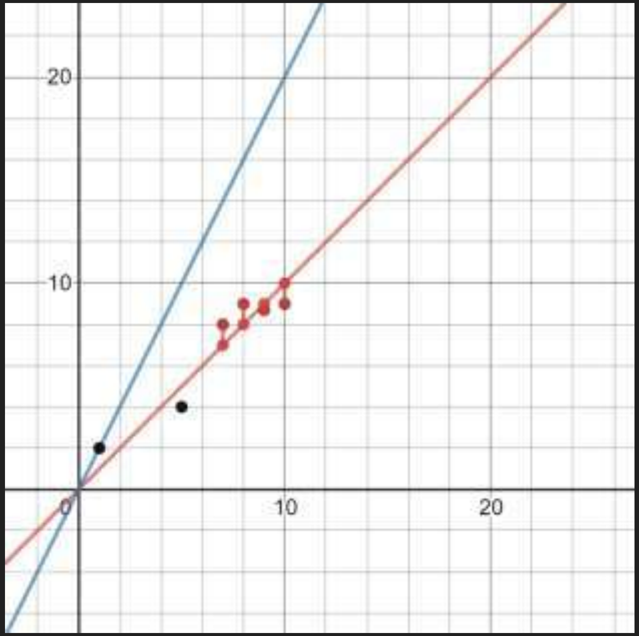
From the set of observations, (x_i, y_i), i=1,2,3,4,5,6.

We can consider the square of the differences

$$\sum_{i=1}^{6}(y_i - x_i)^2 \ \ and \sum_{i=1}^{6}(y_i - 2x_i)^2$$

The first difference is 5.09 and the second difference is 328.49.

Therefore, the first line is better than the second line.



x_i	y_i
1	2
5	4
7	8
8	9
9	8.7
10	9

Distance of a Set of Points from a Line

Apart from perpendicular distance, we can also talk about the distance which is parallel to Y-axis.

Consider the set of points $\{(x_i,y_i)|\ i=1,2,...,n\}$ and a line with equation $y = mx + c$.

Then the **squared sum** of the distance of set of points from the line is defined as

$$SSE = \sum_{i=1}^{n} (y_i - mx_i - c)^2.$$

Least Squares Motivation

- In general, this raises the following question
- Given a set of points, how to find the line that fits the given set of points?
- In other words, what is the equation of the best fit line for given set of points?

In other words, if I need to find the equation of line *y= m x + c*, then the question can be reframed into two questions.

- What is the value of m and c that best fits the given set of points.
- What is a meaning of best fit?

Best Fit: Given a set of *n* points, $\{(x_i,y_i)| \ i =1,2,...,n\}$, define

$$SSE = \sum_{i=1}^{n} (y_i - mx_i - c)^2.$$

Find the value of *m* and *c* that minimizes *SSE*.

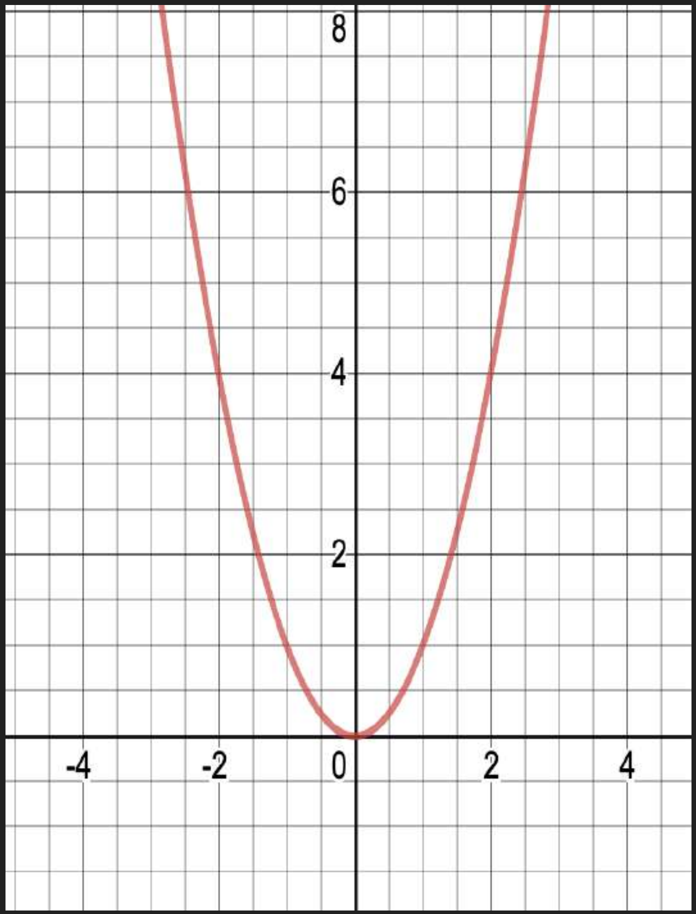
Quadratic Functions

Graphing

Quadratic Function (Definition)

- A quadratic function is described by an equation of the form
  - $f(x) = ax^2 + bx + c$, where $a \neq 0$.

Quadratic term Linear term Constant term



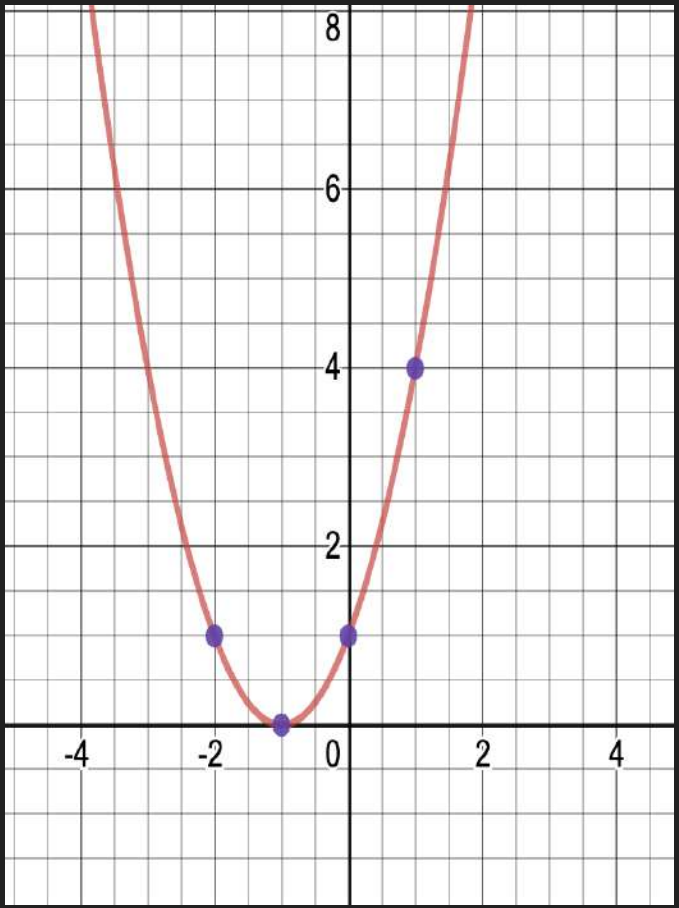
The graph of any quadratic function is called parabola.

To graph a quadratic function, plot the ordered pairs on the coordinate plane that satisfy the function.

Example: Graph a function $f(x) = x^2 + 2x + 1$

1. Generate a table of ordered pairs satisfying the function.
2. Plot the points on the coordinate plane.
3. Connect a smooth curve joining the points.

x	y
-2	1
-1	0
0	1
1	4



Important Observations

- All parabolas have an axis of symmetry. That is, if the graph paper containing the graph of parabola is folded along the axis of symmetry the portion of parabola on either sides will exactly match each other.
- The point at which the axis of symmetry intersects the parabola is called the vertex.
- The y-intercept of a quadratic function is c

Let $f(x) = ax^2 + bx+c$, where $a \neq 0$.

- The y-intercept: $y = a(0)^2+b(0) +c = c$
- The equation of axis of symmetry: $x = -b/(2a)$. (to be derived later).
- The x-coordinate of the vertex: $-b/(2a)$.

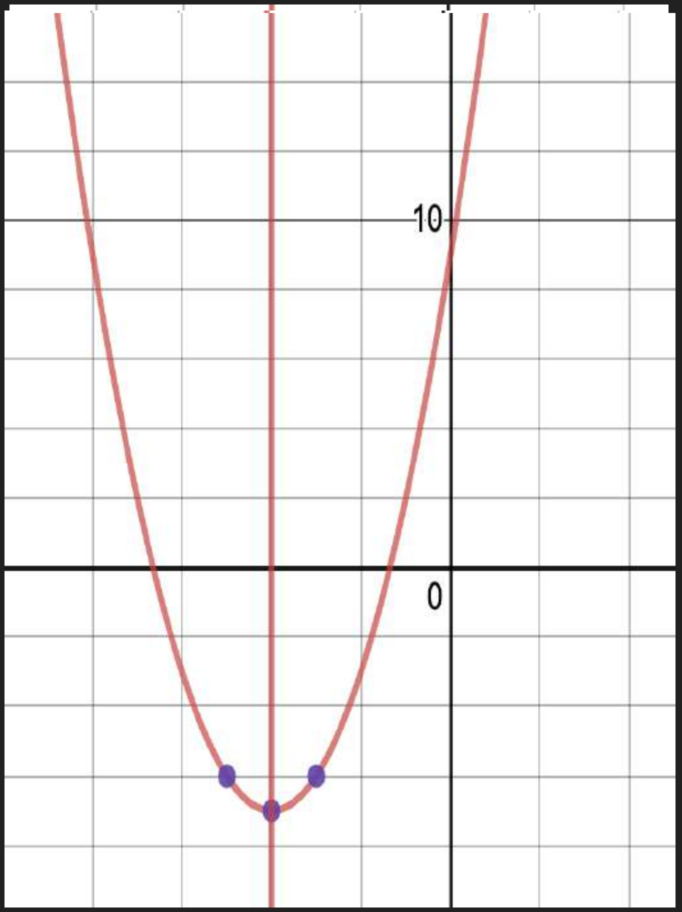
Example: Graph a function $f(x) = x^2+8x+9$

The y-intercept: *9*

The axis of symmetry: $x = -8/(2(1)) = -4$

The vertex: *(-4, -7)*

x	y
-3	-6
-4	-7
-5	-6



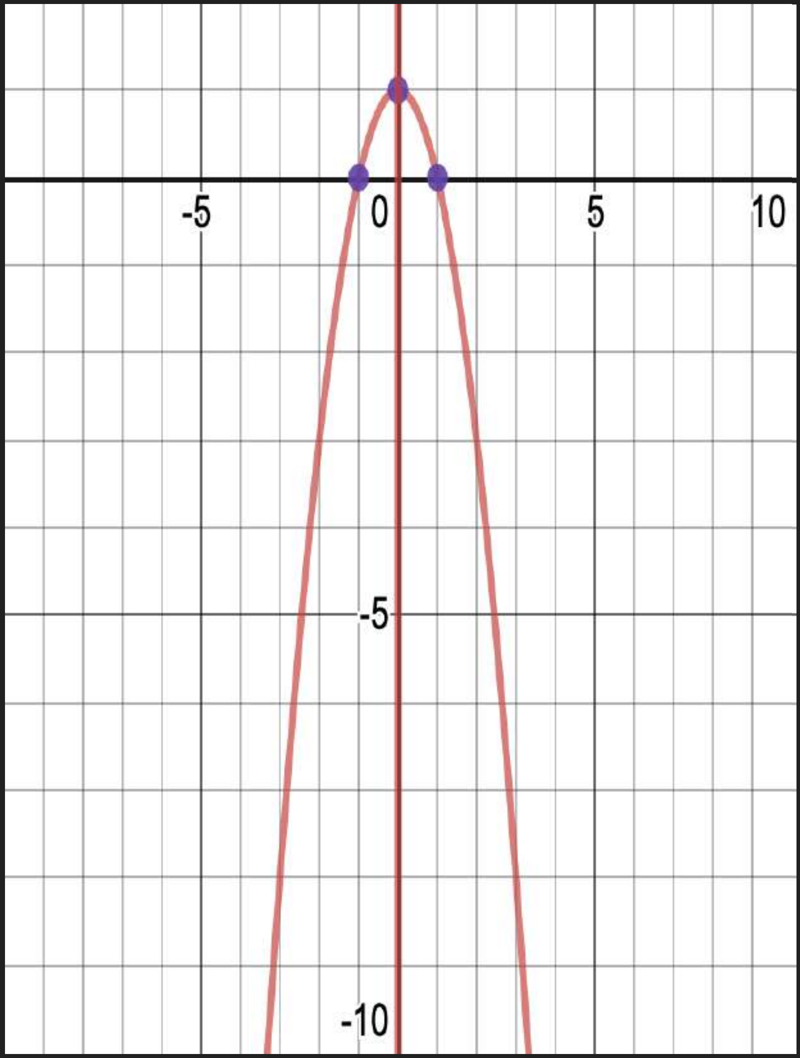
Example: Graph a function $f(x) = -x^2+1$.

The y-intercept: *1*

The axis of symmetry: $x = 0$

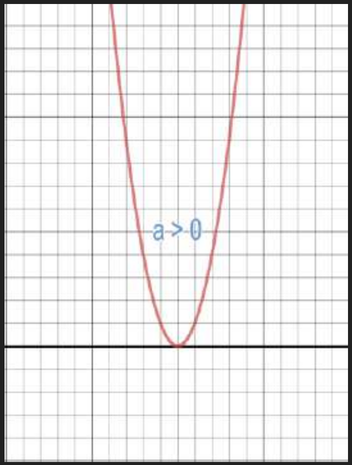
The vertex: *(0, 1)*

x	y
-1	0
0	1
1	0



Maximum and Minimum Values

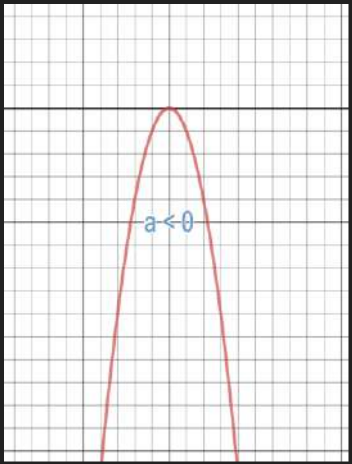
The y-coordinate of the vertex of a given quadratic function is the ***minimum or maximum*** value attained by the function.



The graph of a quadratic function $f(x)= ax^2+bx+c$, where $a \neq 0$ is:

- Opens up and has minimum value, if $a > 0$.
- Opens down and has maximum value if $a < 0$.
- The range of a quadratic function is

$$\mathbf{R}\cap\{f(x)|f(x) \geq f_{min}\} \text{ or } \mathbf{R}\cap\{f(x)|f(x) \leq f_{max}\}.$$



Example

Let $f(x) = x^2-6x+9$.

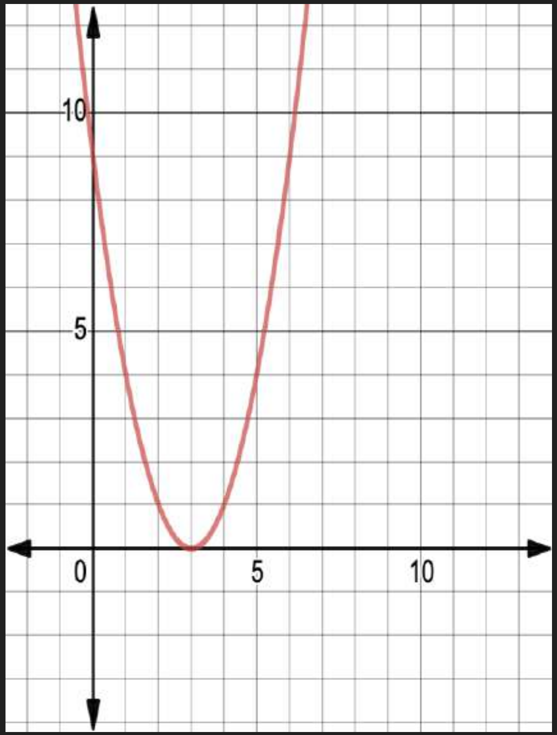
1. Determine whether f has minimum or maximum value. If so, what is the value?
2. State the domain and the range of f.

Observe that $a=1$, $b=-6$, and $c=9$.

Since, $a>0$, the function opens up and has the minimum value.

The minimum value is given by y-coordinate of the vertex. The x-coordinate of the vertex is $-b/(2a) = 3$. Therefore, the minimum value is $f(3) = 0$.

Domain = $\mathbf{R}$ and Range = $\mathbf{R} \cap \{f(x) | f(x) \geq 0\}$.



Example

Let $f(x) = x^2-6x+9$.

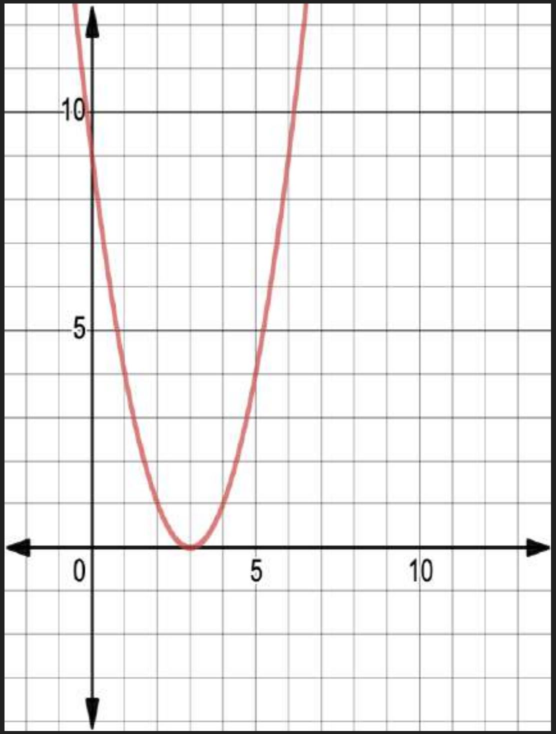
1. Determine whether f has minimum or maximum value. If so, what is the value?
2. State the domain and the range of f.

Observe that $a=1$, $b=-6$, and $c=9$.

Since, $a>0$, the function opens up and has the minimum value.

The minimum value is given by y-coordinate of the vertex. The x-coordinate of the vertex is $-b/(2a) = 3$. Therefore, the minimum value is $f(3) = 0$.

Domain = $\mathbf{R}$ and Range = $\mathbf{R}\cap\{f(x)|f(x) \geq 0\}$.



Example

A tour bus in Chennai serves 500 customers per day. The charge is ₹40/- per person. The owner of the bus service estimate that the company would lose 10 passengers per day for each ₹4/- fare hike.

How much should the fare be in order to maximize the income of the company?

Let x denote the number of ₹4/- fare hike. Then the price per passenger is $40+4x$, and the number of passengers is $(500-10x)$. Therefore, the income is

$$I(x) = (500-10x)(40+4x) = -40x^2+1600x+20000.$$

In this case, $a = -40$, $b = 1600$ and $c=20000$, and the maximum value attained will be $I(-b/(2a)) = I(20) =36000$.

This means the company should make 20 fare hikes of ₹4/- in order to maximize its income. That is the new fare = 40 + 4 x 20= ₹120/-

Slope of a quadratic function

Given a quadratic function, $f(x) = ax^2 + bx + c$, where $a \neq 0$, how to determine the slope of f?

Recall, for a linear function $y = g(x) = mx + c$, we have calculated the ratio of change in y and change in x and observed that it remains constant and is m. We also showed that $m = \tan\theta$, where θ is the inclination with positive X-axis.

Let us use similar analogy for a quadratic function and define slope of a quadratic function.

We now discuss the concept using a simple example.

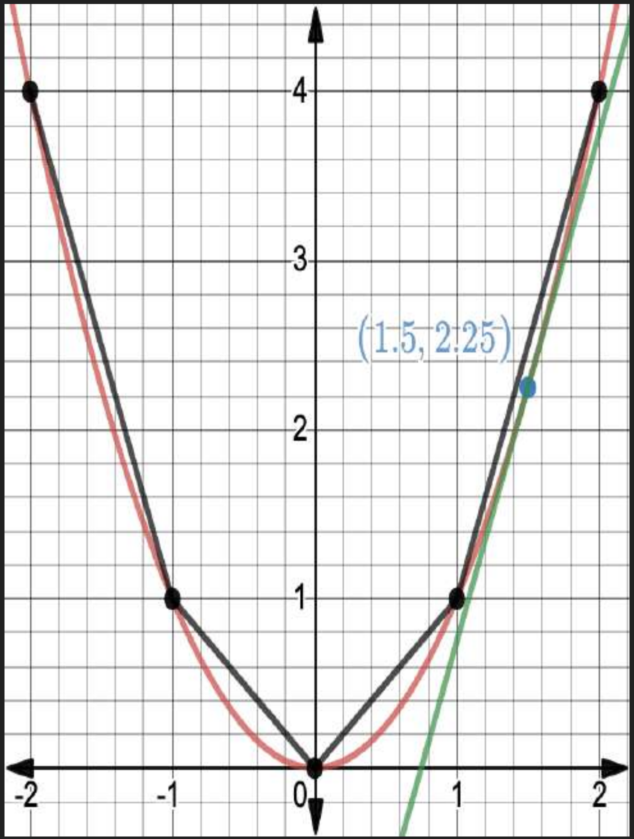
Slope of a quadratic function

Given a quadratic function, $f(x) = ax^2 + bx + c$, where $a \neq 0$, how to determine the slope of f?

Let $y = x^2$ be a quadratic function given.

Let us tabulate the ordered pairs

x_i	y_i	$y_i - y_{i-1}$
-2	4	
-1	1	-3
0	0	-1
1	1	1
2	4	3



The slope of $f(x) = x^2$ is $2x$.

Slope of a quadratic function

Given a quadratic function, $f(x) = ax^2 + bx + c$, where $a \neq 0$, how to determine the slope of f?

x_i	y_i	$y_i - y_{i-1}$	
-2	$4a - 2b + c$		
-1	$a - b + c$	$-3a + b$	
0	c	$-a + b$	$2a$
1	$a + b + c$	$a + b$	$2a$
2	$4a + 2b + c$	$3a + b$	$2a$

From the table, it is clear that the slope of $f = 2ax+b$.

Also note that, the slope denotes the rate of change of y with respect to x.
Hence, slope =0 means the function has either maximum or minimum which happens when $2ax+b = 0$. That is, $x = -b/(2a)$.

Quadratic Equations

Solve by Graphing

Quadratic Equation (Definition)

If a quadratic function is set equal to a value, then the result is a quadratic equation.

Eg. $ax^2+bx+c=0$, and $ax^2+bx+c=5$, where $a \neq 0$ are quadratic equations.

If $ax^2+bx+c=0$, with $a \neq 0$, and a,b,c are integers, then the quadratic equation is said to be in *the standard form*.

Roots of Equations and Zeros of Functions

The solutions to a quadratic equation are called *roots of the equation*.

One method for finding the roots of a quadratic equation is to find zeros of a related quadratic function.

Observe that the zeros of a function are x-intercepts of its graph and these are the solutions of related equation as $f(x)=0$ at these points.

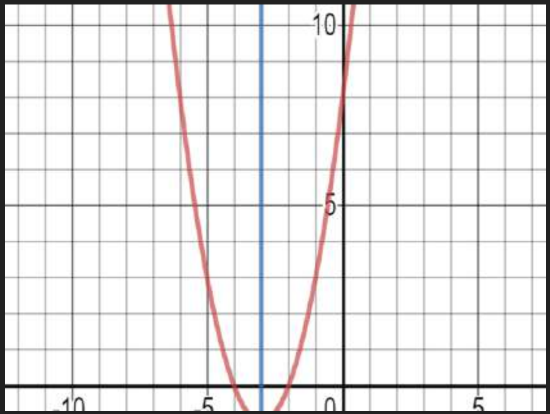
Examples

Find the roots of the following equations.

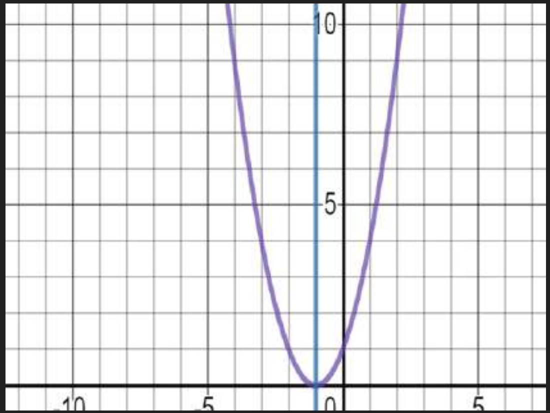
1. $x^2+6x+8=0.$
2. $x^2+2x+1=0.$
3. $x^2+1=0.$

Graph the related quadratic functions using axis of symmetry and vertex.

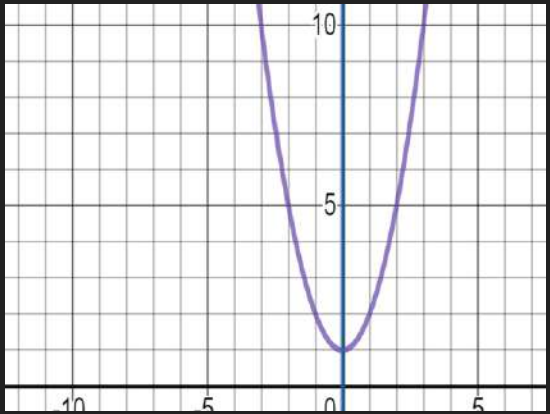
Axis of symmetry: $x = -3$
The roots are -4,-2,
Two real roots.



Axis of symmetry: $x = -1$
The roots are -1, -1
One real root.



Axis of symmetry: $x = 0$
No real roots.



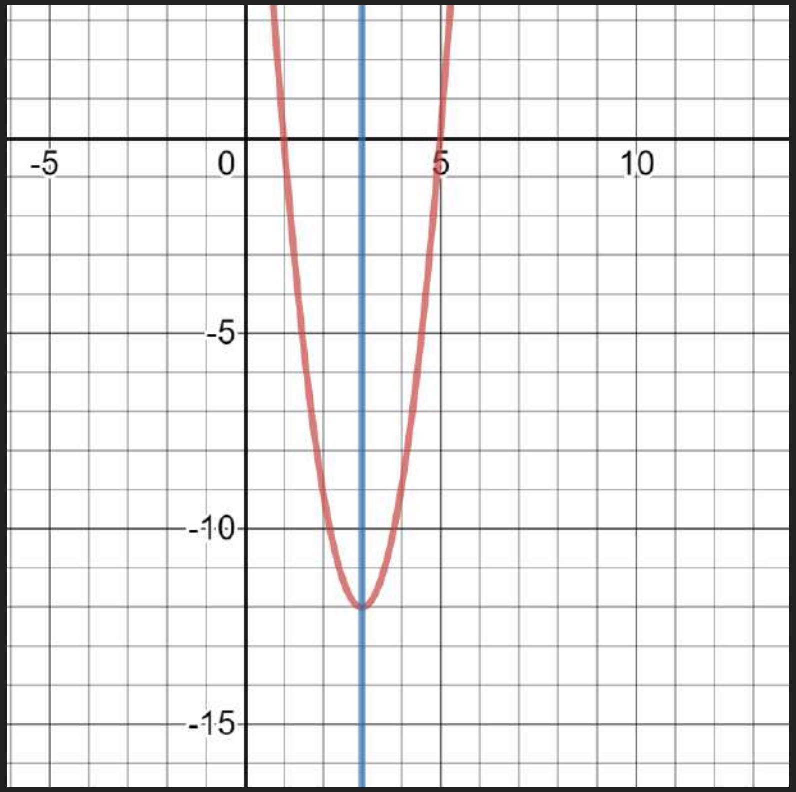
Quadratic Equations

Solve by Factoring

Quadratic Function: Intercept form

Let $y = f(x) = a(x-p)(x-q)$, where p and q represent x-intercepts for the function. Then the form $y = a(x-p)(x-q)$ is called the *intercept form*.

Example: Graph $y=3(x-1)(x-5)$



Question: How will you convert the intercept form into the standard form?

Intercept form to the Standard form

Changing intercept form to standard form requires us to use FOIL method which can be described as follows:

The product of two binomials is the sum of the products of the first(F) terms, the outer(O), the inner(I) and the last(L) terms.

$$(ax+b)(cx+d) = \underbrace{ax.\,cx}_{F} + \underbrace{ax.\,d} + \underbrace{b.\,cx}_{O} + \underbrace{b.\,d}$$

I L

Quick Observations:

The product of coefficient of x^2 and the coefficient of the constant is *abcd*.
The product of the two terms in the coefficient of x is also *abcd*.

Example

Question. Write a quadratic equation with roots, ⅔ and -4, in the standard form.

Recall: By standard form, we mean $ax^2+bx+c=0$, where a,b,c are integers.

By intercept form, we know $(x-\frac{2}{3})(x+4)=0$.

By FOIL method, $(x-\frac{2}{3})(x+4)= x^2+(-\frac{2}{3} +4)x -\frac{2}{3}.4 = x^2+(10/3)x - 8/3=0$

For standard form, multiply both sides by 3, to get

$$3x^2+10x-8=0.$$

Standard form to Intercept form

Example: Convert the function $f(x) = 5x^2 - 13x + 6$ to intercept form.

Let us apply FOIL Method.

$5x^2-13x+6 = (ax+b)(cx+d) = ac\,x^2+(ad+bc)x+bd.$

Therefore, $ac = 5$, $ad+bc=-13$ and $bd = 6$. That is, $abcd = 30$ and $ad+bc=-13$.

$30 = 2x3x5=10x3=(-10)(-3)$. That is, $ad = -10$ and $bc=-3$.

$5x^2-13x+6 = 5x^2-10x - 3x + 6 = 5x(x-2)-3(x-2) = (5x-3)(x-2)= 5(x-\frac{3}{5})(x-2).$

Examples

Solve: $x^2=8x$

That is, $0 = x^2-8x$

$= x(x-8)$

This means 0,8 are the roots of the given quadratic equation.

Solve: $x^2-4x+4=0$.

Using FOIL method, and comparing the coefficients, we get $abcd=4$ and $ad+bc=-4$. Therefore, $ad=-2$ and $bc=-2$.

So,

$$x^2-4x+4 = x^2-2x-2x+4$$
$$= x(x-2)-2(x-2)$$
$$= (x-2)^2=0$$

Hence, 2 is the repeated real root of the given equation.

Solve: $x^2-25=0$

Note $abcd = -25$ and $ad+bc = 0$.

That is, $ad=5$ and $bc=-5$

So,

$$x^2-25= x^2-5x+5x-25$$
$$= x(x-5)+5(x-5)$$
$$=(x+5)(x-5)=0$$

Hence, -5, 5 are the roots of the given quadratic equation.

Quadratic Equations

Solve by Completing the Square

Solving a Quadratic Equations by Completing the Square

Old Method:

$x^2+10x-24=0$

abcd=-24 and ad+bc= 10

ad=12, and bc=-2. So

$x^2+10x-24=x^2+12x - 2x -24$

$=x(x+12)-2(x+12)$

$=(x+12)(x-2)=0$

That is, -12 and 2 are the real roots of the equation.

New Method:

$x^2+10x=24$

Observe that $(x+a)^2= x^2+2ax+a^2$. Using this write 10=2x5 and add 25 on both sides of the equation to get
$x^2+10x + 25= 24 +25 = 49$

$(x+5)^2=7^2$
$(x+5) = \mp 7$

Therefore, x = -5+7=2 and x=-5-7=-12 are the roots of the quadratic equation.

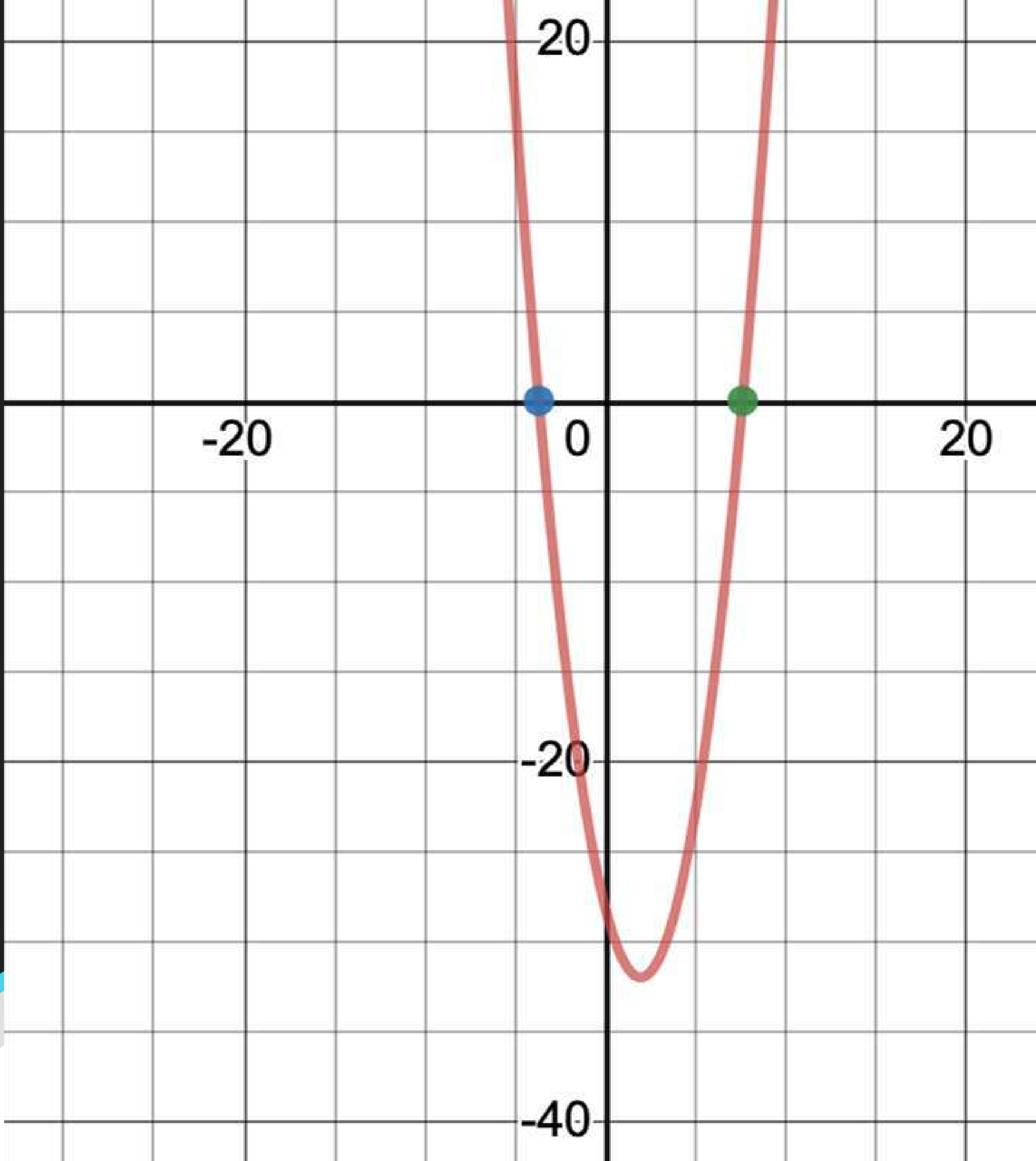
Quadratic Equations with Irrational Roots

Solve: $x^2-4x+4=32$

It can be easily seen that $(x-2)^2=32$.

Hence, $(x-2) = \pm\sqrt{32} = \pm4\sqrt{2}$.

Thus, $x=2\pm4\sqrt{2}$ are the roots of the quadratic equation.



Quadratic Formula

b^2-4ac	roots
>0	2
<0	0
=0	1

$$ax^2+bx+c=0$$

$$x^2+\frac{b}{a}x+\frac{c}{a}=0$$

$$x^2+\frac{b}{a}x+\frac{b^2}{4a^2}=+\frac{b^2}{4a^2}-\frac{c}{a}$$

$$(x+\frac{b}{2a})^2=\frac{b^2-4ac}{2a}$$

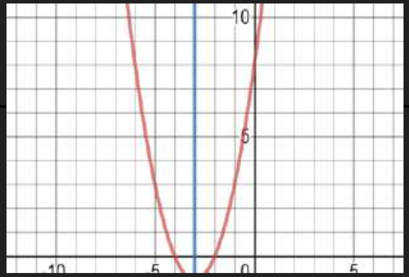
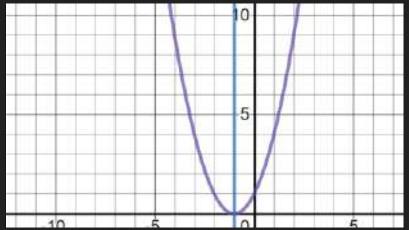
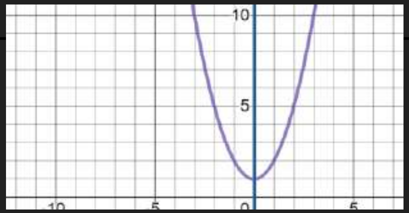
$$(x+\frac{b}{2a})=\pm\frac{\sqrt{b^2-4ac}}{2a}$$

$$x=-\frac{b}{2a}\pm\frac{\sqrt{b^2-4ac}}{2a}$$

The above formula is known as quadratic formula.

The quantity in the square root is known as discriminant.

Summary of Quadratic Formula

Value of the discriminant	Type and number of roots	Example
b^2-4ac >0 perfect square b^2-4ac >0, no perfect square	2 real, rational roots. 2 real, irrational roots.	 
b^2-4ac =0	1 real, rational root.	
b^2-4ac <0	No real root.	

Consider $ax^2+bx+c=0$, where a,b, and c are rational numbers.

Examples

Find the value of the discriminant for each equation and then describe the number and type of the roots for the equation.

1. $9x^2-12x+4=0$
2. $2x^2+16x+33=0$

1. $b^2-4ac = (-12)^2-4(9)(4) = 144 - 144 = 0$ so, it has one rational root.
2. $b^2-4ac = (16)^2-4(2)(33) = 256-264 = -8$ so, it has no real roots.

Axis of Symmetry

Why *x=-b/2a* is the axis of symmetry?

$$\begin{aligned}
f(x) &= ax^2 + bx + c \\
&= a(x^2 + (b/a)x + c/a) \\
&= a(x^2 + (b/a)x + b^2/(4a^2) - b^2/(4a^2) + c/a) \\
&= a(x + b/2a)^2 + (c - b^2/(4a))
\end{aligned}$$

Therefore, the symmetry is about *x=-b/(2a) which is the axis of symmetry.*

Summary of Concepts

Method	Can be used	When preferred
Graphing	Occasionally	Best used to verify the answer found algebraically
Factoring	Occasionally	If constant term is zero or factors are easy to find.
Completing the square	Always	Use when b is even.

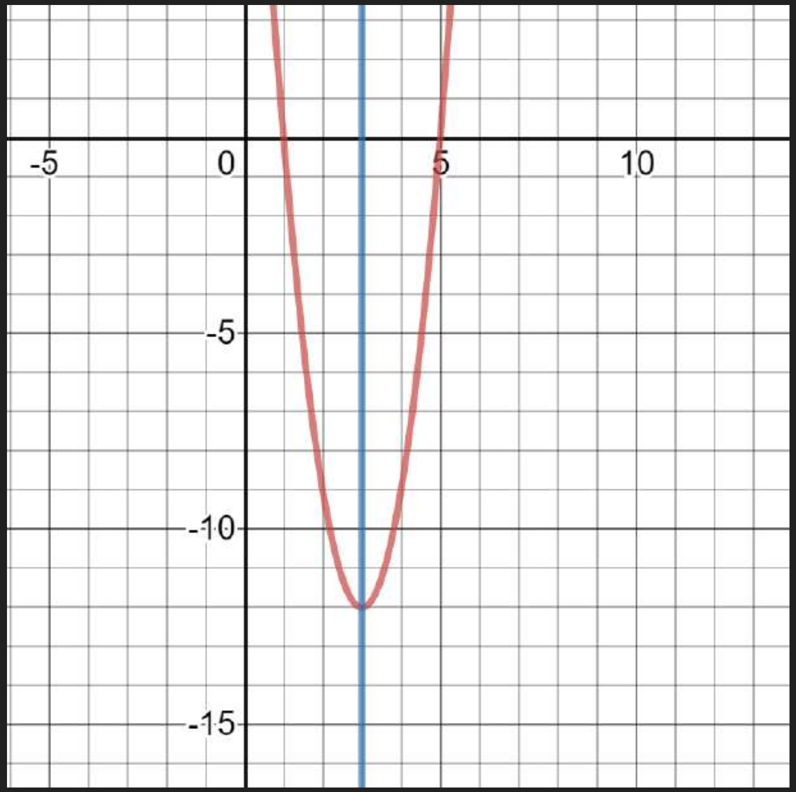
Quadratic Equations

Solve by Factoring

Quadratic Function: Intercept form

Let $y = f(x) = a(x-p)(x-q)$, where p and q represent x-intercepts for the function. Then the form $y = a(x-p)(x-q)$ is called the *intercept form*.

Example: Graph $y=3(x-1)(x-5)$



Question: How will you convert the intercept form into the standard form?

Intercept form to the Standard form

Changing intercept form to standard form requires us to use FOIL method which can be described as follows:

The product of two binomials is the sum of the products of the first(F) terms, the outer(O), the inner(I) and the last(L) terms.

$$(ax+b)(cx+d) = \underbrace{ax.\,cx} + \underbrace{ax.\,d} + \underbrace{b.\,cx} + \underbrace{b.\,d}$$

F O

I L

Quick Observations:

The product of coefficient of x^2 and the coefficient of the constant is *abcd*.
The product of the two terms in the coefficient of x is also *abcd*.

Example

Question. Write a quadratic equation with roots, ⅔ and -4, in the standard form.

Recall: By standard form, we mean $ax^2+bx+c=0$, where a,b,c are integers.

By intercept form, we know $(x-\frac{2}{3})(x+4)=0$.

By FOIL method, $(x-\frac{2}{3})(x+4)= x^2+(-\frac{2}{3} +4)x -\frac{2}{3}.4 = x^2+(10/3)x - 8/3=0$

For standard form, multiply both sides by 3, to get

$$3x^2+10x-8=0.$$

Standard form to Intercept form

Example: Convert the function $f(x) = 5x^2-13x+6$ to intercept form.

Let us apply FOIL Method.

$5x^2-13x+6 = (ax+b)(cx+d) = ac\,x^2+(ad+bc)x+bd.$

Therefore, $ac = 5$, $ad+bc = -13$ and $bd = 6$. That is, $abcd = 30$ and $ad+bc = -13$.

$30 = 2 \times 3 \times 5 = 10 \times 3 = (-10)(-3)$. That is, $ad = -10$ and $bc = -3$.

$5x^2-13x+6 = 5x^2-10x - 3x + 6 = 5x(x-2)-3(x-2) = (5x-3)(x-2) = 5(x-\frac{3}{5})(x-2).$

Examples

Solve: $x^2=8x$

That is, $0 = x^2-8x$

$= x(x-8)$

This means 0,8 are the roots of the given quadratic equation.

Solve: $x^2-4x+4=0$.

Using FOIL method, and comparing the coefficients, we get $abcd=4$ and $ad+bc=-4$. Therefore, $ad =-2$ and $bc=-2$.

So,

$$
\begin{aligned}
x^2-4x+4 &= x^2-2x -2x+4 \\
&= x(x-2)-2(x-2) \\
&= (x-2)^2=0
\end{aligned}
$$

Hence, 2 is the repeated real root of the given equation.

Solve: $x^2-25=0$

Note $abcd = -25$ and $ad+bc = 0$.

That is, $ad=5$ and $bc=-5$

So,

$$
\begin{aligned}
x^2-25 &= x^2-5x+5x-25 \\
&= x(x-5)+5(x-5) \\
&=(x+5)(x-5)=0
\end{aligned}
$$

Hence, -5, 5 are the roots of the given quadratic equation.

Quadratic Equations

Solve by Completing the Square

Solving a Quadratic Equations by Completing the Square

Old Method:

$x^2+10x-24=0$

abcd=-24 and ad+bc= 10

ad=12, and bc=-2. So

$x^2+10x-24=x^2+12x - 2x -24$

$=x(x+12)-2(x+12)$

$=(x+12)(x-2)=0$

That is, -12 and 2 are the real roots of the equation.

New Method:

$x^2+10x=24$

Observe that $(x+a)^2= x^2+2ax+a^2$. Using this write 10=2x5 and add 25 on both sides of the equation to get
$x^2+10x + 25= 24 +25 = 49$

$(x+5)^2=7^2$
$(x+5) = \mp 7$

Therefore, x = -5+7=2 and x=-5-7=-12 are the roots of the quadratic equation.

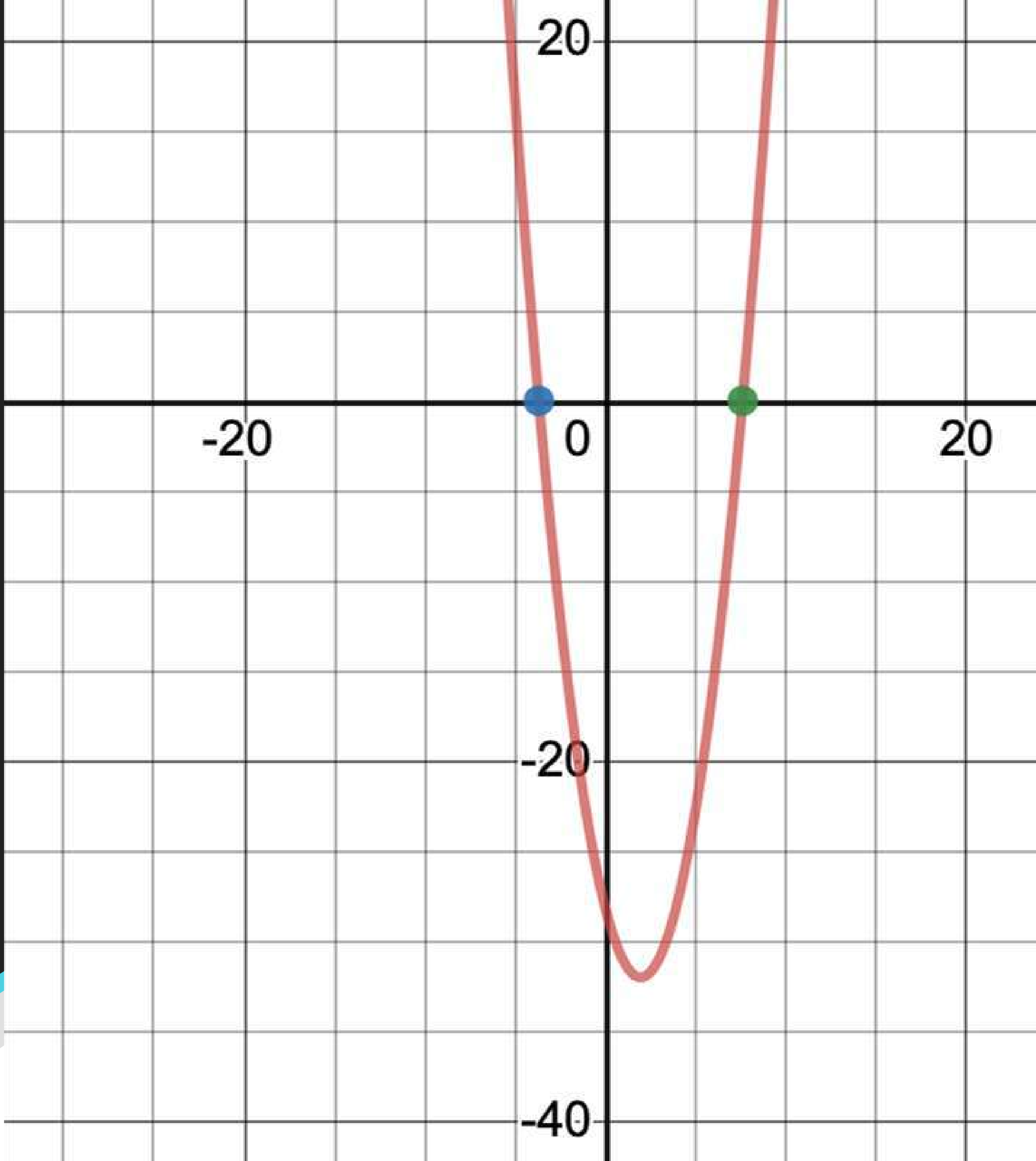
Quadratic Equations with Irrational Roots

Solve: $x^2-4x+4=32$

It can be easily seen that $(x-2)^2=32$.

Hence, $(x-2) = \pm\sqrt{32} = \pm4\sqrt{2}$.

Thus, $x=2\pm4\sqrt{2}$ are the roots of the quadratic equation.



Quadratic Formula

b^2-4ac	roots
>0	2
<0	0
=0	1

$$ax^2 + bx + c = 0$$

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = +\frac{b^2}{4a^2} - \frac{c}{a}$$

$$(x + \frac{b}{2a})^2 = \frac{b^2 - 4ac}{2a}$$

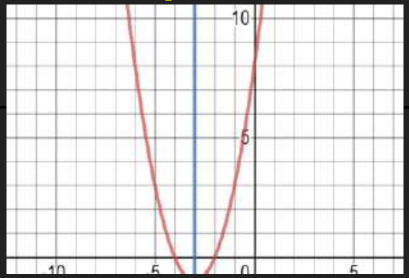
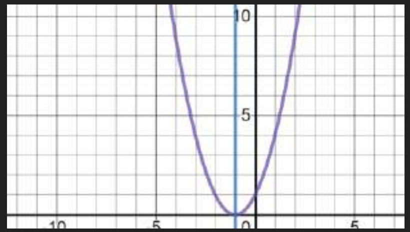
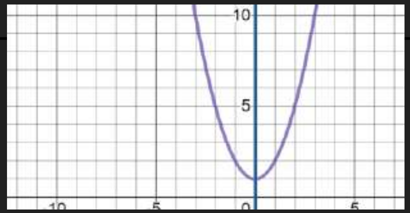
$$(x + \frac{b}{2a}) = \pm\frac{\sqrt{b^2 - 4ac}}{2a}$$

$$x = -\frac{b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a}$$

The above formula is known as quadratic formula.

The quantity in the square root is known as discriminant.

Summary of Quadratic Formula

Value of the discriminant	Type and number of roots	Example
$b^2-4ac > 0$ perfect square $b^2-4ac > 0$, no perfect square	2 real, rational roots. 2 real, irrational roots.	
$b^2-4ac = 0$	1 real, rational root.	
$b^2-4ac < 0$	No real root.	

Consider $ax^2+bx+c=0$, where a,b, and c are rational numbers.

Examples

Find the value of the discriminant for each equation and then describe the number and type of the roots for the equation.

1. $9x^2-12x+4=0$
2. $2x^2+16x+33=0$

1. $b^2-4ac = (-12)^2-4(9)(4) = 144 - 144 = 0$ so, it has one rational root.
2. $b^2-4ac = (16)^2-4(2)(33) = 256-264 = -8$ so, it has no real roots.

Axis of Symmetry

Why *x=-b/2a* is the axis of symmetry?

$$
\begin{aligned}
f(x) &= ax^2 + bx + c \\
&= a(x^2 + (b/a)x + c/a) \\
&= a(x^2 + (b/a)x + b^2/(4a^2) - b^2/(4a^2) + c/a) \\
&= a(x + b/2a)^2 + (c - b^2/(4a))
\end{aligned}
$$

Therefore, the symmetry is about *x=-b/(2a) which is the axis of symmetry.*

Summary of Concepts

Method	Can be used	When preferred
Graphing	Occasionally	Best used to verify the answer found algebraically
Factoring	Occasionally	If constant term is zero or factors are easy to find.
Completing the square	Always	Use when b is even.

Polynomials

Introduction

What is a Polynomial?

A Layman’s Perspective:

A Polynomial is one kind of mathematical expression which is a sum of several mathematical terms.

Each term in this expression is called monomial and the term can be a number, a variable or product of several variables.

Definition: (A mathematician’s Perspective)

A polynomial is an algebraic expression in which the only arithmetic is addition, subtraction, multiplication and “natural” exponents of the variables.

Examples:

Why do we call them Polynomials?

The word "Polynomial" is derived from two words

Poly + Nomen

many

name

Each term is called monomial.
A polynomial having two terms is called binomial.
A polynomial with three terms is called trinomial.

Eg. A polynomial in one variable can be represented as

Coefficient of the term

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = \sum_{m=0}^{n} a_m x^m .$$

Exponent

variable

Identification of Polynomials

Identify whether the following are polynomials or not.

1. x^2+4x+2
2. $x+x^{1/2}$
3. $x+y+xy+x^3$

Analysis:

Types of Polynomials

Polynomials in one variable

Eg. x^4+1

Polynomials in two variables

Eg. $x^4 + y^5 + xy$

Polynomials in more than two variables

Eg. $xyz + x^2z^5$

- The degree of zero polynomial is undefined.

The Degree of the Polynomial

- The exponent on the variable in a term is called the degree of that variable in that term.
- The degree of that term is the sum of the degrees of the variables in that term.
- The degree of the polynomial is the largest degree of any one of the terms with non-zero coefficients.

Examples: $x = x^1$ and $c = c.x^0$

Classification based on the degree of the polynomial

Degree	Name	Example
0	Constant Polynomial	c, 1, 5
1	Linear Polynomial	$2x+4$, $ax+b$
2	Quadratic Polynomial	$3x^2+2$, $4xy+2x$
3	Cubic Polynomial	$3x^3$, $4x^2y + 2y+1$
4	Quartic Polynomial	$10x^4+y^4$, $x^4 +10x+1$

Polynomia

Algebra with Polynomia

Polynomials in One Variable

Description: As seen earlier, the polynomial of degree n, is represented as

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0.$$

This expression can be treated as a function from IR $\longrightarrow$ IR.

That is, the domain of $p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ is IR, and the range depends on the function.

Addition of Polynomials

Add the following polynomials:

1. $p(x) = x^2+4x+4,\ q(x) = 10$
2. $p(x) = x^4+4x,\ q(x) = x^3+1$
3. $p(x)= x^3+2x^2+x,\ q(x)= x^2+2x+2$

$$p(x) = 1x^2+4x+4$$
$$q(x) = 0x^2+0x+10$$

$$p(x)+q(x) = x^2+4x+14$$

$$p(x) = 1x^4+0x^3+0x^2+4x+0$$
$$q(x) = 0x^4+ x^3+0x^2+0x+1$$

$$p(x)+q(x) = x^4+x^3+4x+1$$

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

$$p(x)+q(x) = x^3+(2+1)x^2+(1+2)x+2 = x^3+3x^2+3x+2$$

Let $p(x) = \sum_{k=0}^{n} a_k x^k$, and $q(x) = \sum_{j=0}^{m} b_j x^j$. Then $p(x) + q(x) = \sum_{k=0}^{m \vee n} (a_k + b_k) x^k$.

Subtraction of Polynomials

Subtract the following polynomials:

1. $p(x) = x^2+4x+4,\ q(x) = 10$
2. $p(x) = x^4+4x,\ q(x) = x^3+1$
3. $p(x)= x^3+2x^2+x,\ q(x)= x^2+2x+2$

$$\begin{aligned} p(x) &= 1x^2+4x+4 \\ -q(x) &= -0x^2-0x-10 \end{aligned}$$

$$p(x)-q(x) = x^2+4x-6$$

$$\begin{aligned} p(x) &= 1x^4+0x^3+0x^2+4x+0 \\ -q(x) &= -0x^4-x^3-0x^2-0x-1 \end{aligned}$$

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

$$\begin{aligned} p(x) &= 1x^3+2x^2+x+0 \\ -q(x) &= -0x^3-1x^2-2x-2 \end{aligned}$$

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

$$\textit{Let } p(x) = \sum_{k=0}^{n} a_k x^k, \textit{ and } q(x) = \sum_{j=0}^{m} b_j x^j. \textit{ Then}$$

$$p(x) - q(x) = \sum_{k=0}^{m \vee n} (a_k - b_k) x^k.$$

Multiplication of Polynomials

Multiply the following polynomials

$p(x) = x^2+x+1$ and $q(x) = 2x^3$

$$p(x)q(x) = (x^2+x+1)(2x^3)$$

$$= 2x^{3+2}+2x^{1+3}+2x^3$$

$$= 2x^5+2x^4+2x^3.$$

Multiply the following polynomials

$p(x) = x^2+x+1$ and $q(x) = 2x+1$

$$p(x)q(x) = (x^2+x+1)(2x+1)$$

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

$$= 2x^{1+2}+2x^{1+1}+2x + x^2+x+1$$

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

$$= 2x^3+3x^2+3x+1.$$

Multiplication of Polynomials

Multiply the polynomials $p(x) = a_2x^2+a_1x+a_0$ *and* $q(x) = b_1x+b_0$.

$$p(x)q(x) = (a_2x^2+a_1x+a_0)(b_1x+b_0)$$

$$= (a_2x^2+a_1x+a_0)(b_1x) + (a_2x^2+a_1x+a_0)b_0 \,.$$

$$= (a_2b_1x^{2+1}+a_1b_1x^{1+1}+a_0b_1x) + (a_2b_0x^2+a_1b_0x+a_0b_0)$$

$$= a_2b_1x^3+(a_1b_1+a_2b_0)x^2+(a_0b_1+a_1b_0)x+a_0b_0.$$

$Let\ p(x) = \sum_{k=0}^{n} a_k x^k,\ and\ q(x) = \sum_{j=0}^{m} b_j x^j.\ Then$

$$p(x)q(x) = \sum_{k=0}^{m+n} \sum_{j=0}^{k} (a_j b_{k-j}) x^k.$$

Multiplication of Polynomials

k	a_k	b_k
0	1	1
1	1	2
2	1	1

Multiply the polynomials $p(x) = x^2+x+1$ and $q(x) = x^2+2x+1$

$$Let\ p(x) = \sum_{k=0}^{n} a_k x^k,\ and\ q(x) = \sum_{j=0}^{m} b_j x^j.\ Then$$

$$p(x)q(x) = \sum_{k=0}^{m+n} \sum_{j=0}^{k} (a_j b_{k-j}) x^k.$$

The resultant polynomial is:

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

k	Coefficient	Calculations
0	a_0b_0	1
1	$a_1b_0+a_0b_1$	1+2=3
2	$a_0b_2+a_1b_1+a_2b_0$	1+2+1=4
3	$a_0b_3+a_1b_2+a_2b_1+a_3b_0$	0+1+2+0=3
4	$a_0b_4+a_1b_3+a_2b_2+a_3b_1+a_4b_0$	0+0+1+0+0=1

Division of Polynomials

Division of a polynomial by a monomial

$$\frac{3x^2+4x+3}{x} = 3x + 4 + \frac{3}{x}$$

Division of a polynomial by another polynomial

$$\frac{3x^2+4x+3}{2x+1} = ???$$

Division of Polynomials

$$\frac{3x^2+4x+1}{x+1} = (3x+1)$$

Divide $p(x) = x^4+2x^2+3x+2$ by $q(x)=x^2+x+1$.

Dividend

Quotient

Remainder

$$\frac{p(x)}{q(x)} = x^2-x+2 + \frac{2x}{q(x)}$$

Divisor

Division of Polynomials

$$\frac{3x^2+4x+1}{x+1} = (3x+1)$$

Divide $p(x) = x^4+2x^2+3x+2$ by $q(x)=x^2+x+1$.

$$\frac{\overset{\text{Dividend}}{p(x)}}{\underset{\text{Divisor}}{q(x)}} = \overset{\text{Quotient}}{x^2-x+2} + \frac{\overset{\text{Remainder}}{2x}}{q(x)}$$

Division of Polynomials

Find $\dfrac{2x^3+3x^2+1}{2x+1} = x^2+x-½ + 3/2/(2x+1)$

Division Algorithm

Step 1. Arrange the terms in descending order of the degree and add the missing exponents with 0 as coefficient.

Step 2. Divide the first term of the dividend by the first term of the divisor and get the monomial

Step 3. Multiply the monomial with divisor and subtract the result from the dividend.

Step 4.Check if the resultant polynomial has degree less than divisor.If true, write the remainder else Go to Step 2.

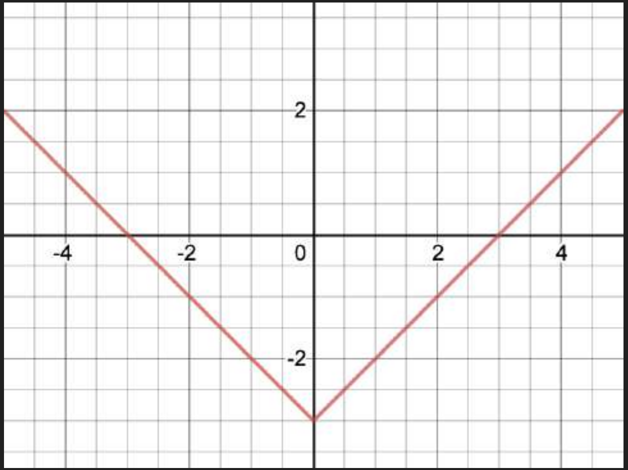
Polynomials

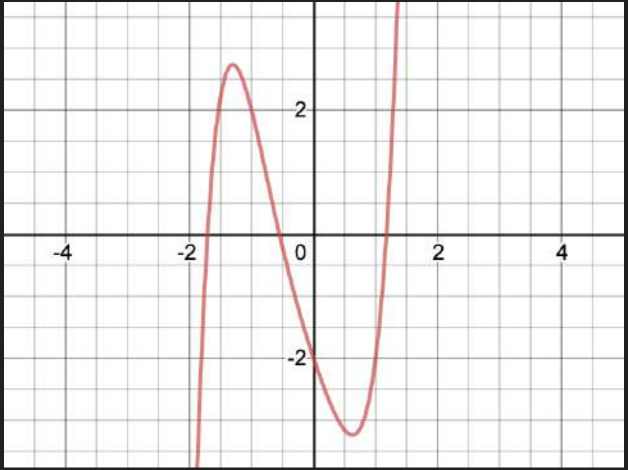
Graphs of Polynomials

Characterization of Graphs of Polynomial Functions

Polynomial functions of second degree or higher have graphs that do not have sharp corners. That is, the graphs are smooth curves.

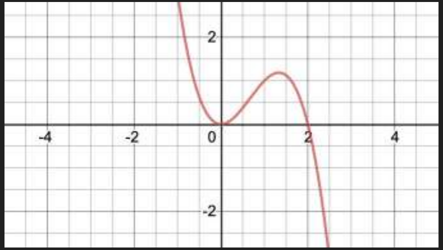
Polynomial functions also display graphs that have no breaks. Curves with no breaks are called continuous.

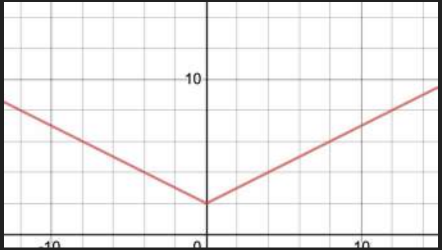


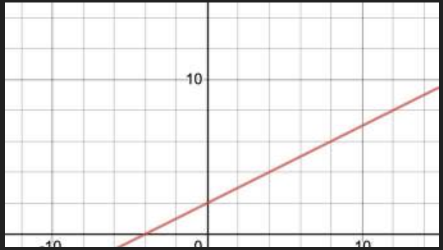


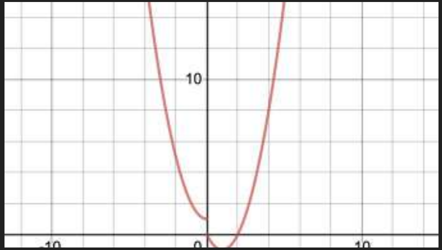
Identification of Polynomial Functions

Which of the graphs given below, represent polynomial functions?









Zeros of Polynomial Functions

Recall: If f is a polynomial function, the values of x for which $f(x)=0$ are called **zeros** of f.

If the equation of the polynomial function can be factored, we can set each factor equal to zero and solve for the zeros.

Also, any value $x=a$ that is a zero of a polynomial function yields a factor of the polynomial, of the form $(x-a)$.

Given the equation of a polynomial function, we can use this method to find x-intercepts because at the x-intercepts we find the input values whose output value is zero.

For general polynomials, this can be a challenging prospect. However quadratic functions can be solved using the quadratic formula.

The corresponding formulas for cubic and fourth-degree polynomials are not simple enough to remember. And formulas do not exist for general higher-degree polynomials.

Zeros of Polynomial Functions and Factoring

- The polynomial can be factored using known methods:
    a. greatest common factor,
    b. factor by grouping, and
    c. trinomial factoring.
- The polynomial is given in factored form.
- Technology is used to determine the intercepts.

x-intercept of Polynomial Function by Factoring

1. Set $f(x)=0$.
2. If the polynomial function is not given in factored form:

   a. Factor out any common monomial factors.

   b. Factor any factorable binomials or trinomials.

3. Set each factor equal to zero and solve to find the x-intercepts.

Example

Find x-intercepts of $f(x) = x^6 - 8x^4 + 16x^2$.

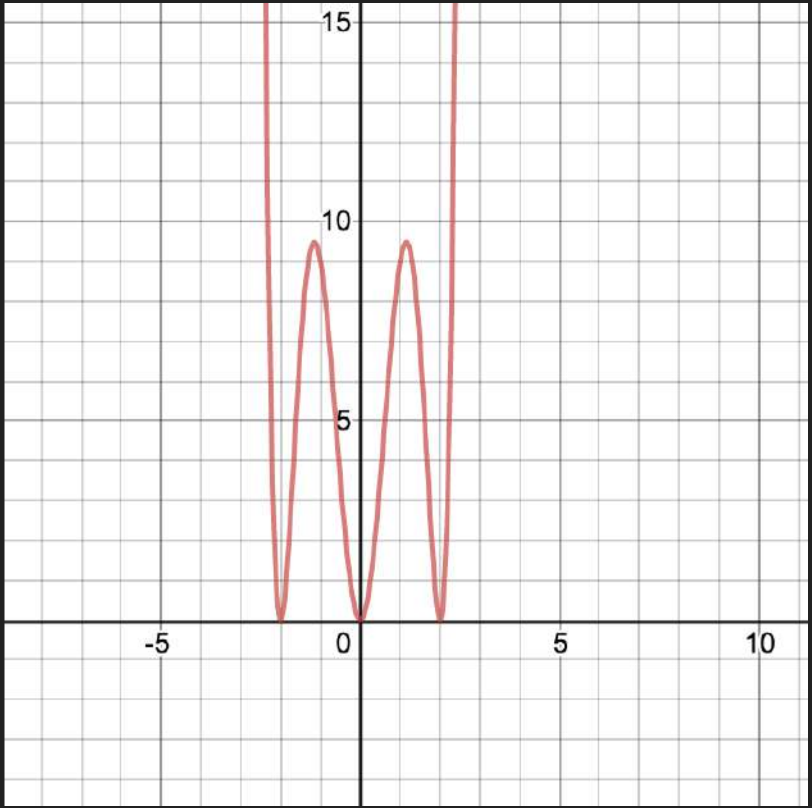
Set f(x)=0

$$x^6 - 8x^4 + 16x^2 = 0$$

$$x^2(x^4 - 8x^2 + 16) = 0$$

$$x^2(x^2 - 4)^2 = 0$$

$x = 0,\ 2, -2$ are the x-intercepts of f.



Example

Find x-intercepts of $f(x) = x^3 - 4x^2 - 3x + 12$.

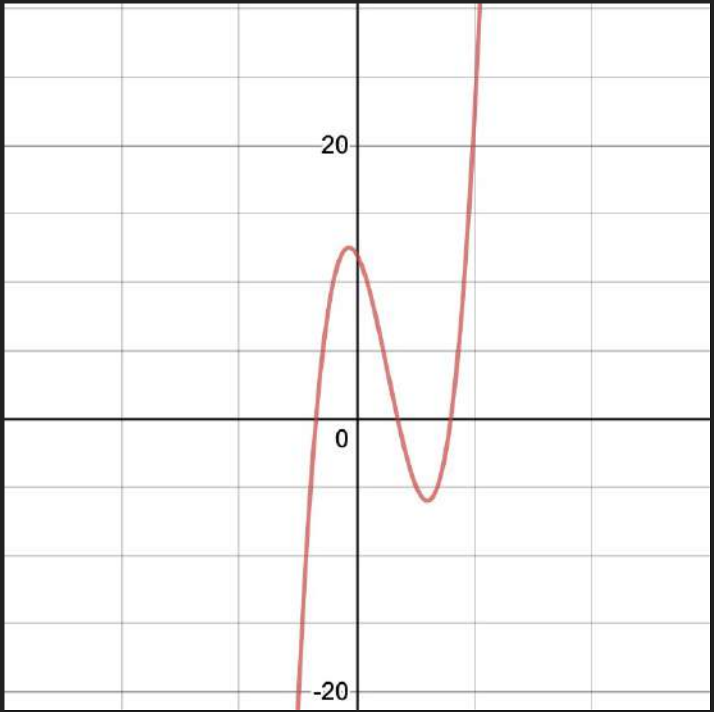
Set f(x)=0

$$x^3 - 4x^2 - 3x + 12 = 0$$

$$x^2(x-4) - 3(x-4) = 0$$

$$(x^2-3)(x-4)=0$$

$x=4, \sqrt{3}, -\sqrt{3}$ *are the x-intercepts of f.*



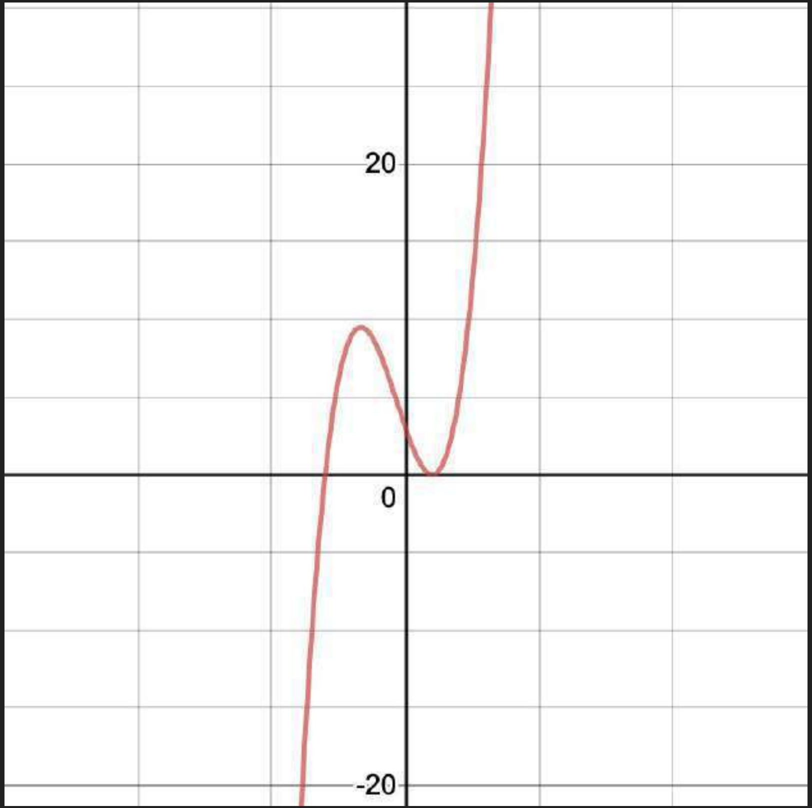
Example

Find the y- and x-intercepts of $g(x)=(x-1)^2(x+3)$.

Set $g(x)=0$

$x = 1, -3$ are the x-intercepts of f.

For y-intercept, $g(0) = 3$



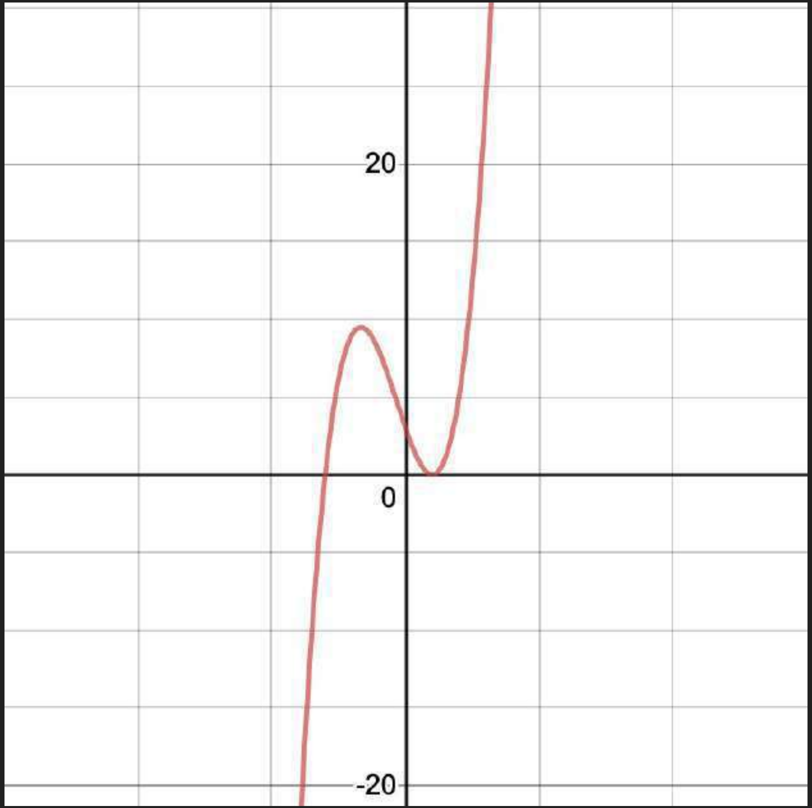
Example

Find the y- and x-intercepts of $g(x)=(x-1)^2(x+3)$.

Set $g(x)=0$

$x = 1, -3$ *are the* x*-intercepts of* f.

For y-intercept, $g(0) = 3$

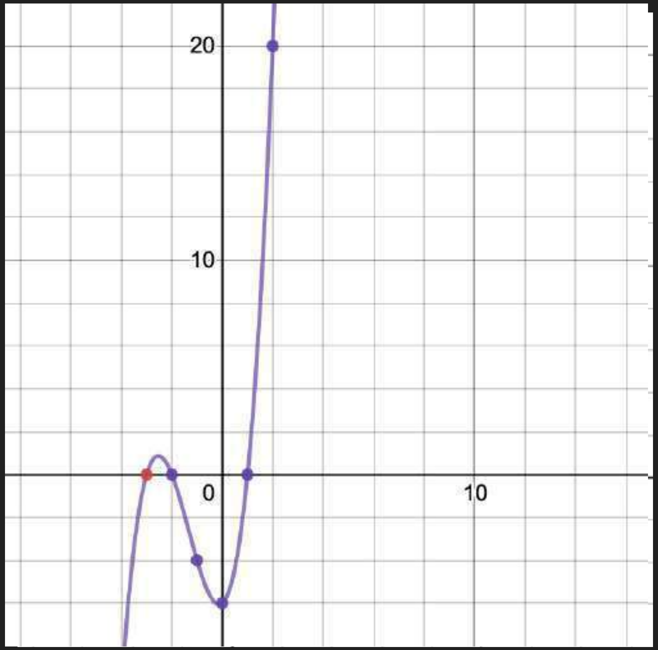


x-intercept of Polynomial Function using Graph

Find x-intercept of $f(x) = x^3+4x^2+x-6$

In this case, the polynomial is not in factored form, has no common factors, and does not appear to be factorable using techniques previously discussed.
The only option is to generate the pair of values as done in quadratic case.

x	y
-2	0
-1	-4
0	-6
1	0
2	20



From table, x=-2, 1 are the x-intercepts of f.
The third zero can be found by dividing f(x) by (x+2)(x-1).
The third zero of f is x=-3.
Therefore, join the points smoothly to get the graph.

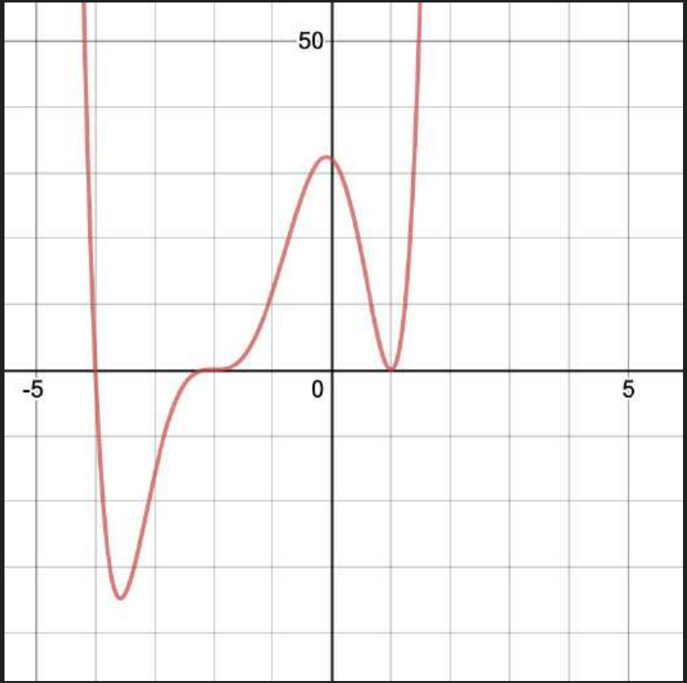
Identification of Zeros and Their Multiplicities

Graphs behave differently at various x-intercepts.

Sometimes, the graph will cross over the horizontal axis at an intercept.

Other times, the graph will touch the horizontal axis and "bounce off."

Suppose, for example, we graph the function $f(x) = (x-1)^2(x+2)^3(x+4)$.



Identifying Zeros and their Multiplicities

The x-intercept -4 is the solution of the equation $(x+4)=0$. The graph passes directly through the x-intercept at $x=-4$. The factor is linear (has a degree of 1), so the behavior near the intercept is like that of a line — it passes directly through the intercept. We call this a single zero because the zero corresponds to a single factor of the function.

The x-intercept 1 is the repeated solution of the equation $(x-1)^2=0$. The graph touches the axis at the intercept and changes direction. The factor is quadratic (degree 2), so the behavior near the intercept is like that of a quadratic — it bounces off of the horizontal axis at the intercept.

The x-intercept -2 is the repeated solution of the equation $(x+2)^3=0$.The graph passes through the axis at the intercept, but flattens out a bit first. This factor is cubic (degree 3), so the behavior near the intercept is like that of a cubic — with the same S-shape near the intercept as the toolkit function $f(x)=x^3$. We call this a triple zero, or a zero with multiplicity 3.

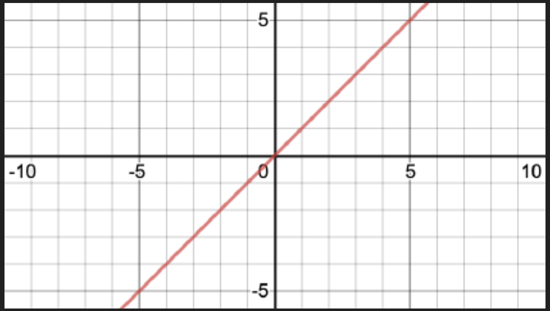
IDENTIFYING ZEROS AND THEIR MULTIPLICITIES

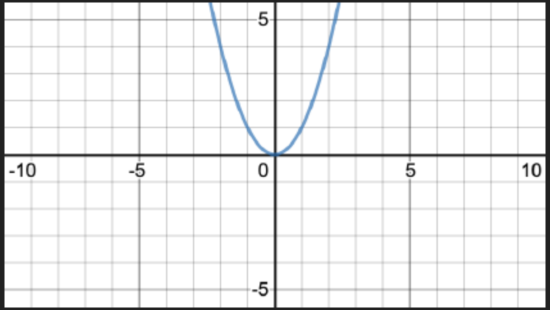
For **zeros** with even multiplicities, the graphs touch or are **tangent** to the x-axis.

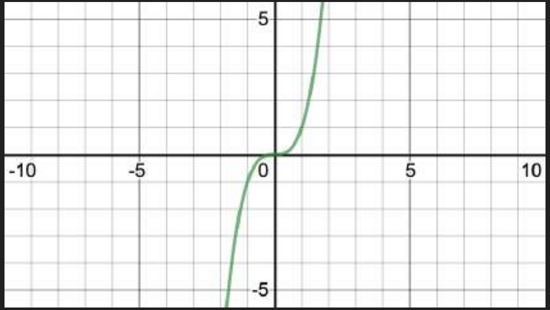
For zeros with odd multiplicities, the graphs cross, or **intersect**, the x-axis.

For higher even powers, such as 4, 6, and 8, the graph will still touch and bounce off of the horizontal axis but, for each increasing even power, the graph will appear flatter as it approaches and leaves the x-axis.

For higher odd powers, such as 5, 7, and 9, the graph will still cross through the horizontal axis, but for each increasing odd power, the graph will appear flatter as it approaches and leaves the x-axis.







Graphical Behavior of Polynomials at x-Intercepts

If a polynomial contains a factor of the form $(x-a)^m$, the behavior near the x-intercept a is determined by the exponent m. We say that x=a is a zero of **multiplicity m**.

The graph of a polynomial function will touch but not cross the x-axis at zeros with even multiplicities. The graph will cross the x-axis at zeros with odd multiplicities.

The sum of the multiplicities is no greater than the degree of the polynomial function.

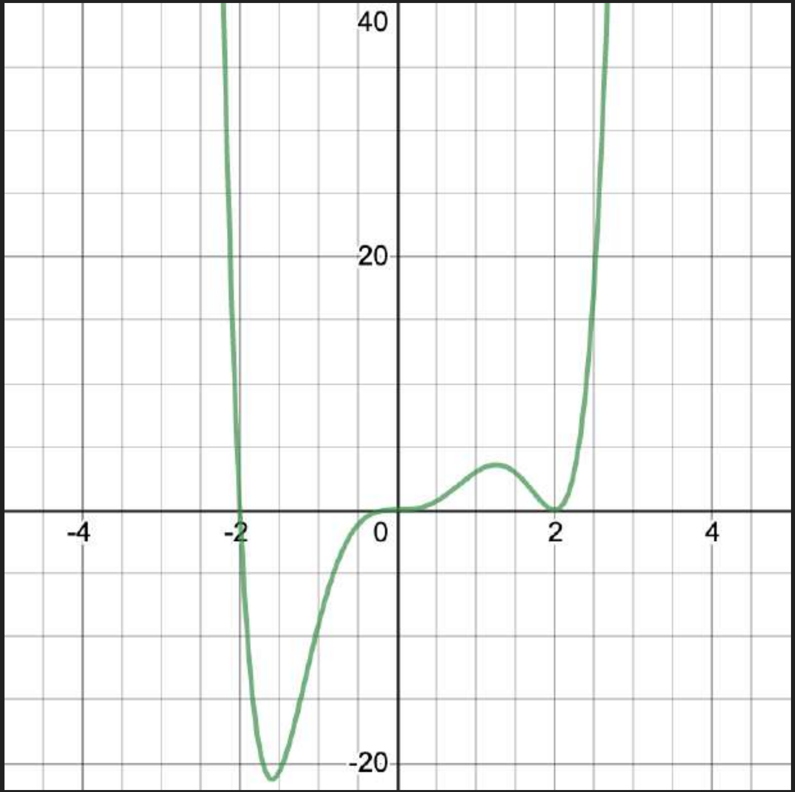
Graphical Behavior of Polynomials at x-Intercepts

Given the graph of a polynomial of degree n, how can one identify zeros and their multiplicities?

1. If the graph touches the x-axis and bounces off of the axis, it is a zero with even multiplicity.
2. If the graph crosses the x-axis, it is a zero with odd multiplicity.
3. If the graph crosses the x-axis and appears almost linear at the intercept, it is a single zero.
4. The sum of all the multiplicities is no greater than n.

Example

Use the graph of the function of degree 6 to identify the zeros of the function and their possible multiplicities.



x = -2, 0, 2

x=-2, linear, 1

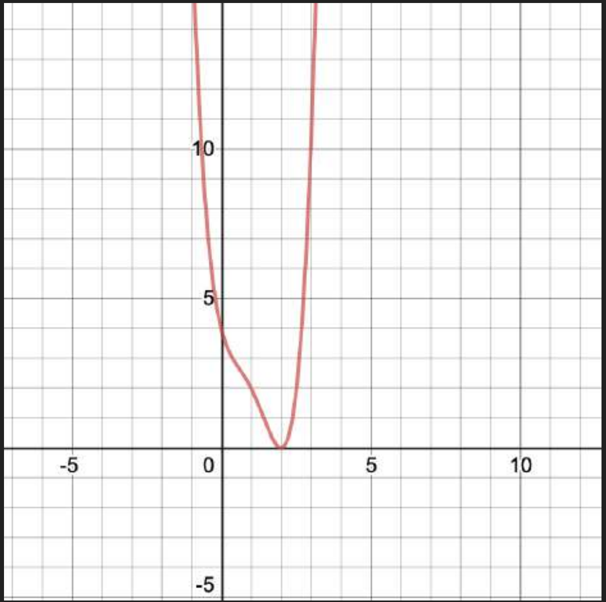
x=0, odd degree, 3 or 5

x=2, even degree, 2 or 4

x=0 with multiplicity 3 and x=2 with multiplicity 2 and x=-2 with multiplicity 1.

Example

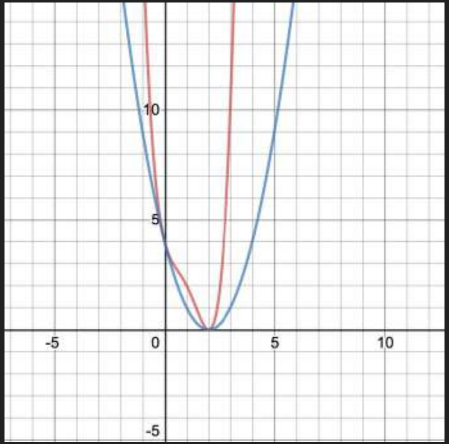
Use the graph of the function of degree 4 to identify the zeros of the function and their possible multiplicities.

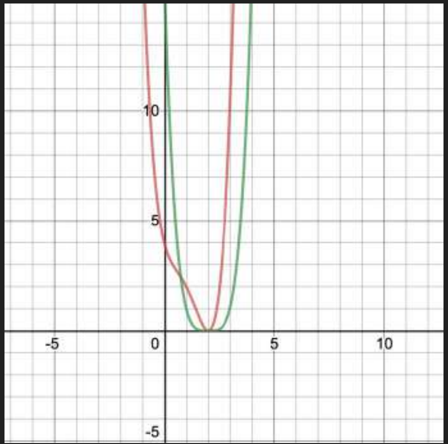


x = 2

x=2, even degree, 2 or 4

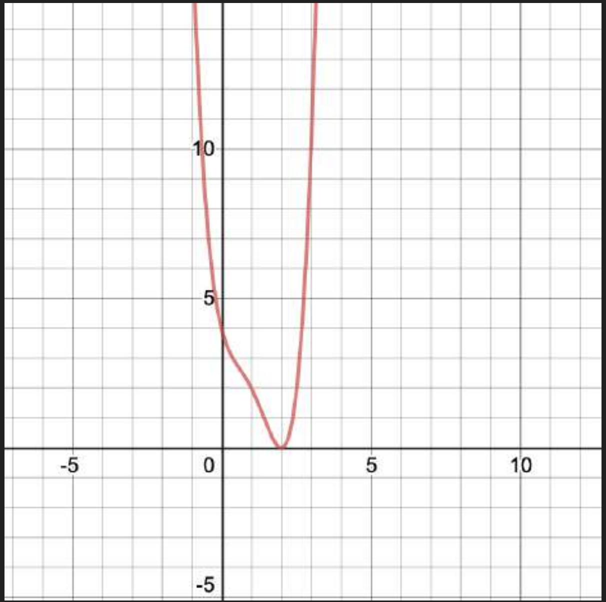
Hence, the function f(x) must have a factor $(x-2)^2$.





Example

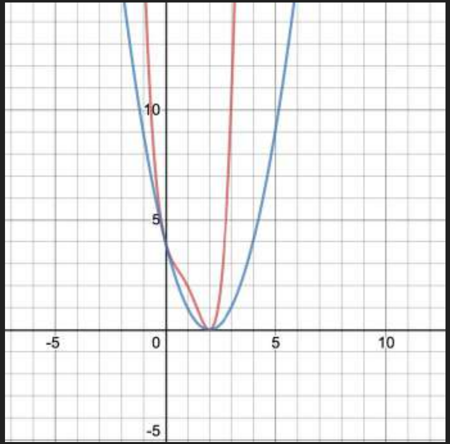
Use the graph of the function of degree 4 to identify the zeros of the function and their possible multiplicities.

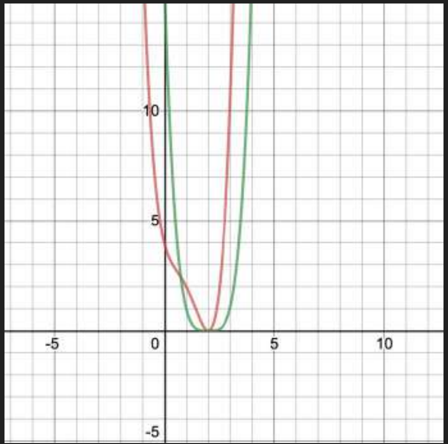


$x = 2$

$x=2$, even degree, 2 or 4

Hence, the function $f(x)$ must have a factor $(x-2)^2$.





End-Behavior of Polynomials

As we have already observed, the behavior of polynomial function

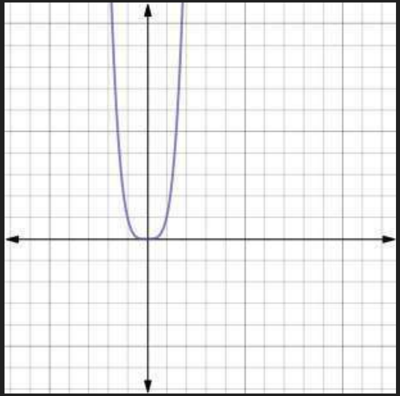
$$f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0$$

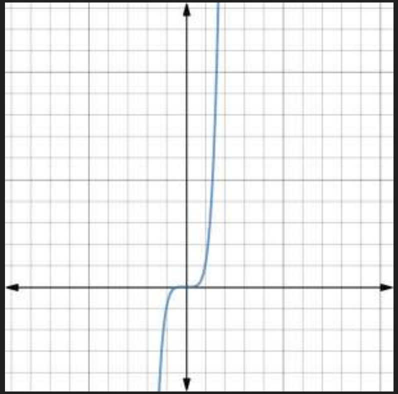
is either increasing or decreasing as the the value of x increases which is mainly due to the fact that the leading terms dominate the behavior of polynomial. This behavior is known as End behavior of the function.

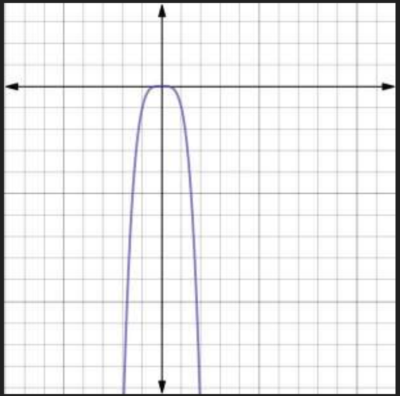
As observed in quadratic equations, if the leading term of a polynomial function, $a_n x^n$, is an even power function and $a_n > 0$, then as x increases or decreases, f(x) increases and is unbounded.
When the leading term is an odd power function, as x decreases, f(x) also decreases and is unbounded; as x increases, f(x) also increases and is unbounded.

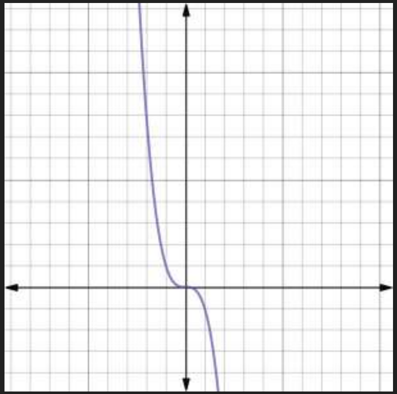
End-Behavior of Polynomials

	Even Degree	Odd Degree
$a_n > 0$	$x \to \infty,\ f(x) \to \infty$ $x \to -\infty,\ f(x) \to \infty$	$x \to \infty,\ f(x) \to \infty$ $x \to -\infty,\ f(x) \to -\infty$
$a_n < 0$	$x \to \infty,\ f(x) \to -\infty$ $x \to -\infty,\ f(x) \to -\infty$	$x \to \infty,\ f(x) \to -\infty$ $x \to -\infty,\ f(x) \to \infty$









Graphing a Polynomial Function

1. Find the x- and y- intercepts, if possible.
2. Check for symmetry. If the function is an even function, its graph is symmetrical about the y-axis, that is, $f(-x)=f(x)$. If a function is an odd function, its graph is symmetrical about the origin, that is, $f(-x)=-f(x)$.
3. Use the multiplicities of the zeros to determine the behavior of the polynomial at the x-intercepts.
4. Determine the end behavior by examining the leading term.
5. Use the end behavior and the behavior at the intercepts to sketch a graph.
6. Ensure that the number of turning points does not exceed one less than the degree of the polynomial.
7. Optionally, use graphing tools to check the graph.

Example

Sketch a graph of $f(x)=-(x+2)^2(x-5)$.

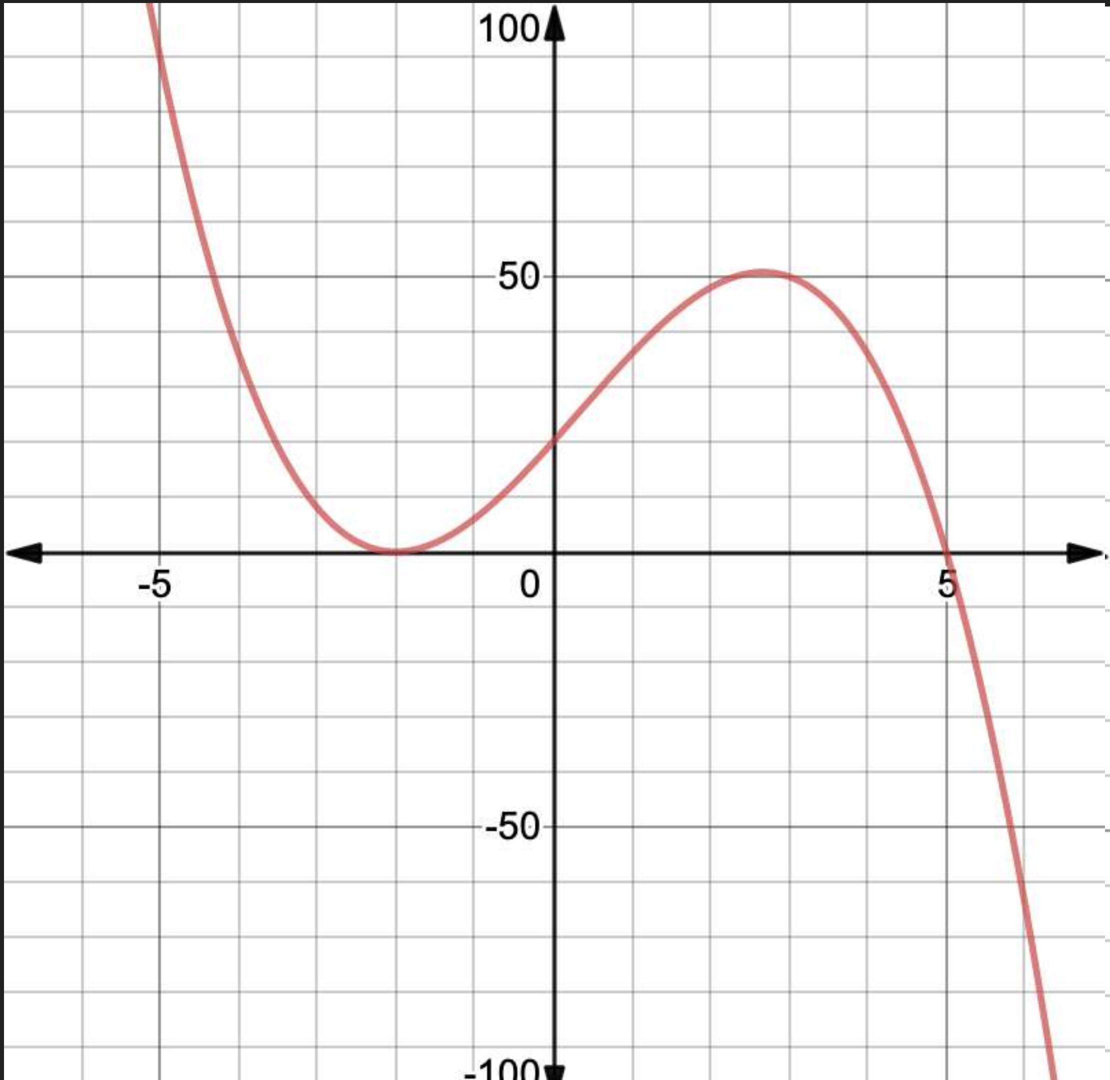
x-intercepts are x= -2, 5

x= -2 has multiplicity 2, quadratic
x=5 has multiplicity 1, linear

y-intercept f(0)=20.

The leading term is $-x^3$. Therefore, the odd degree polynomial with negative leading coefficient has the following end-behavior
$x \to \infty, \quad f(x) \to -\infty$
$x \to -\infty, \quad f(x) \to \infty$

f can have at most 3-1=2 turning points.



Intermediate Value Theorem

Let f be a polynomial function. The **Intermediate Value Theorem** states that if $f(a)$ and $f(b)$ have opposite signs, then there exists at least one value c between a and b for which $f(c)=0$.

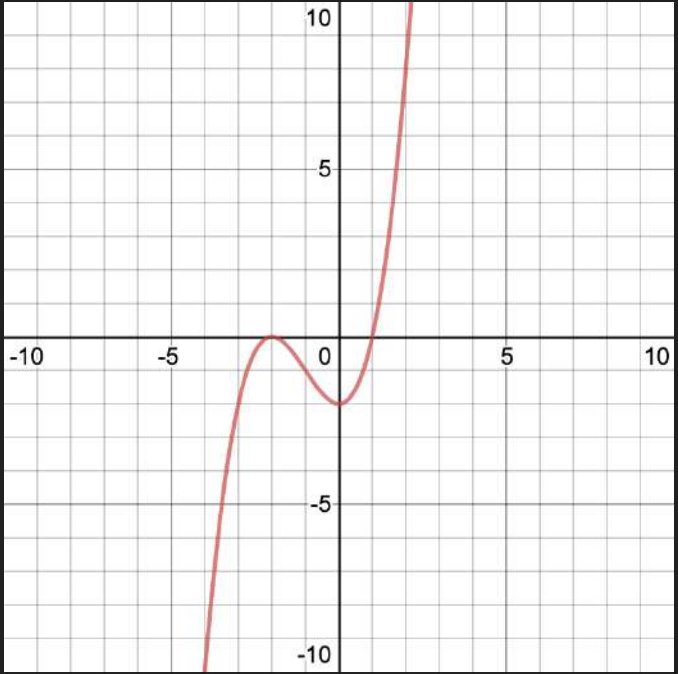
Deriving Formula for Polynomial Functions

Given the graph, how to find the formula for polynomial function?

1. Find the x-intercepts of the graph to find the factors of the polynomial.
2. Understand the behavior of the graph at the x-intercepts to determine the multiplicity of each factor.
3. Find the polynomial of least degree containing all the factors found in the previous step.
4. Use any other point on the graph (the y-intercept may be easiest) to determine the stretch factor (?).

Example

Write the formula for polynomial given in the graph.



x= -2,1 are the x-intercepts and the function has two turning points. The end behavior is similar to odd degree polynomial with positive leading term. That is, it may be be polynomial of degree 3.

The behavior at x =1 is linear and x=-2 is of even degree and hence quadratic. The resultant polynomial is of degree 3 with zeros -2 and 1 with multiplicities 2 and 1 respectively.

The polynomial has form $f(x) = a(x+2)^2(x-1)$.

To determine a, use y-intercept. From the graph, f(0) = -2
From the form f(0) = -4a. Therefore, a = ½.

Hence, the function must be $f(x) = \frac{1}{2}(x+2)^2(x-1)$.

Logarithmic Functions

Monday, 14 September 2020 9:56 AM

Recall. $f(x) = a^x$ $(a > 0, a \neq 1)$

is one-to-one, it has its inverse

Def^n. The logarithmic function (to the base

a) in standard form is

$$y = \log_a(x)$$

and is defined to be the inverse of

$f(x) = a^x$

$\log_a x = y$ $a^y = x$

7-rule

$$y = \log_a x \iff x = a^y$$

$$a^{\log_a x} = x \quad \& \quad \log_a a^x = x$$

$$f(f^{-1}(x)) = x \qquad f^{-1}(f(x)) = x$$

$\mathrm{Dom}(a^x) = \mathbb{R}$ $\qquad$ $\mathrm{Range}(a^x) = (0,\infty)$

$\qquad \| \qquad\qquad\qquad\qquad \|$

$\mathrm{Range}(\log_a)$ $\qquad$ $\mathrm{Dom}(\log_a)$

☆ $\mathrm{Dom}(\log_a) = \mathrm{Range}(a^x) = (0,\infty)$

☆ $\mathrm{Dom}(a^x) = \mathrm{Range}(\log_a) = \mathbb{R}$

$Dom(a^x) = Range(\log_a) = \mathbb{R}$

Example. $\boxed{f(x) = \log_4(1-x)}$ Find the domain of f

$\boxed{Dom(\log_4) = (0, \infty)}$ $\qquad Dom(f) = (-\infty, 1)$

$1-x > 0 \iff 1 > x \qquad x < 0$

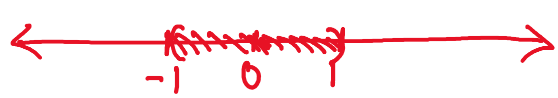
$1 - x > 0$

Example. $g(x) = \log_3\left(\frac{1+x}{1-x}\right)$, $x \neq 1$

$Dom(g) = (-1, 1)$ ✓

$Dom(\log_3) = (0, \infty)$

$\frac{1+x}{1-x} > 0$



Example $y = \log_3 x$

$$\underline{3}^{y} = \underline{3}^{\underline{\log_3 x}} = \underline{\underline{x}}$$

$$\underline{3^{y}} = \underline{x}$$

Example. $(1.3)^2 = m$

$$\log_{1.3} (1.3)^2 = \log_{1.3} m$$

$$a^{\log_a x} = x$$

$$2 = \log_{1.3} m$$

Observe $a^u = a^v \ (a > 0, a \neq 1)$

$$\Rightarrow u = v$$

Find $\log_3(\frac{1}{9})$

$$\log_3\left(\tfrac{1}{9}\right) = \log_3\left(3^{-2}\right) = \boxed{-2}$$

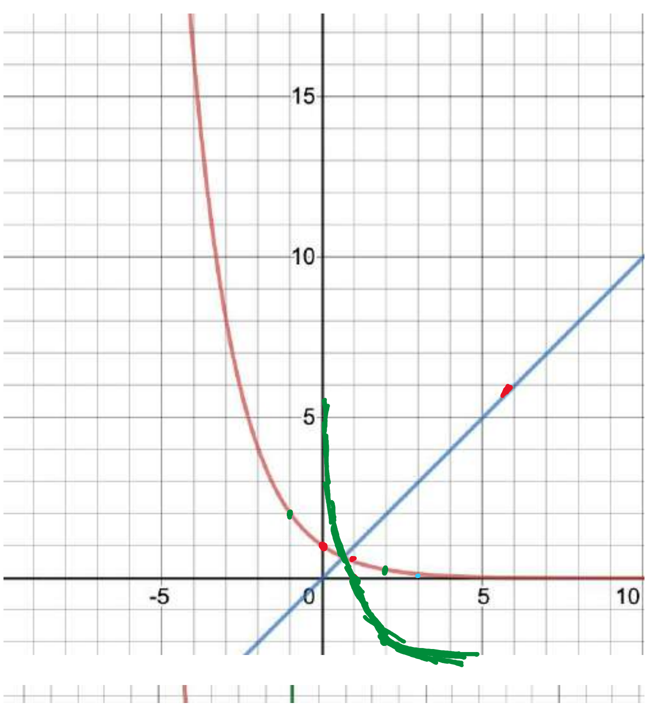
$$3^2 = 9 \Rightarrow 3^{-2} = \frac{1}{9}$$

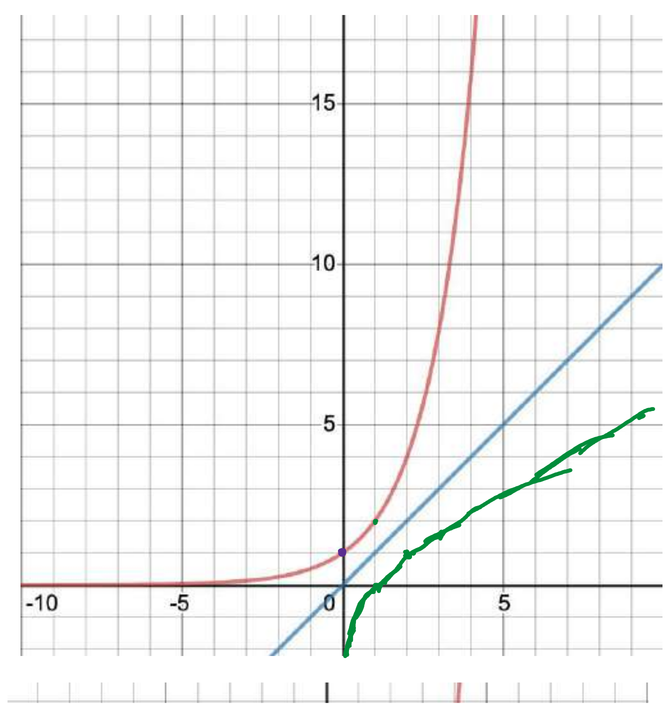
Graph $\underline{f(x) = \log_a x}$

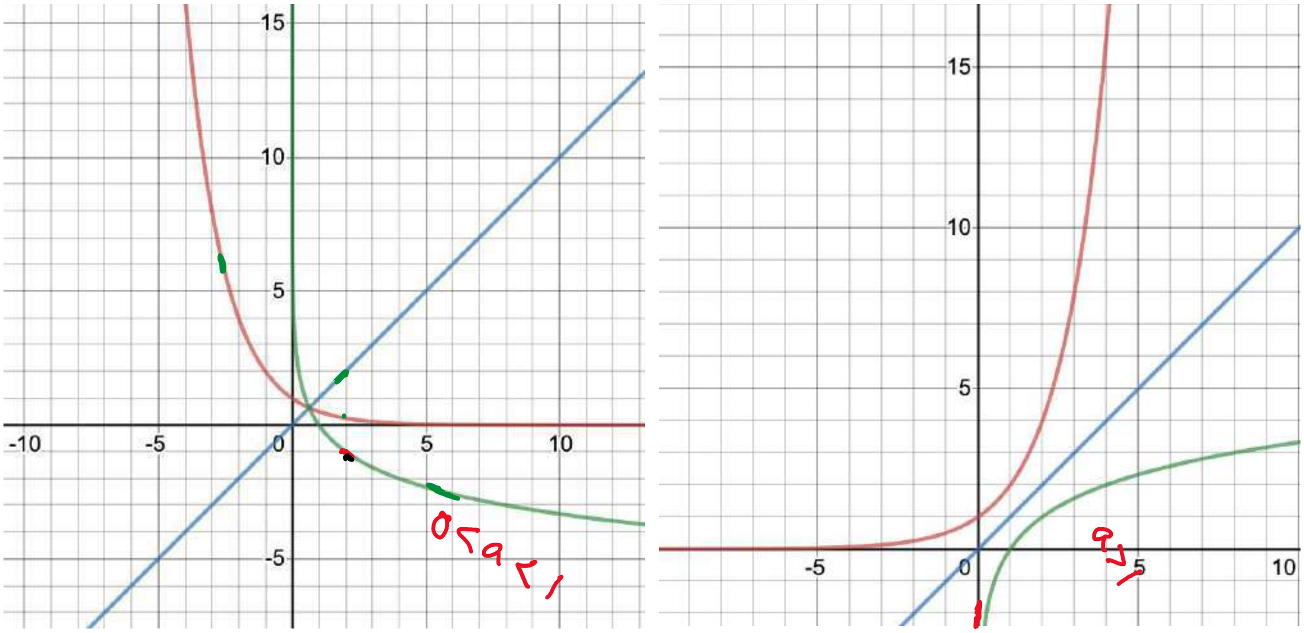
a   1   a>1

$*\ 0 < a < 1$

$a > 1$







Properties for $f(x) = \log_a(x)$

$Dom(f) = (0, \infty)$ $Range(f) = \mathbb{R}$

x-intercept : $(1, 0)$

y-intercept : Nil

Vertical asymptote at $x = 0$ (y-axis)

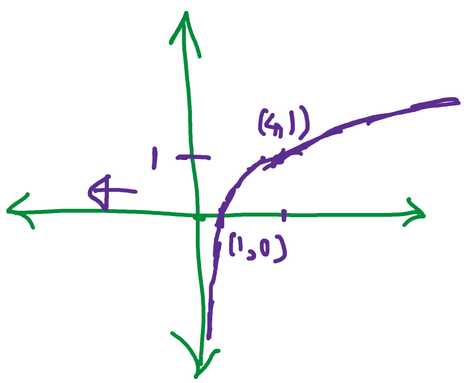
- f is one-to-one & passes through $(1,0)$ & $(a,1)$
- $0 < a < 1$, f is decreasing
- $a > 1$, f is increasing

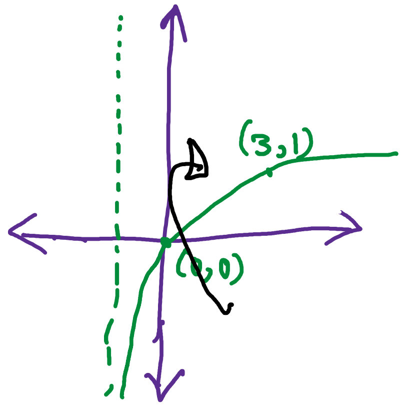
Example. Draw graphs of the functions

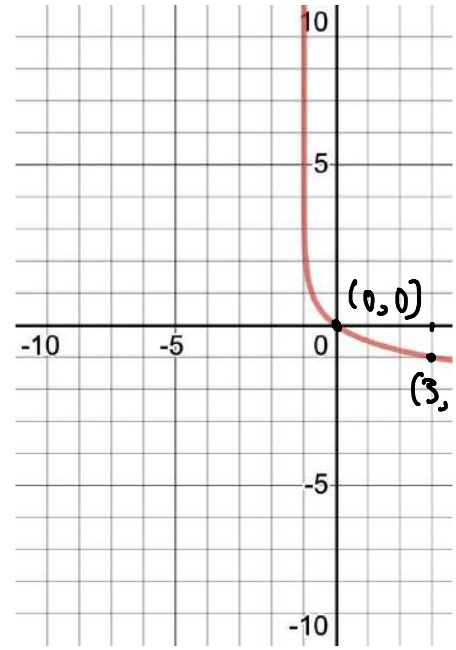
$f(x) = -\log_2(x+1)$

$$g(x) = \log_{1/4}(-x) + 1$$

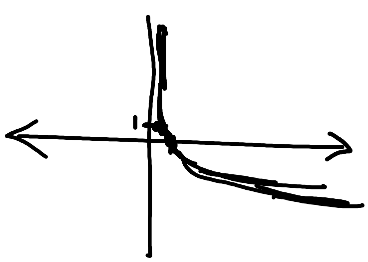
$\log_4(x)$ $\quad$ $-\log_4(x+1)$

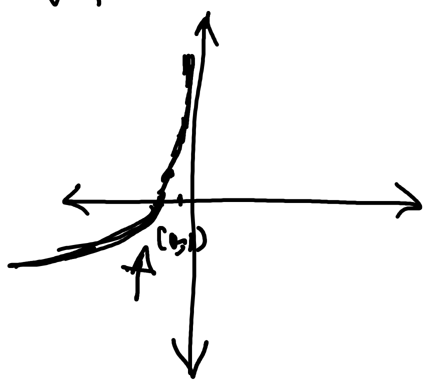


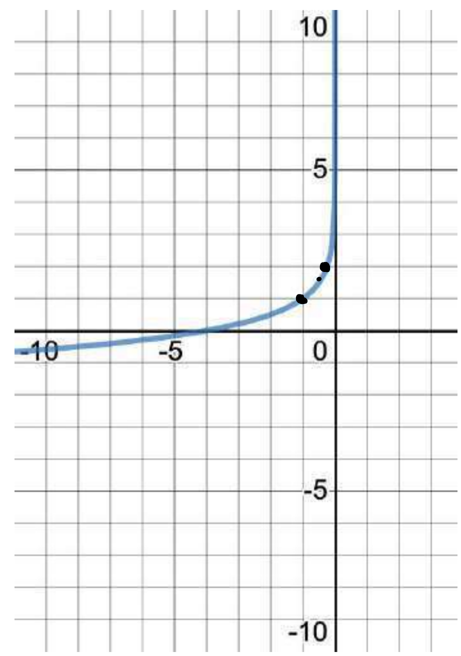




$\log_{1/4}(x)$ $\quad$ $\log_{1/4}(-x)+1$







Properties of Logarithmic Function

$a \in (0,1)$ or $a > 1$

Recall. $\boxed{\log_a 1 = 0} \longleftarrow (1,0)$ $\quad$ $(0,1)$

$$\boxed{\log_a a = 1}$$

$\cdot (a,1)$ $(1,a)$

$$a^{\log_a x} = x \quad \log_a(a^x) = x$$

$$3^{\log_3(\pi/2)} = \boxed{\pi/2} \qquad 4^{\log_4(1)} = 1$$

Laws of Logarithm

Let $x \in \mathbb{R}$, $0<a<1$ or $a>1$; $M, N > 0$.

Then

1. $\log_a(MN) = \log_a M + \log_a N$.
2. $\log_a(M/N) = \log_a M - \log_a N$.
3. $\log_a(1/N) = -\log_a N$
4. $\log_a(M^x) = x \log_a M$

Proof of 1. Put $A = \log_a M$ & $B = \log_a N$

$$A+B = \log_a M + \log_a N$$

$$a^{A+B} = a^{\log_a M + \log_a N} = a^{\log_a M} a^{\log_a N}$$
$$= MN$$

$A-B$

M/N

$$a^{A+B} = MN$$

$$\log_a(a^{A+B}) = \log_a(MN)$$

$$\log_a(a^{A-B}) = \log_a(M/N)$$

$A+B = \log_a(MN)$

$A+B = \log_a(MN)$

1. $\log_a(M) + \log_a(N) = \log_a(MN)$ //

2. $\log_a(M) - \log_a(N) = \log_a(M/N)$.

3. $\log_a(1/N) = \log_a 1 - \log_a N = -\log_a N$.

4. $\log_a(M^x) = x \log_a M$

$x \in \mathbb{N} := \{0,1,2,\ldots\}$

Partially proved

$$\log_a(M^x) = \log_a(\underbrace{M\ldots M}_{x \text{ times}}) = \log_a M + \ldots + \log_a M$$
$$= x \log_a M$$

$\boxed{\begin{array}{l} x \in \mathbb{Q} \\ x \in \mathbb{R} \end{array}}$ $\boxed{\log_a(M^\pi) = \pi \log_a M}$

Applications of Laws of logarithm

Simplify using logs.

$$\log_a \left[\frac{x^3 \sqrt{x^2+1}}{(x+3)^4}\right]$$

$$= \log_a\left(x^3 (x^2+1)^{1/2}\right) - \log_a\left[(x+3)^4\right]$$

$$= \log_a(x^3) + \log_a\left[(x^2+1)^{1/2}\right] - 4\log_a(x+3)$$

$$= 3\log_a(x) + \frac{1}{2}\log_a(x^2+1) - 4\log_a(x+3)$$

Warning: $\log_a(M+N) \neq \log_a M + \log_a N = \log_a(MN)$

$\log_a(M-N) \neq \log_a M - \log_a N = \log_a(M/N)$

Combine using logs

$$\underbrace{2\log_a x + \log_a 9}\ + \underbrace{\log_a(x^2+1) - \log_a 5}$$

$$= \log_a(9x^2) + \log_a\left(\frac{x^2+1}{5}\right)$$

$$= \log_a\left(\frac{9x^2(x^2+1)}{5}\right)$$

$$\boxed{\begin{aligned} \log_a(M+N) &= ? \\ \log_a(M-N) &= ? \end{aligned}}$$

X X

X X

Thm. Let $0<a<1$ or $a>1$ and $M, N>0$.

$$M=N \iff \log_a M = \log_a N$$

Two important values of a are 'e' & '10'

ln log

Natural Common

natural $\boxed{\ln}$ $\boxed{\log}$ common

Change of base Rule

old base new base

Thm. If $0<a<1$ or $a>1$ & $0<b<1$ or $b>1$.

Then, for $x>0$,

$$\boxed{\log_a x = \frac{\log_b x}{\log_b a}}$$

Proof. $M=\log_a x$ $N=\log_b x$ $R=\log_b a$

$\boxed{a^M = x}$ $\boxed{b^N = x}$ $\boxed{b^R = a}$

$(b^R)^M = x$

$b^{RM} = x$

$\log_b(b^{RM}) = \log_b x$

$$\log_a x = \frac{\log_b x}{\log_b a}$$

$$RM = \log_b x \Leftrightarrow (\log_b a)(\log_a x) = \log_b x$$

Examples. $\log_5 89 = \dfrac{\ln 89}{\ln 5} \approx 2.78$ ✓

$$\log_{\sqrt{2}} \sqrt{5} = \frac{\ln\sqrt{5}}{\ln\sqrt{2}} \approx 2.32$$ ✓

$\boxed{\ln\left(\frac{x}{a}\right) = ?}$ $\boxed{\dfrac{\ln x}{\ln a}}$ Warning

Graph: $\log_2 x$

common $\rightarrow \log_{10} x$

natural $\rightarrow \ln x$ ✓ ~~" $\log_e x$ "~~

$$f(x) = \log_2 x = \frac{\ln x}{\ln 2} = \left(\frac{1}{\ln 2}\right)\ln x$$

Prove that $\boxed{\dfrac{1}{\log_2 \pi} + \dfrac{1}{\log_6 \pi} > 2}$ ✓ ✓

$$LHS = \frac{1}{\log_2 \pi} + \frac{1}{\log_6 \pi} = \frac{\ln 2}{\ln \pi} + \frac{\ln 6}{\ln \pi}$$

$$= \frac{\ln 2 + \ln 6}{\ln \pi} = \frac{\ln(12)}{\ln \pi} \overset{?}{>} 2$$

$$\frac{\ln(12)}{\ln \pi} \overset{?}{>} 2$$

$$\ln(12) \overset{?}{>} 2(\ln\pi) = \ln(\pi^2)$$

$$\ln(12) \overset{?}{>} \ln(\pi^2)$$

Exponentiate with e $\quad e^{\ln(12)} \overset{?}{>} e^{\ln(\pi^2)}$

$$\boxed{a^{\log_a x} = x}$$

$$12 \overset{?}{>} \pi^2$$

3.141...

$\pi^2 < (3.15)^2$

$\pi^2 < 10 < 12$

Solve Logarithmic Equations

$2\log_{0.5} x = \log_{0.5} 4$ Solve for x.

$x = 2$

$$2\log_{0.5} x = \log_{0.5} 4$$

$$\log_{0.5} x^2 = \log_{0.5} 4$$

$$\boxed{a^{\log_a x} = x}$$

$$(0.5)^{\log_{0.5} x^2} = (0.5)^{\log_{0.5} 4}$$

$$x^2 = 4$$

$$x = \pm 2$$

$$\boxed{x = \cancel{-2}, 2}$$

$$\boxed{x = 2}$$

$\cancel{-3+1=-2}$ $\quad \cancel{-3-1=-4}$

$3+1=4$ $\quad +3-1=2$

Solve for x: $\log_8 \underset{3+1}{(x+1)} + \log_8 \underset{3-1}{(x-1)} = 1$

$$\underbrace{\log_8(x+1) + \log_8(x-1)} = 1$$

$$\log_8[(x+1)(x-1)] = 1$$

$$8^{\log_8[(x^2-1)]} = 8^1$$

$$x^2 - 1 = 8 \iff x^2 - 9 = 0$$

$x = \pm 3$ possible sol^n

$$\boxed{x = 3}$$

Solve for x.

$$\log_3 x + \log_4 x = 4$$

$$\log_3 x + \log_4 x = 4$$

$$\frac{\ln x}{\ln 3} + \frac{\ln x}{\ln 4} = 4$$

$$\ln x \left[\frac{1}{\ln 3} + \frac{1}{\ln 4}\right] = 4$$

$$\ln x = 4\left[\frac{1}{\frac{1}{\ln 3} + \frac{1}{\ln 4}}\right]$$

$$= 4\left[\frac{\ln 3 \cdot \ln 4}{\ln 3 + \ln 4}\right]$$

$$\ln x = 4\,\frac{\ln 3 \cdot \ln 4}{\ln(12)}$$

$$\boxed{x = e^{4\frac{\ln 3 \cdot \ln 4}{\ln 12}}}$$

Example. Solve for x.

$$\ln(x^2) = \boxed{(\ln x)^2}$$

$$\ln(x^2) = (\ln x)^2$$

$$2(\ln x) = (\ln x)^2 \qquad \boxed{\ln x = t}$$

$$2t = t^2$$

$$2t - t^2 = 0 \qquad t(2-t) = 0$$

$$\boxed{t=0} \text{ or } \boxed{t=2}$$

$$\ln x = 0 \text{ or } \ln x = 2$$

$$e^{\ln x} = e^0 \text{ or } e^{\ln x} = e^2$$

$$x = 1 \quad \text{or} \quad x = e^2$$

$$\boxed{x=1} \text{ or } \boxed{x=e^2}$$

The Natural Logarithmic Function

Saturday, 10 October 2020 10:28 AM

Defn. The natural logarithmic function is

$$f(x) = \log_e(x),$$

where the base is "e".

It is always denoted by $\ln(x)$ i.e.

$$f(x) = \boxed{\ln(x)}.$$

Remark.

$$\ln(e^x) = x, \quad \forall x \in \mathbb{R} = \mathrm{Dom}(e^x)$$

$$e^{(\ln x)} = x \quad \forall x \in (0,\infty) = \mathrm{Dom}(\ln x)$$

Common Logarithm

$$\boxed{\log x} = \boxed{\log_{10}(x)}$$

$$\boxed{\ln x = \log_e x}$$

$\boxed{\ln}$ $\boxed{\log}$

Solving Exponential Equations

Saturday, 10 October 2020 10:36 AM

$\log_a(a^x) = x$

Example 1. Solve for x. $\boxed{2^{x+1} = 64}$

$64 = 16 \times 4 = 2^4 \times 2^2 = 2^6$

$2^{x+1} = 2^6 \Leftrightarrow 2^x = 2^5$

$\Leftrightarrow \log_2 2^x = \log_2 2^5$

$\Leftrightarrow \boxed{x = 5}$

Example 2. Solve $e^{-x^2} = (e^x)^2 \frac{1}{e^3}$

ln

$e^{-x^2} = e^{2x-3}$

$-x^2 = 2x - 3$

$0 = x^2 + 2x - 3 \Leftrightarrow 0 = (x+3)(x-1)$

$\boxed{x = -3, 1}$

Example 3. Solve $9^x - 2\,3^{x+1} - 27 = 0$

$(3^2)^x \quad 6\,3^x$

$(3^x)^2$

$9^x - 2\,3^{x+1} - 27 = 0$

$(3^x)^2 - 6(3^x) - 27 = 0$

$t^2 - 6t - 27 = 0$

$t^2 - 9t + 3t - 27 = 0 \Leftrightarrow (t-9)(t+3) = 0$

$(3^x - 9)(3^x + 3) = 0$

$\boxed{3^x = 9}$ or $\boxed{\cancel{3^x = -3}}$

$\boxed{x = 2}$

Example 4. Solve $5^{x-2} = 3^{3x+2}$

$\ln(5^{x-2}) = \ln(3^{3x+2})$

$$(x-2)\ln(5) = (3x+2)\ln(3)$$

$$-2(\ln(5)+\ln(3)) = 3x(\ln(3)) - x\ln(5)$$

$$-2(\ln(15)) = x[3\ln(3) - \ln(5)]$$

$$x = \frac{-2\ln(15)}{\ln(27)-\ln(5)} = \left[\frac{\ln(1/225)}{\ln(27/5)}\right]$$

Example 5. Solve $x + e^x = 2$

$$x + e^x = 2$$

$$e^x = 2 - x$$

$$\boxed{x = \ln(2-x)}$$

$$\ln(2-x) - x = 0$$

$$\boxed{x \approx 0.443}$$

Outline

Sunday, 9 August 2020 9:51 AM

Exponential Functions

- One-to-One Functions.
- Exponential Function
- The Natural Exponential Function.

One-to-One Functions

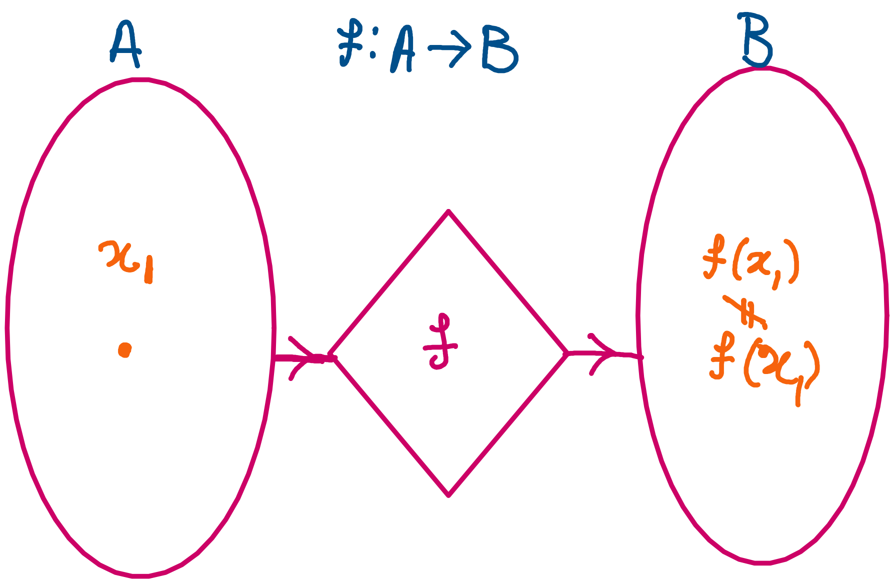
10 March 2020 08:28

$y = f(x)$ ✓

$f : A \to B$

$A, B \subseteq \mathbb{R}$

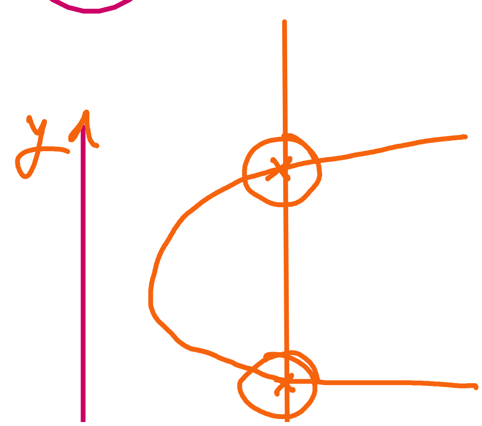
	Domain (A)	Codomain (B)	
✓1	One x	More than one $f(x)$	? Not a function
✓2	More than one x	One $f(x)$	? It is a function but it is not reversible
✓3	One x	One $f(x)$	? It is a function It is reversible

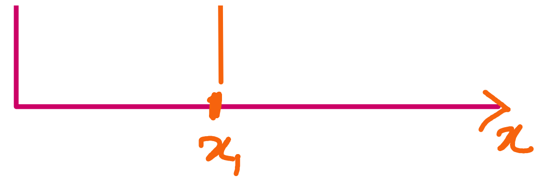


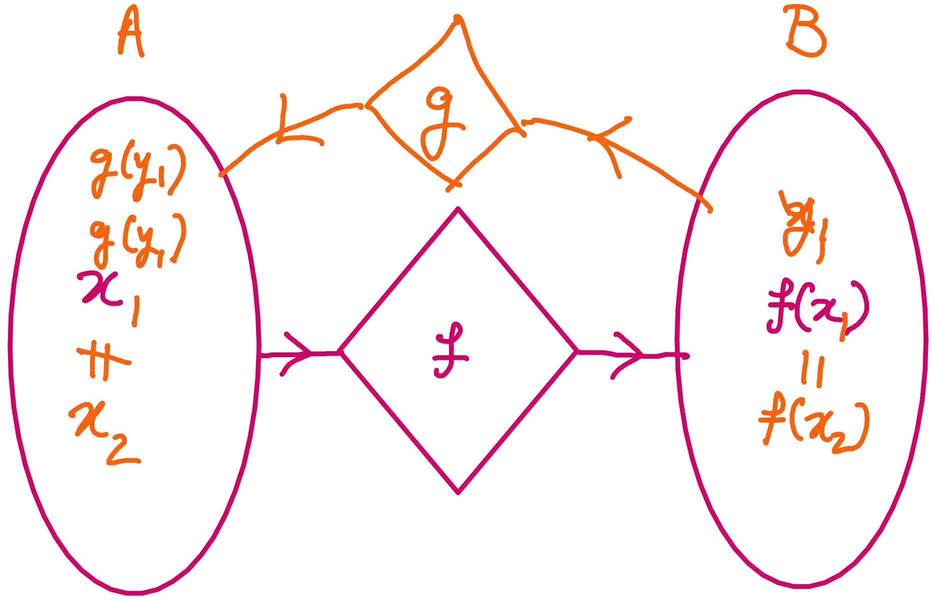
Vertical Line Test

$x = \text{const.}$

Vertical line test fails





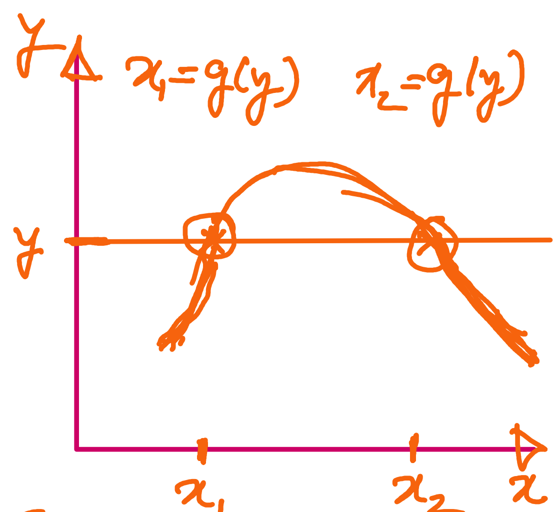


Horizontal Line Test

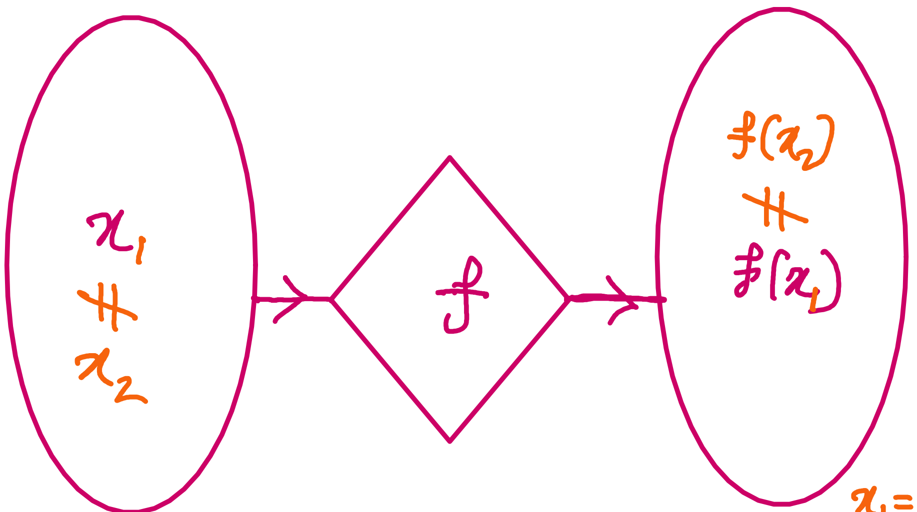
$$y = f(x) \not\Rightarrow x = f(y)$$

$\boxed{\text{Not 'Reversible'}}$ (?)

Horizontal line test fails. ✓

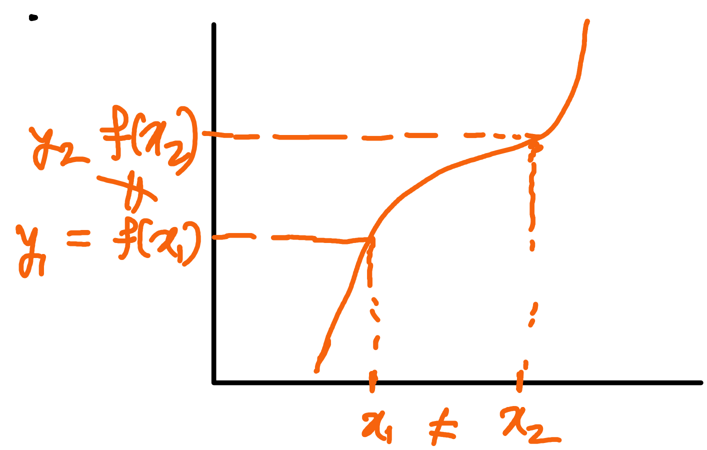


· Observe f is NOT "Reversible"



$\boxed{\text{"Reversible"}}$

$x_1 = g(y_1)$
$x_2 = g(y_2)$



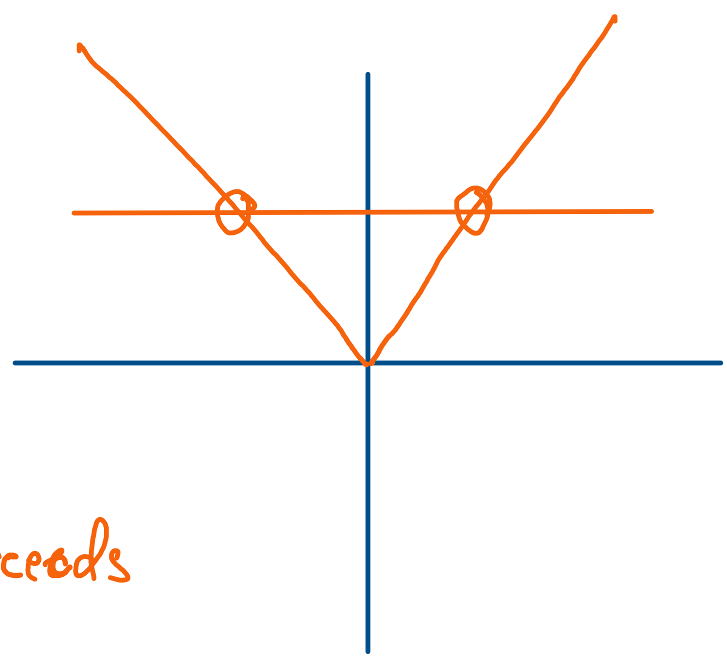
Definition (One-to-One Function)

A function $f : A \to B$ is called one-to-one if, for any $x_1 \neq x_2 \in A$, then $f(x_1) \neq f(x_2)$.

$$\boxed{\begin{array}{l} f(x_1) = f(x_2) \\ \Rightarrow x_1 = x_2 \end{array}}$$

Example.

$$f(x) = |x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$$



Vertical line test succeeds

$2, -2$

$f(2) = 2 = f(-2)$

$\boxed{\text{NOT one-to-one}}$

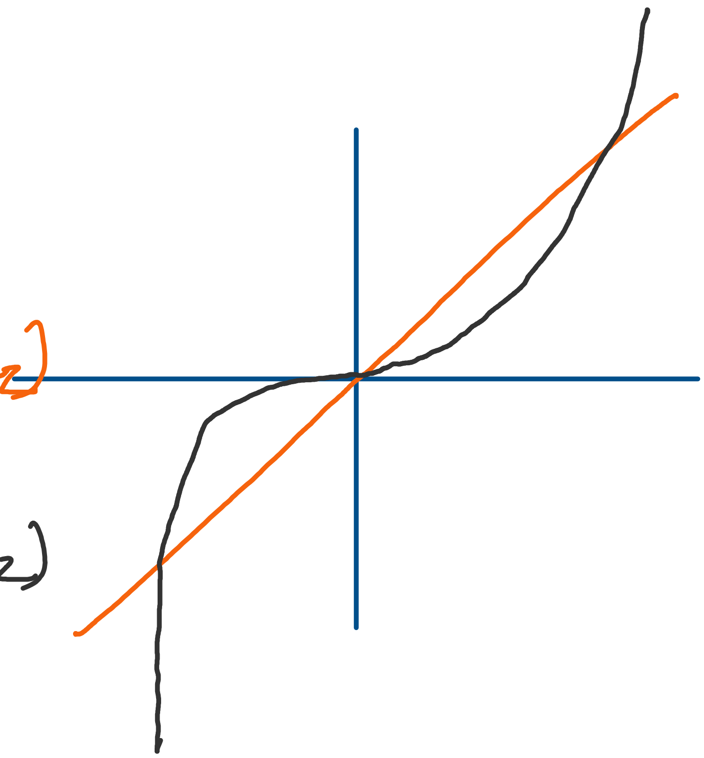
Example

$f(x) = x$

$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$

$f(x) = x^3$

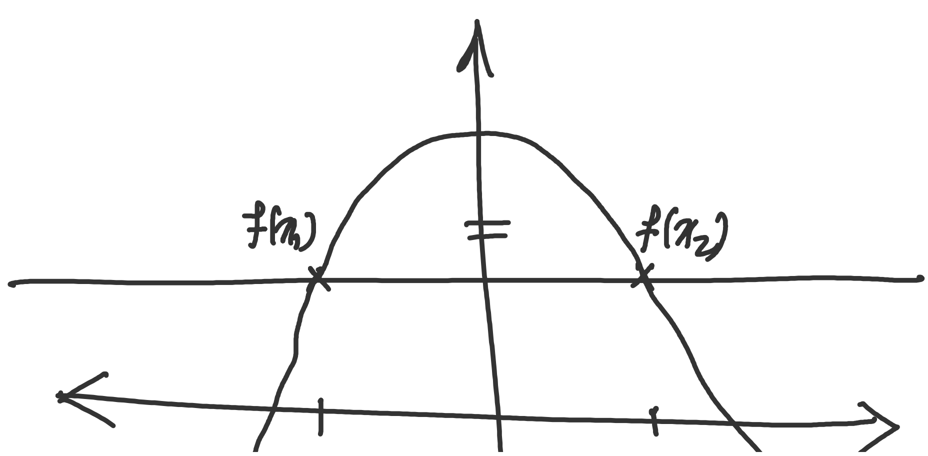
$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$



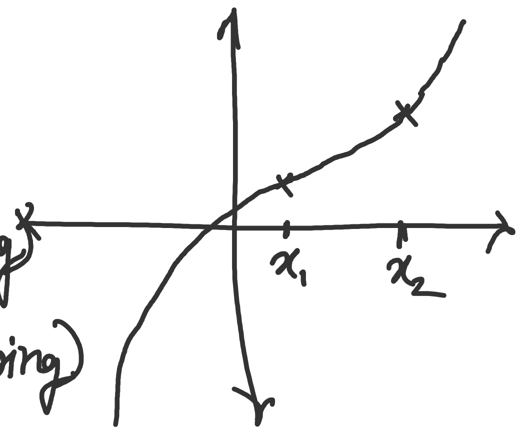
Theorem. (The Horizontal Line Test)

If any horizontal line intersects the graph of a function f in at most one point, then f is one-to-one.

Proof.



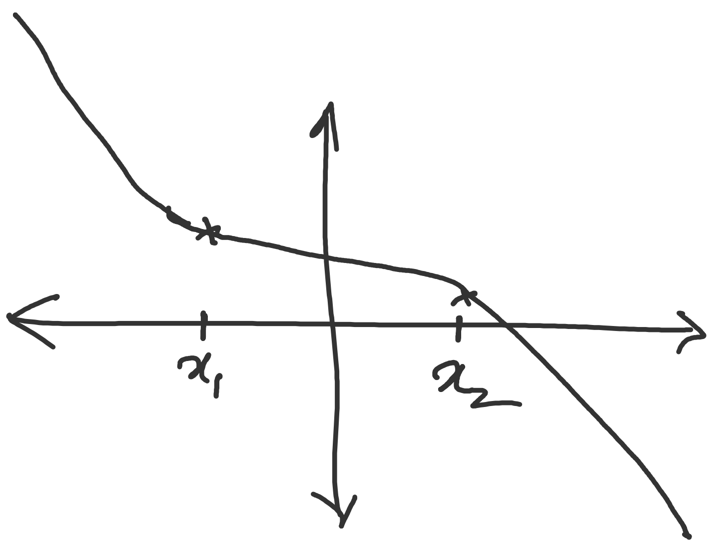
$\boxed{Q.}$ Can we identify the class of functions that are one-to-one?



For every $x_1, x_2 \in A$,

$x_1 \lneq x_2 \Rightarrow f(x_1) \lneq f(x_2)$ (increasing)

$x_1 < x_2 \Rightarrow \underline{f(x_1) > f(x_2)}$ (decreasing)



Theorem.

$\boxed{\text{If}}$ f is an increasing or decreasing function then f is one-to-one.

Exponential function.

12 March 2020 10:58

Exponent

Recall a^{r} exponent 'base'

$a > 0$, $r \in \mathbb{Q}$

What if $r \in \mathbb{R}\setminus\mathbb{Q}$? Irrational?

Why $a > 0$?

$a = -1$

$a^{1/2} = (-1)^{1/2} = \boxed{i \in \mathbb{C}}$

a^{r}, $r \in \mathbb{R}\setminus\mathbb{Q}$

Q. Can we define a^{x} $(a>0)$ for $x \in \mathbb{R}\setminus\mathbb{Q}$?

Eg. $2^{\sqrt{2}}$, 5^{π}

$\sqrt{2} = 1.41\ldots$

2^{1}

$2^{1.4}$

$2^{1.41}$

$\vdots$

a^{x} is defined for $\boxed{x \in \mathbb{R}}$

$\pi = 3.14159\,2635\ldots$ (Non-repeating)

$5^{\pi} = ?$

5^{3}

$5^{3.1}$

$5^{3.14}$

$5^{3.141\ldots}$

$\vdots$

$\downarrow$ seq^n

Existence of 5^{π} is assured. 5^{π}

Laws of Exponents.

For $s, t \in \mathbb{R}$ and $a, b > 0$,

(i) $a^s . a^t = a^{s+t}$

(ii) $(a^s)^t = a^{st}$

(iii) $(ab)^s = a^s b^s$

Recall. $1^s = 1$, $a^{-s} = \frac{1}{a^s} = \left(\frac{1}{a}\right)^s$ and $a^0 = 1$, $a > 0$

0^0 is undefined

Definition.

An exponential function in standard form is given by $f(x) = a^x$, where $a > 0$, $\boxed{a \neq 1}$

Observations.

(i) Domain of f is $\mathbb{R}$

(ii) $a \neq 1$? $f(x) = 1^x = 1$ (constant)

Exercise.

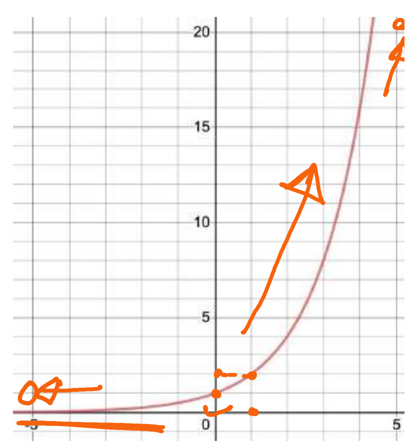
Graph the following functions (Graphing tool)

1. (a) 2^x (b) 3^x (c) 5^x together

2. (a) $(1/2)^x$ (b) $(1/3)^x$ (c) $(1/5)^x$. together

Identify properties of the graphs.

1 (a) $f(x) = 2^x$



Domain of f = $\mathbb{R}$

$2 > 1$

$\boxed{2^x > 2^0 = 1 \quad 0 < 2^x < 1}$ $x > 0$

Range of f = $(0, \infty)$

y-intercept = $(0, 1)$

x-intercept = Nil

$\boxed{y = 0 \quad \text{Horizontal Asymptote.}}$

End-behavior $x \to \infty$ $2^x \to \infty$

$x \to -\infty$ $2^x \to 0$

Roots $\boxed{\text{No roots}}$

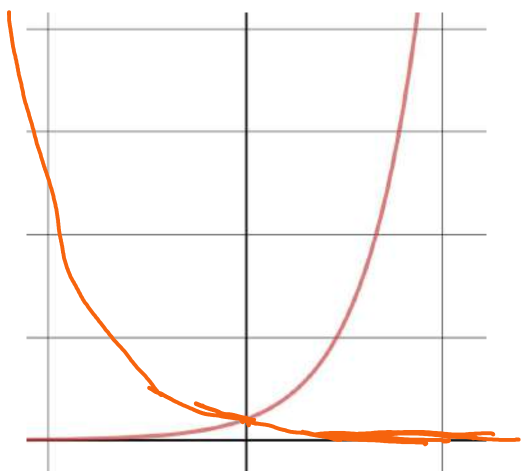
$x_1 \neq x_2$ $x_1 < x_2$

increase/decrease $\quad 2 \quad$ increasing $\boxed{2^{x_1} < 2^{x_2}}$

Fact.

Every $f(x) = a^x$, $a > 1$ has same properties as 2^x.

Graph of $f(x) = a^x$, $a > 1$.



$0 < a < 1 \qquad g(x) = a^x$

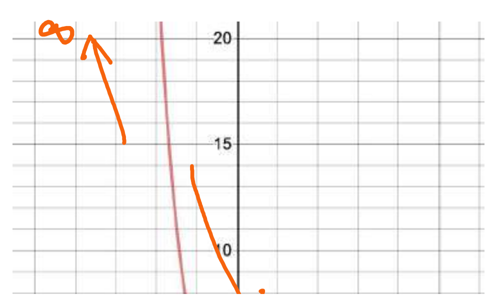
2(c) $\quad g(x) = (1/5)^x = 5^{-x}$

Compare with $\boxed{5^x}$

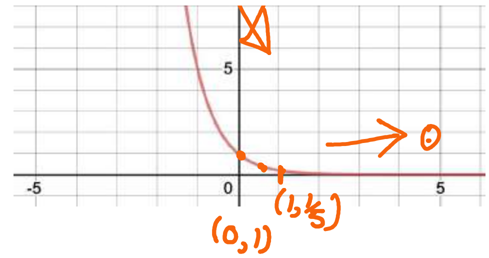
Domain $= \mathbb{R}$

Range $= (0, \infty)$

y-intercept $= (0, 1)$



y-intercept = (0, 1)



x-intercept = Nil

Roots   No roots

End-behavior

$x \to \infty$ $\left(\frac{1}{5}\right)^x \to 0$

$x \to -\infty$ $\left(\frac{1}{5}\right)^x \to \infty$

Increase / decrease   Decreasing f^n

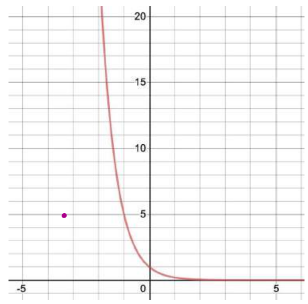
Fact.

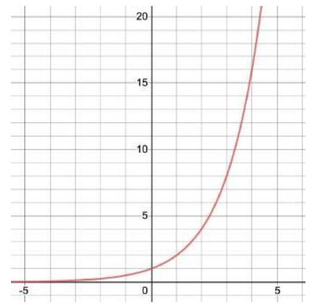
Every $f(x) = a^x$, $0<a<1$ has same properties as $\left(\frac{1}{5}\right)^x$.

Summary

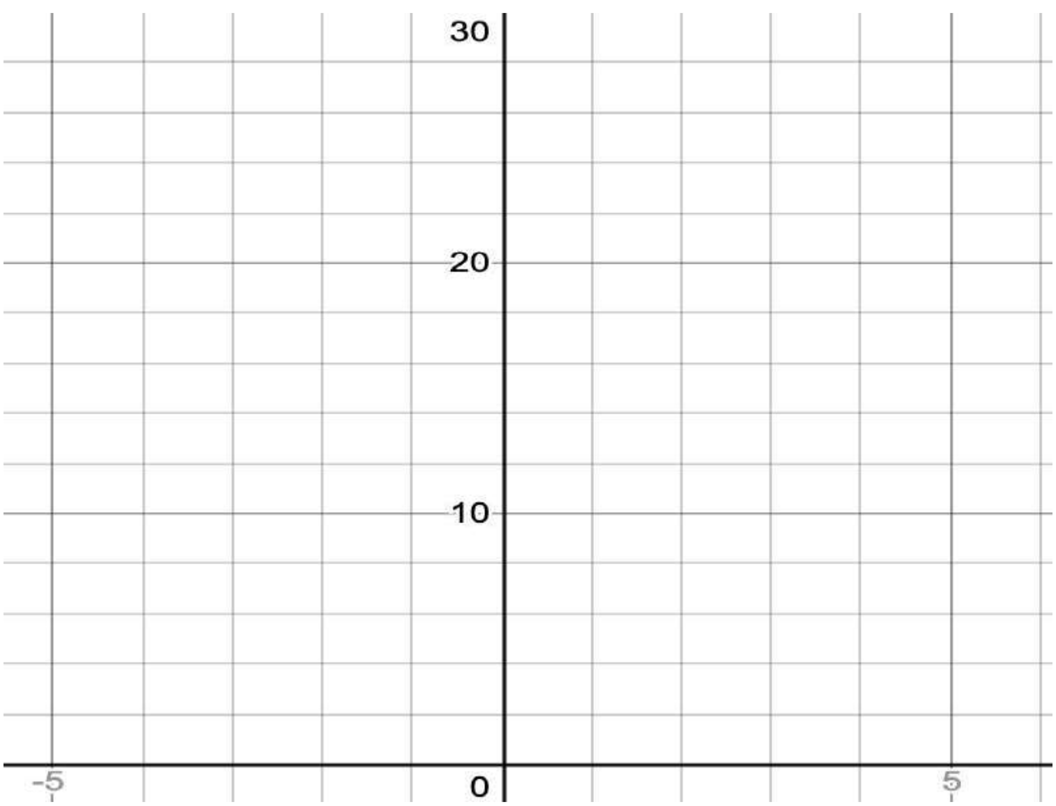
$f(x) = a^x$	$0<a<1$	$a>1$
Domain	$\mathbb{R}$	$\mathbb{R}$
Range	$(0, \infty)$	$(0, \infty)$
x - intercept	Nil	Nil
y - intercept	(0, 1)	(0, 1)
Horizontal Asymptote	$y=0$	$y=0$
Increase / decrease	decreasing	increasing
End behavior		
$x \to \infty$	$f(x) \to 0$.	$f(x) \to \infty$.
$x \to -\infty$	$f(x) \to \infty$.	$f(x) \to 0$.

Graphs





Example. Graph $f(x) = 3^{-x} + 2$.



The Natural Exponential Function.

Sunday, 9 August 2020 11:41 AM

From the theory of limits, it is known that

$$\left(1+\frac{1}{n}\right)^n \to e \quad \text{as } n \to \infty$$

Existence of 'e' is studied in calculus.

e is irrational number.

$e \approx \underline{2.718}28\ldots$

Q Why is 'e' so important?

Interest Rate Calculation

Cont^s Compounding

n	$\left(1+\frac{1}{n}\right)^n$
1	2
10	2.5937
100	2.7048
1000	2.7169
10,000	2.7181
100,000	2.7182

e

Euler's number

Re 1

$$\left(1+\frac{x}{n}\right)^{nt}$$

$$\boxed{e^{xt}}$$

1%

$$\left(1+\frac{1}{100}\right)^n$$

$$\left(1+\frac{0.01}{4}\right)^4$$

$$\left(1+\frac{0.01}{12}\right)^{12}$$

$$\left(1+\frac{0.01}{n}\right)^n \quad n\to\infty$$

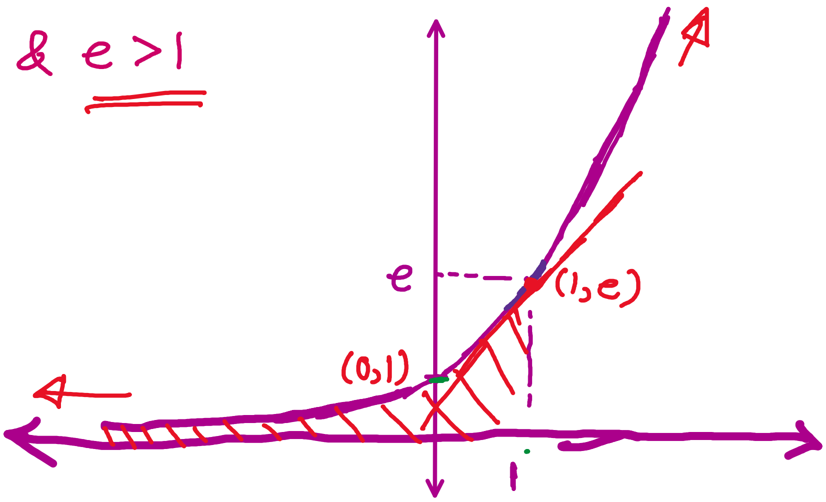
$$\boxed{1\times e^{0.01t}}$$

Definition.

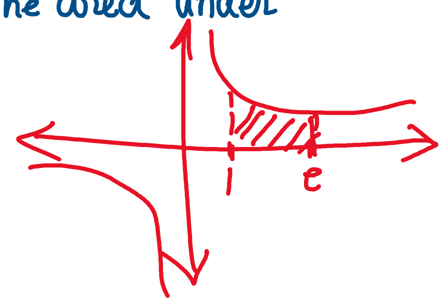
The natural exponential function is defined as $f(x) = e^{x}$.

Properties. Domain of $f = \mathbb{R}$, Range $= (0, \infty)$

& $e > 1$



- e is the slope of the tangent line to $f(x) = e^{x}$ at $(1, e)$.
- The area under the $f(x) = e^{x}$ from $(-\infty, 1)$ is e.
- For $f(x) = 1/x$, $x \in (1, e)$, the area under the curve is 1.



Example.

Let R be the percent of people who respond to affiliate links under YouTube descriptions &

purchase the product in t minutes is given by

$$\boxed{R(t) = 50 - 100e^{-0.2t}}$$

ⓐ What is the percentage of people responding after 10 minutes?

ⓑ What is the highest percent expected? 50%

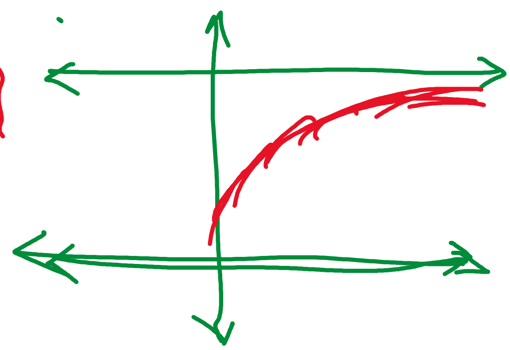
ⓒ How long before $R(t)$ exceeds 30%?

ⓐ $R(10) = 50 - 100e^{-0.2\times 10}$

$= 50 - 100e^{-2} = 36.46$

ⓑ $R(t) = 50 - \cancel{100e^{-0.2t}}$

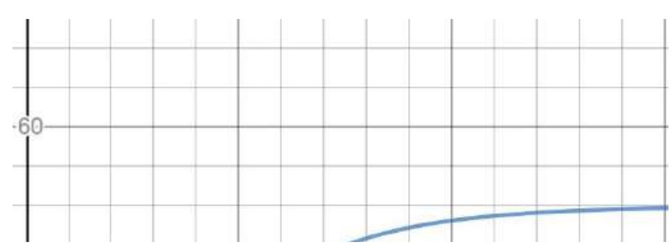
50%



ⓒ $R(t) = 50 - 100e^{-0.2t}$

• $R(t) = 30$

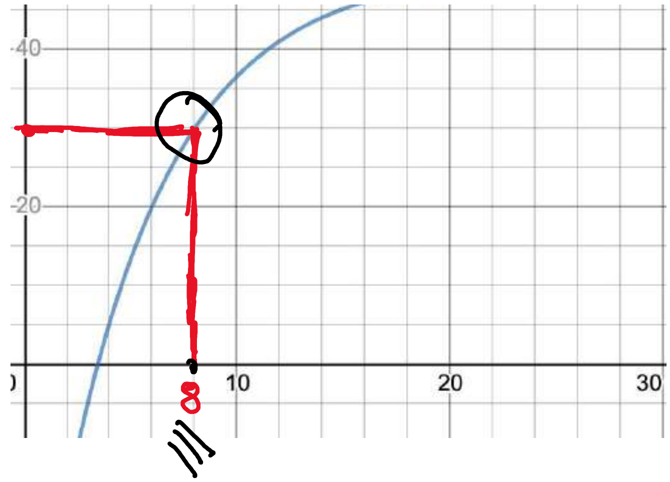
$30 = 50 - 100e^{-0.2t}$



$$20 = 100e^{-0.2t}$$

$$\frac{1}{5} = e^{-0.2t} \qquad \boxed{\text{STOP}}$$

$\to$ $t \approx 8$ minutes



Composite Functions

Monday, 9 November 2020 10:06 AM

Computer store

85% of price

3000 off ✓

→ 85% of price − 3000 ✓

Let x denote the item price (MRP)

$$\boxed{f(x) = 0.85x}$$

$$g(x) = x - 3000$$

★ $h(x) = 0.85x - 3000$

$h(x) = f(x) - 3000$ ✓

$= g(f(x))$

$g(\square) = \square - 3000$

$$\boxed{h(x) = (g \circ f)(x)}$$

$x = 14\,000$ Evaluation of $(g\circ f)-n$

$$(g\circ f)(x) = g(f(x))$$

$$= f(x) - 3000$$

$$= 0.85x - 3000$$

$$g(f(14\,000))$$

$$= 0.85 \times 14000 - 3000$$

$$= 11900 - 3000$$

$$= 8{,}900$$

The composition of Functions

The composition of the functions f & g is denoted $f\circ g$ & is defined by

$$(f\circ g)(x) = \boxed{f(g(x))}$$

The domain of the composite function $f\circ g$ is the set of all x such that

- (i) x is in the domain of g.
- (ii) $g(x)$ is in the domain of f.

$$(f \circ g)(x) = f(g(x))$$

$$x \xrightarrow{\;g\;} g(x) \xrightarrow{\;f\;} f(g(x))$$

Dom(g) Dom(f) Range(f)

$$x \to \boxed{g} \to g(x) \to \boxed{f} \to f(g(x))$$

Example. Given $f(x) = 3x - 4$, $g(x) = x^2$,

find (a) $(g \circ f)(x)$ (b) $(f \circ g)(x)$.

Sol^n.

$$(g \circ f)(x) = g(f(x))$$
$$= (f(x))^2$$
$$= (3x-4)^2$$

$g(\square) = \square^2$

$$(g \circ f)(x) = g(f(x)) = g(3x-4)$$
$$= (3x-4)^2$$

Replace $f(x) = 3x-4$

$$(f \circ g)(x) = f(g(x))$$
$$= 3g(x) - 4$$
$$= 3x^2 - 4$$

$f(\Delta) = 3\Delta - 4$

Exercise. $f(x) = x+1$ $\quad g(x) = x^2 - 1$

find $(g \circ f)(x)$ $\quad (f \circ g)(x)$.

Determination of the domain for composite f^n.

$$(f \circ g)(x) = f(g(x))$$

The following values must be excluded from input x.

- $x \notin \mathrm{Dom}(g) \Rightarrow x \notin \mathrm{Dom}(f \circ g)$
- $\{x \mid g(x) \notin \mathrm{Dom}(f)\}$ must not be included in $\mathrm{Dom}(f \circ g)$.

Example. $f(x) = \boxed{\dfrac{2}{x-1}}$ $\quad g(x) = \boxed{\dfrac{3}{x}}$

$(f \circ g)(x)$ & $\mathrm{Dom}(f \circ g)$

$$\boxed{(f \circ g)(x)} = f(g(x)) \qquad f(\square) = \frac{2}{\square - 1}$$

$$= \frac{2}{g(x) - 1}$$

$$= \frac{2}{3/x - 1} = \boxed{\frac{2x}{3-x}}$$

- $(f \circ g)(x) = \boxed{\frac{2x}{3-x}} = \frac{0}{3} = 0$

Rule 1. $x \notin \mathrm{Dom}(g) \Rightarrow x \notin \mathrm{Dom}(f \circ g)$

$g(x) = \frac{3}{x}$, $x \neq 0$

$x = 0 \notin \mathrm{Dom}(g) \Rightarrow x = 0 \notin \mathrm{Dom}(f \circ g)$

Rule 2. $g(x) \notin \mathrm{Dom}(f)$

$f(x) = \frac{2}{x-1}$ $\boxed{x \neq 1}$

$\mathrm{Dom}(f \circ g) = \{ x \mid x \neq 0 \overset{\text{and}}{,} x \neq 3 \}$

Exercise. $f(x) = \frac{1}{x+1}$ $g(x) = \frac{1}{x}$

$(f \circ g)(x)$ and $\mathrm{Dom}(f \circ g)$

Inverse Functions

Monday, 14 September 2020 9:30 AM

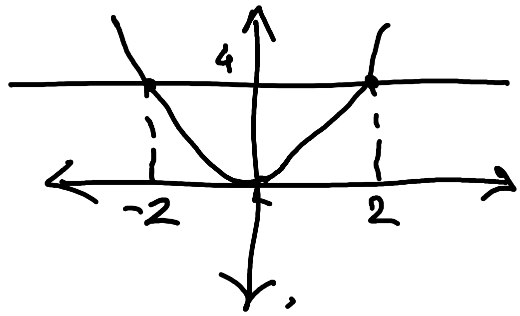
$f : \mathbb{R} \to \mathbb{R}$

$f(x) = x^2$

★ Not 'Reversible'

$f(-2) = f(2) = 4$

Reversible?



★ We now look at one-to-one functions

$\mathbb{R} \to \mathbb{R}$

$\boxed{g(x) = 4x}$

$\boxed{h(x) = \frac{x}{4}}$ $\mathbb{R} \to \mathbb{R}$

$y = 4x$

$\frac{y}{4} = x$

$y = \frac{x}{4}$

$\boxed{g(x) = 4x}$ $4y = x$

$I(x) = g \circ h(x) = g(\underbrace{h(x)}) = 4\,h(x)$
$= 4\,\frac{x}{4} = x$

$I(x) = h \circ g(x) = h(g(x)) = \frac{g(x)}{4} = \frac{4x}{4} = x$

$$\boxed{g \circ h(x) = I(x) = h \circ g(x)}$$

Def^n. The Inverse of a function f,

f^{-1} is a function such that

$$f^{-1} \circ f(x) = f^{-1}(f(x)) = x \quad \forall x \in \mathrm{Dom}(f) = \mathrm{Range}(f^{-1})$$

$$\& \quad f \circ f^{-1}(x) = f(f^{-1}(x)) = x \quad \forall x \in \mathrm{Dom}(f^{-1}) = \mathrm{Range}(f)$$

$$f : \mathbb{R} \to [0,\infty) \qquad f^{-1} : [0,\infty) \to \mathbb{R}$$

Remark. f is one-to-one function

$\Rightarrow$ f^{-1} exists for f.

Warning: $f^{-1} \neq \frac{1}{f}$

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

Example. $g(x) = x^3$, $\mathbb{R} \to \mathbb{R}$ & $g^{-1}(x) = \sqrt[3]{x} = x^{1/3}$, $\mathbb{R} \to \mathbb{R}$

Verify

$$g^{-1}(g(x)) = g^{-1}(x^3) = (x^3)^{1/3} = x.$$

$$g(g^{-1}(x)) = g(x^{1/3}) = (x^{1/3})^3 = x.$$

Example 2. Verify f is the inverse of g

$$f(x) = \frac{x-5}{2x+3} \quad \& \quad g(x) = \frac{3x+5}{1-2x}$$

$$f(g(x)) = \frac{g(x)-5}{2g(x)+3} = \frac{\frac{3x+5}{1-2x} - 5}{2\,\frac{3x+5}{1-2x} + 3}$$

$$= \frac{3x+\cancel{5} - 5(\cancel{1}-2x)}{2(3\cancel{x}+5)+3(1-2\cancel{x})} = \frac{13x}{13} = x.$$

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

$$= \frac{3(x-\cancel{5})+5(2x+\cancel{3})}{\cancel{2x}+3-2(\cancel{x}-5)} = \frac{13x}{13} = x.$$

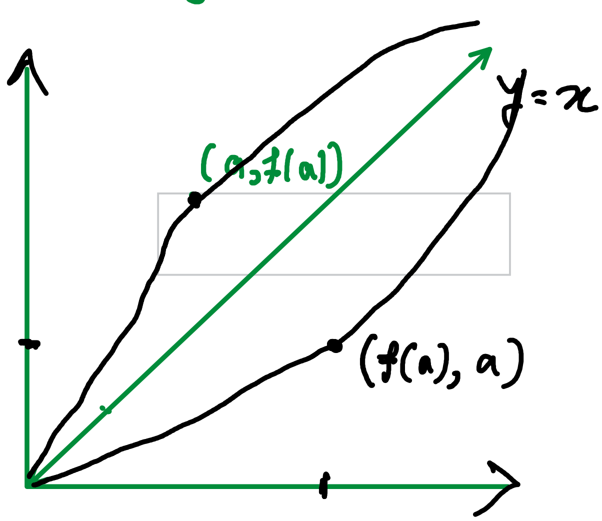
Graph f & f^{-1}

$$x = a \xrightarrow{\;f\;} f(a) = y$$

$$a = f^{-1}(f(a)) \xleftarrow{\;f^{-1}\;} f(a)$$

If $\underline{(a, f(a))}$ is on the graph of f

then $\underline{(f(a), a)}$ is on the graph of f^{-1}



Theorem. The graphs of f & f^{-1} are symmetric across $\underline{y=x}$ line

Some topics from Maths 1

- Straight lines
- Quadratic equations
- Polynomials
- Functions
- Exponentials and logarithms

What is a function of one variable?

$f : D \longrightarrow \mathbb{R}$ where $D \subseteq \mathbb{R}$. $f(x)$ where $x \in \mathbb{R}$.

A function is defined to be a relation from a set of inputs to a set of possible outputs where each input is related to exactly one output.

This means that if the object x is in the set of inputs (called the domain) then a function f will map the object x to exactly one object $f(x)$ in the set of possible outputs (called the codomain).

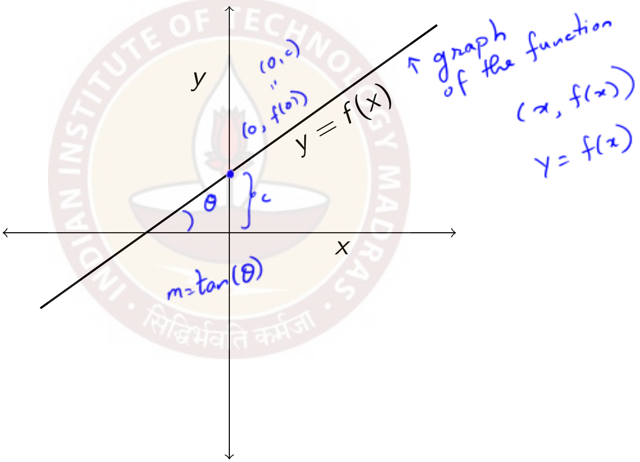
$$f : X \to Y$$

Domain : X Codomain : Y Range : $\{f(x) \mid x \in X\}$

Linear functions

slope

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = mx + c \qquad m, c \in \mathbb{R}.$$

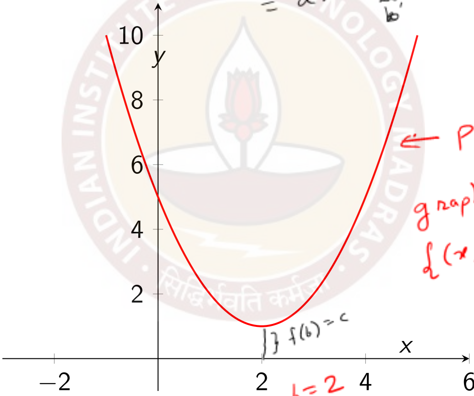


Quadratic functions

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = \underline{\overset{>0}{a}(x-b)^2} + c \qquad a, b, c \in \mathbb{R}.$$

$= a(x^2 - 2bx + b^2) + c$

$= ax^2 \underbrace{-2ab}_{b'} x + \underbrace{ab^2 + c}_{c'} = ax^2 + b'x + c'.$



← parabola

graph of f(x)

$\{(x, f(x)) \mid x \in \mathbb{R}\}$

Polynomial functions

$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \quad ; \quad a_i \in \mathbb{R}.$

Example : Graph of $f(x) = x^3 - 4x$

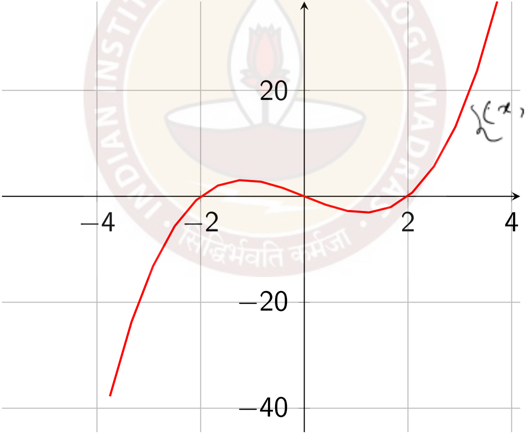
$f(x)=0$

$x^3-4x=0$

$x(x^2-4)=0$

$x(x-2)(x+2)=0$

$f(x)=0$ when exactly $x=-2,0,2$.



$\{(x, f(x)) \mid x \in \mathbb{R}\}$

$= \{(x, x^3-4x) \mid x \in \mathbb{R}\}$

The exponential and logarithmic functions

$g,\ f : \mathbb{R} \to \mathbb{R}$ $\boxed{f(x) = a^{x}}$ $g(x) = log_a(x).$

$2^2 = 4.$

$(0.5)^2 = 0.25$

$\pi^2 = \pi \times \pi.$

$2^{-2} = \frac{1}{2^2} = \frac{1}{4}$

$(0.5)^{-2} = \frac{1}{(0.5)^2} = \frac{1}{1/4} = 4.$

$\pi^{-2} = \frac{1}{\pi^2}.$

$2^{\sqrt{2}}$

Take a rational no. q which approximates $\sqrt{2}$ as closely as we desire.

2^q

2^{π}

$2^{\sqrt{2}}$

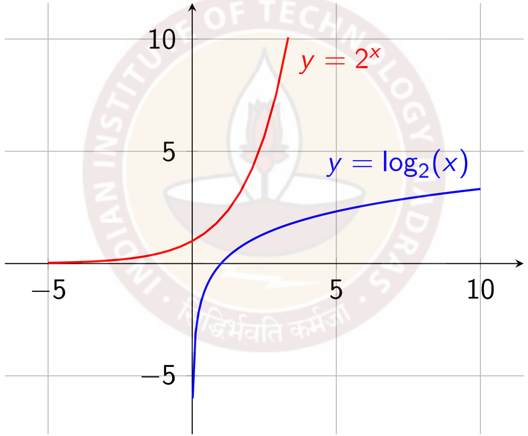
$2^{1/3} = \sqrt[3]{2}$

$(0.5)^{5/6}$

The exponential and logarithmic functions

$f : \mathbb{R} \to \mathbb{R}$ $\qquad f(x) = a^x$ $\qquad g(x) = log_a(x)$.

$a > 0$.

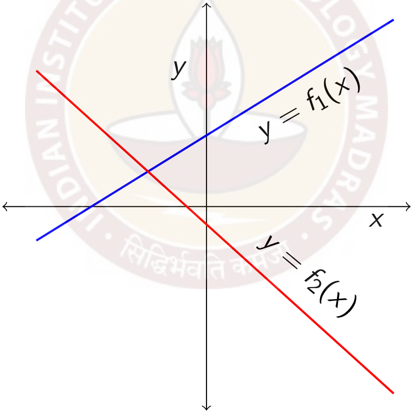


$\log_2(2^x) = x$.

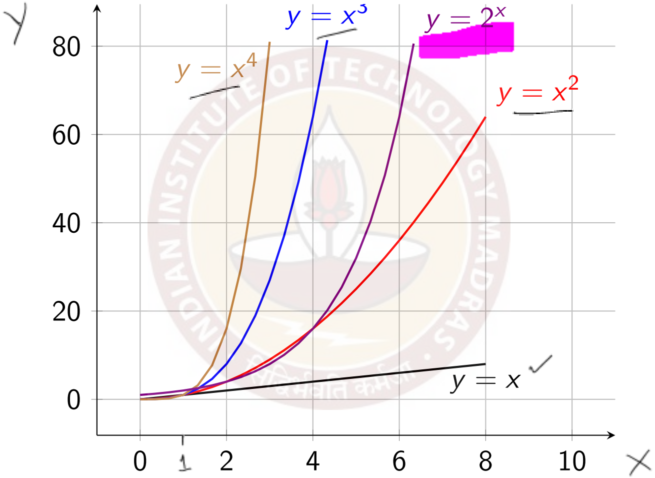
Monotonicity of functions

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be monotone increasing if $x_1 \leq x_2$ implies $f(x_1) \leq f(x_2)$.

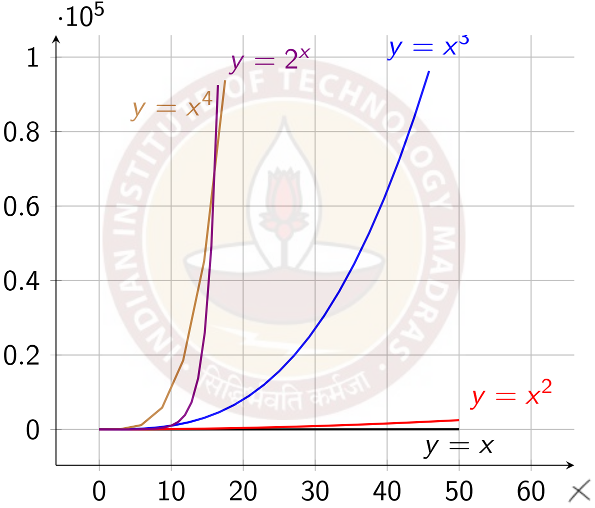
A function $f : \mathbb{R} \to \mathbb{R}$ is said to be monotone decreasing if $x_1 \leq x_2$ implies $f(x_1) \geq f(x_2)$.



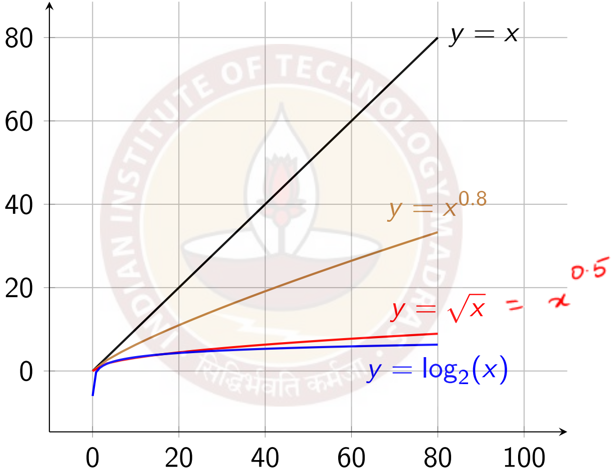
Comparing various functions : fast growth, close to 0



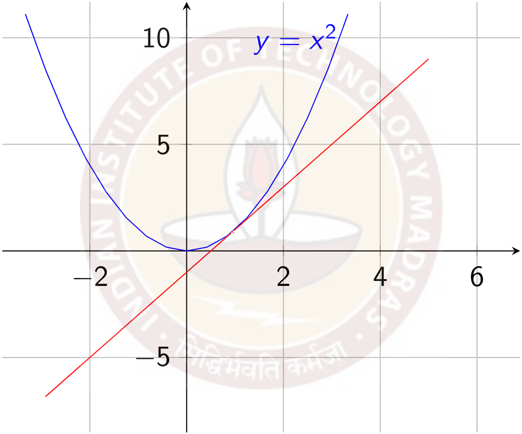
Comparing various functions : fast growth



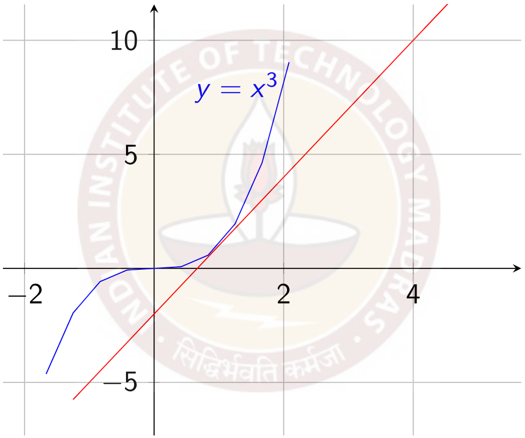
Comparing various functions : slow growth



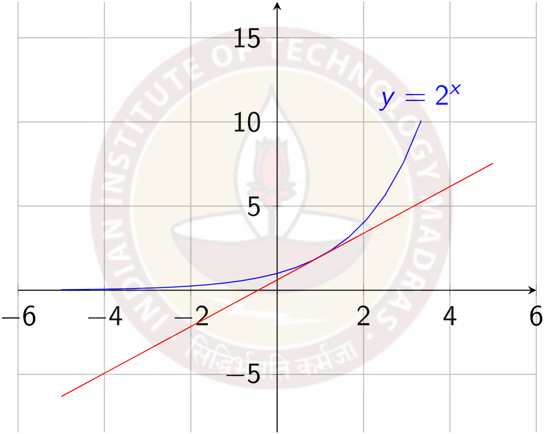
Tangent lines : Example 1



Tangent lines : Example 2



Tangent lines : Example 3

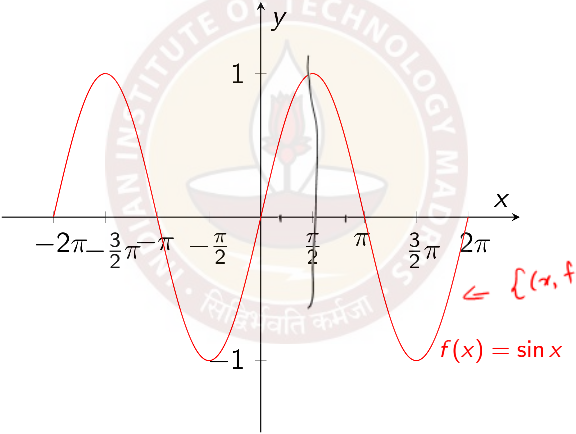


Trigonometric functions : the sine function

$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = \sin x$

$0 \le x \le \pi/2$

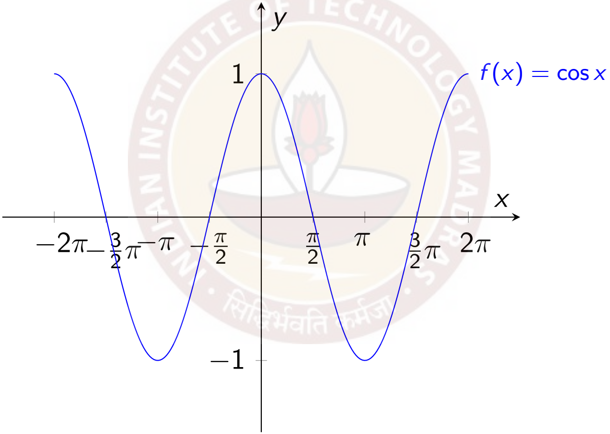
$\frac{|opp|}{|hyp|} = \sin(x)$



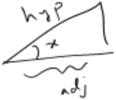
$\leftarrow \{(x, f(x)) \mid x \in \mathbb{R}\}$

Trigonometric functions : the cosine function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = cos\ x$$

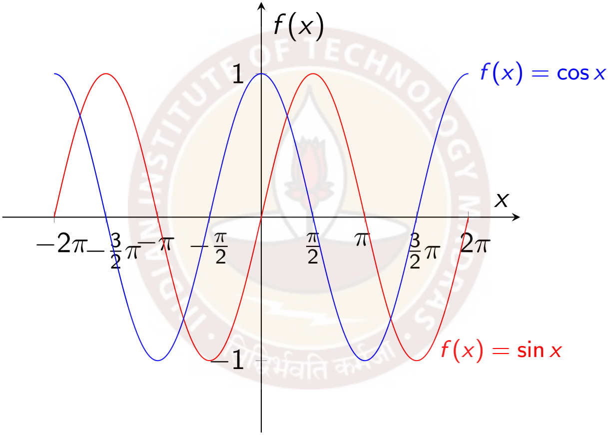


$0 \leq x \leq \pi/2$



$\cos(x) = \frac{|adj|}{|hyp|}$.

Comparing the *sin* and *cos* functions

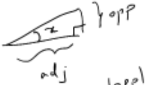


Trigonometric functions : the tangent function

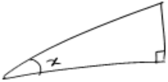
$\mathbb{R} \setminus \{ n\frac{\pi}{2} \mid n$ runs over odd integers$\}$.

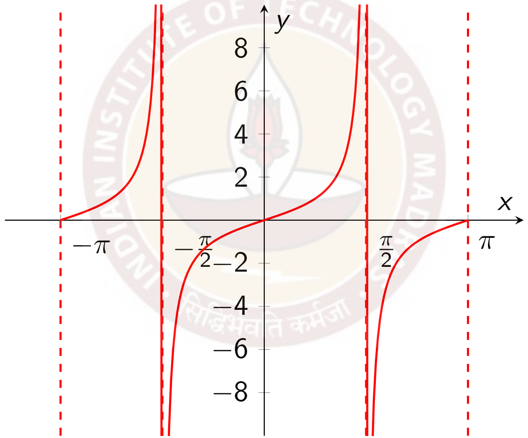
$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = tan\ x$$

$0 \le x \le \pi/2$



$\tan(x) = \dfrac{|opp|}{|adj|}$





Trigonometric functions : other functions and identities

Other trigonometric functions :

- the cotangent function $cot(x)$
- the secant function $sec(x)$
- the cosecant function $cosec(x)$

Important identities :

- $sin(-x) = -sin(x)$
- $cos(-x) = cos(x)$
- $tan(-x) = -tan(x)$
- $sin^2(x) + cos^2(x) = 1$
- $\frac{sin(x)}{cos(x)} = tan(x)$
- $\frac{1}{cos(x)} = sec(x)$
- $\frac{1}{sin(x)} = cosec(x)$

Arithmetic operations on functions

Let $D \subset \mathbb{R}$ and $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$ be functions on D.

i) The sum function $f + g$ is defined on D by
$(f + g)(x) = f(x) + g(x)$, $x \in D$.

ii) The product function fg is defined on D by
$fg(x) = f(x) \times g(x)$, $x \in D$.
$f(x)g(x)$

iii) Let $c \in \mathbb{R}$. The function cf is defined on D by
$(cf)(x) = c \times f(x)$, $x \in D$.
$c f(x)$

$h(x) = \frac{x}{x^2+1}$

$\mathbb{R} \to \mathbb{R}$

iv) If $g(x) \neq 0$, $x \in D$, the quotient f/g is defined on D by
$(f/g)(x) = f(x)/g(x)$, $x \in D$.

Functions obtained by composition

Let $D \subset \mathbb{R}$ and $f : D \to \mathbb{R}$ be a function.

Let $g : E \to \mathbb{R}$ be a function on E where $Range(f) \subset E \subset \mathbb{R}$.

Then for each $x \in D$, $f(x) \in E$ and therefore $g(f(x))$ yields a well-defined number in $\mathbb{R}$.

Thus, we obtain a function $g \circ f : D \to \mathbb{R}$ called the composition of f and g defined as $g \circ f(x) = g(f(x))$, $x \in \mathbb{R}$.

Example : $f(x) = x^2 + 1$ is a function from $\mathbb{R}$ to $\mathbb{R}$. $g(x) = \sqrt{x}$ is a function from $D = \{x \in \mathbb{R} \mid x \geq 0\}$ to $\mathbb{R}$.

Then $g \circ f(x) = \sqrt{x^2 + 1}$.

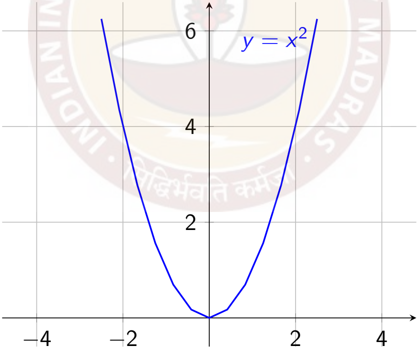
$Range(f) = \{y \in \mathbb{R} \mid y = x^2+1 \text{ for some } x\} = [1, \infty) \subseteq Domain(g) = [0, \infty)$.

Recall : the graph of a function

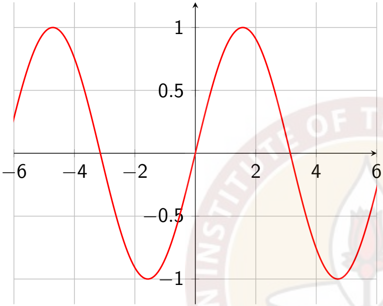
Let $f : X \to Y$ be a function. Then the graph of f is the subset
$$\Gamma(f) = \{(x, f(x)) | x \in X\} \subseteq X \times Y.$$

Let $f : D \to \mathbb{R}$ be a function where $D \subseteq \mathbb{R}$. Then the graph of f can be drawn as a in $\mathbb{R}^2$ by considering points $\{(x, y) | y = f(x)\}$.

Example :

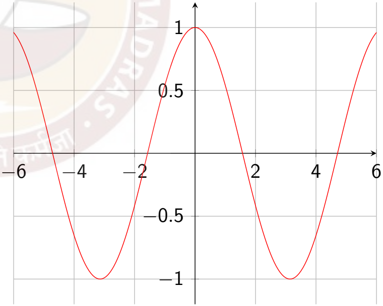


Graphs of functions : recall the trigonometric functions

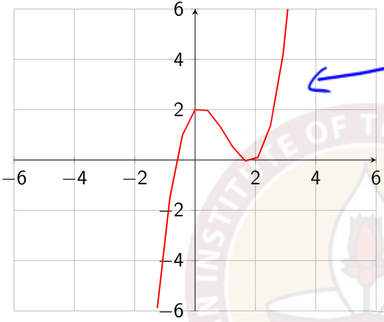


Graph of $y = sin(x)$

Graph of $y = cos(x)$



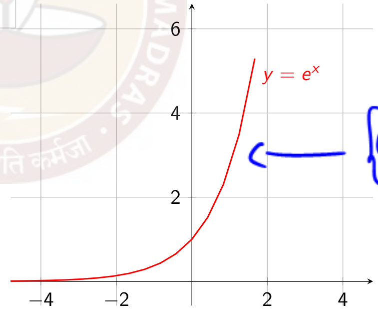
Graphs of functions : more examples



$\Gamma(x^3 - 3x^2 + x + 2)$

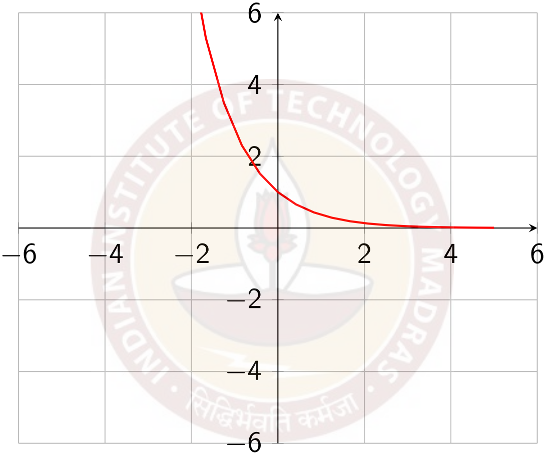
$\{(x, x^3 - 3x^2 + x + 2) \mid x \in \mathbb{R}\}$

Graph of $y = e^x$



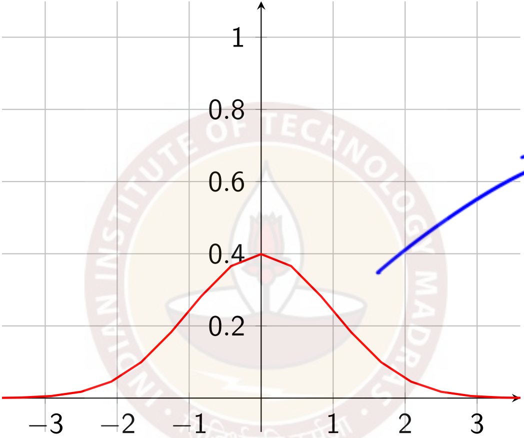
$\{(x, e^x) \mid x \in \mathbb{R}\}$

Graphs of functions : exponential decay



Graph of $y = e^{-x}$

Graphs of functions : the normal distribution



$\left\{\left(x, \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right) \mid x \in \mathbb{R}\right\}$

Graph of $y = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

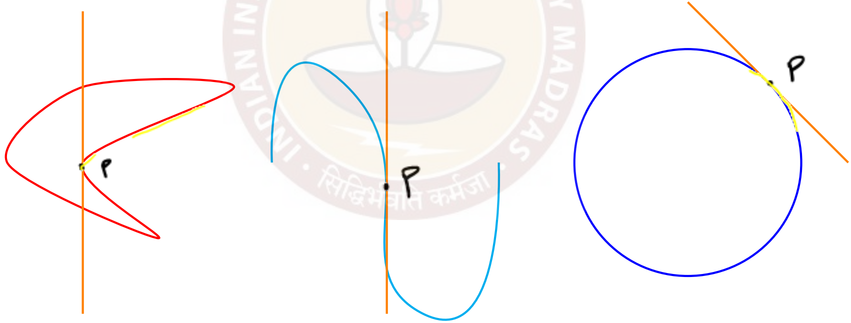
Curves (contd.)

A curve can be thought of as a figure obtained by bending a line at various places.

The intuition of a tangent line to a curve

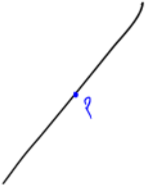
A tangent line to a curve C at a point p (on C) is a line which represents the *instantaneous* direction in which the curve C moves at the point p.

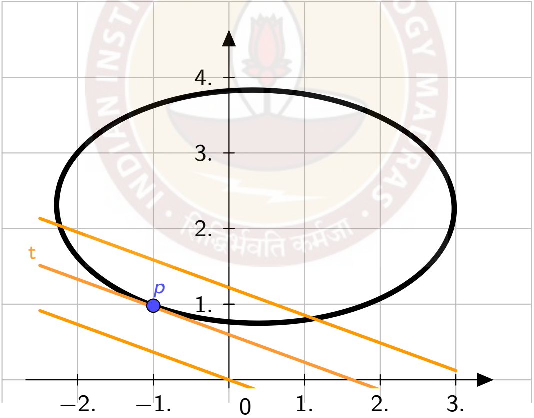
Traditionally, it was thought of as a line which *just touches* the curve at that point.

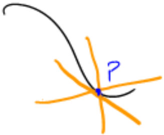


Tangent lines : some means of identification

Often, a tangent line to a curve C at a point p (on C) has the property that it passes through the point p but does not intersect the curve C in any other point *close to* the point p, and lines parallel and close to it either do not intersect C *close to* p, or intersect the curve C in two (or more) points close to p.







What is a tangent (line) to a function?

Let $f : D \to \mathbb{R}$ be a function where D is a subset of $\mathbb{R}$. Assume that $\Gamma(f)$, the graph of f is a curve. Let $x \in D$.

Then a tangent (line) to f at x is a tangent (line) to $\Gamma(f)$ at the point $(x, f(x))$.

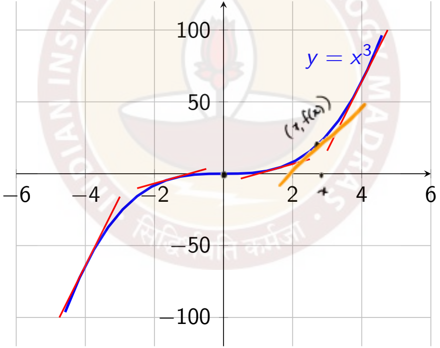
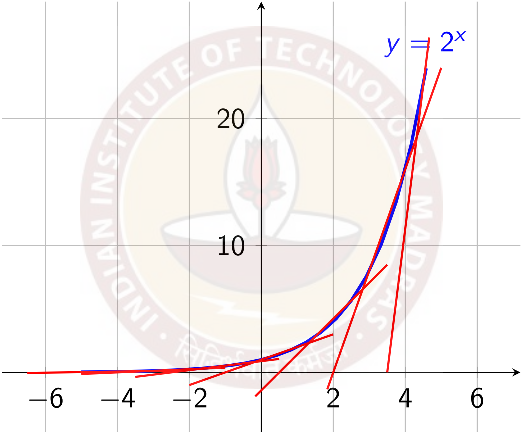


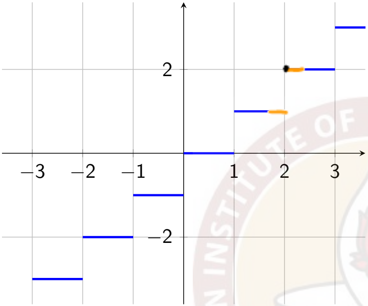
Figure: Tangent lines for $y = x^3$

$f(x) = x^3$

Tangent line for $y = 2^x$

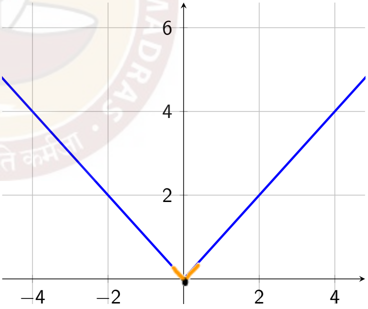


Examples advising caution



$\Gamma\left(\lfloor x \rfloor\right)$

Graph of $y = |x|$



Example : The limit of a sequence

Consider the sequence of numbers $2 - \frac{1}{n}$ as n increases.

More examples : Limits of sequences

The sequence $2 - \frac{1}{n^2}$ as n increases :

More examples : Limits of sequences

The sequence $2 - \frac{1}{(1+log(n))^{1.1}}$ as n increases :

What is the limit of a sequence?

Let $\{a_n\}$ be a sequence of real numbers. We say that $\{a_n\}$ has limit $a \in \mathbb{R}$ if as n increases, the numbers a_n come closer and closer to a.

Other equivalent terminology :

1. $\{a_n\}$ tends to a.
2. $\{a_n\}$ converges to a.
3. $\lim\limits_{n\to\infty} a_n = a$.
4. $a_n \to a$.
5. $a_n \xrightarrow{n\to\infty} a$.
6. $\lim a_n = a$.
7. $\lim\limits_{n\to\infty} \{a_n\} = a$.
8. $\lim \{a_n\} = a$.

sequence $\{a_n\}$ has limit a.

Convergent and divergent sequences

A sequence $\{a_n\}$ is called convergent if it converges to some limit (i.e. a real number).

Example : the sequence $\{\frac{1}{n}\}$ is convergent and has limit 0.

A sequence $\{a_n\}$ is called divergent if it is not convergent.

Example : the sequence $\{(-1)^n\}$ is divergent.

Subsequences

A subsequence of a sequence is a new sequence formed by (possibly) excluding some entries of a sequence.

Example :
Sequence : $-1 \quad \mathbf{1} \quad -1 \quad \mathbf{1} \quad -1 \quad \mathbf{1} \quad -1 \quad \mathbf{1} \quad -1 \quad \mathbf{1} \quad \ldots$
Subsequence : $1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad \ldots$

More examples of convergent and divergent sequences

1. The sequence $\{n\}$ is divergent.

2. The sequence $\{-n\}$ is divergent.

3. Let $x \in \mathbb{R}$. Then the sequence $\left\{\sum_{k=0}^{n} \frac{x^k}{k!}\right\}$ is convergent and converges to e^x.

4. Let $x \in \mathbb{R}$. Then $\left\{\left(1 + \frac{x}{n}\right)^n\right\}$ converges to e^x.

5. The sequence $\left\{n\left(\frac{\sqrt{2\pi n}}{n!}\right)^{\frac{1}{n}}\right\}$ converges to e.

6. The sequence $\left\{\frac{n}{\sqrt[n]{n!}}\right\}$ converges to e.

Useful rules regarding convergence of sequences

1. If $a_n \to a$, then every subsequence of $\{a_n\}$ also converges to a.
2. If $a_n \to a$ and $b_n \to b$, then $a_n + b_n \to a + b$.
3. If $a_n \to a$ and $c \in \mathbb{R}$, then $ca_n \to ca$.
4. If $a_n \to a$ and $b_n \to b$, then $a_n - b_n \to a - b$.
5. If $a_n \to a$ and $b_n \to b$, then $a_n b_n \to ab$.
6. If $a_n \to a$ and f is a polynomial function in one variable, then $f(a_n) \to f(a)$.
7. If $a_n \to a$ and $b_n \to b$ and $b \neq 0$, then $\dfrac{a_n}{b_n} \to \dfrac{a}{b}$.
8. If $a_n \to a$ and $c \in \mathbb{R}$, then $c^{a_n} \to c^a$.
9. If $a_n \to a$ and $c \in \mathbb{R}$ such that $a_n > 0 \forall n$ and $a, c > 0$, then $log_c(a_n) \to log_c(a)$.
10. The sandwich principle : If $a_n \to a$ and $b_n \to a$ and $\{c_n\}$ is a sequence such that $a_n \leq c_n \leq b_n$, then $c_n \to a$.

Some examples of applying the rules

1. $\left\{ \dfrac{(-1)^n}{n} \right\}$ converges to 0.

$\frac{-1}{1}, \frac{1}{2}, \frac{-1}{3}, \frac{1}{4}, \frac{-1}{5},$

$c_n = \frac{(-1)^n}{n}$

$\underbrace{\frac{1}{n}}_{b_n} \to 0, \quad \underbrace{\frac{-1}{n}}_{a_n} \to 0$

$a_n \leq c_n \leq b_n \to 0$

$a_n \to 0, \quad c_n \to 0$

2. $\left\{ \dfrac{\overset{a_n}{\dfrac{1}{\ln(1+n)}} + \overset{b_n}{\dfrac{5n^2}{1+n^2}}}{\left(1+\frac{1}{n}\right)^{2n}} \right\}$ converges to $\dfrac{5}{e^2}$

$\ln(1+n) \to \infty$ (diverges to ∞)

$\frac{1}{\ln(1+n)} \to 0$

$\left(1+\frac{1}{n}\right)^{2n} = \left\{ \underbrace{\left(1+\frac{1}{n}\right)^n}_{\to e} \right\}^2 \to e^2$

$\frac{5n^2}{1+n^2} = \frac{5}{\frac{1}{n^2}+1} \to \frac{5}{1} = 5$

$1+\frac{1}{n^2} \to 1$

Examples

Recall that for any convergent sequence $a_n \to a$, we obtain that $a_n^2 \to a^2$.

Consider the function $f(x) = x^2$. Then we can rewrite the above statement as $f(a_n) \to f(a)$ whenever $a_n \to a$.

In contrast, consider the floor function $g(x) = \lfloor x \rfloor$.

If one takes a sequence a_n decreasing to 2, then indeed $g(a_n) \to g(2) = 2$.

However, if one takes a sequence a_n increasing to 2, then $g(a_n) \to 1$.

Note also that for $g(x)$ this happens at each integer value and if a is a non-integer value, then indeed $g(a_n) \to g(a)$ whenever $a_n \to a$.

Another example

Consider the function $f(x) = \begin{cases} 1 & \text{if } x \text{ is a rational number} \\ 0 & \text{if } x \text{ is not a rational number} \end{cases}$.

Recall that $\sqrt{2}$ is an irrational number. We can construct a sequence of rational numbers which tends to $\sqrt{2}$.

Then $f(a_n) = 1 \forall n$ whereas $f(\sqrt{2}) = 0$. Thus even though $a_n \to a$, $f(a_n) \not\longrightarrow f(\sqrt{2})$.

Limit of a function at a point from the left

Let f be a function and a be a point such that $a_n \to a$ where a_n belongs to the domain of definition of f.

If there is a real number L such that $f(a_n) \to L$ for all sequences a_n such that $a_n \to a$ and $a_n < a$, then we say that the limit of f at a from the left exists and equals L.

We denote this by $\lim\limits_{x \to a^-} f(x) = L$.

If there is no such number L then we say that the limit of f at a from the left does not exist.

An equivalent way of thinking of $\lim\limits_{x \to a^-} f(x) = L$ is that as x comes closer and closer to a from the left, $f(x)$ eventually comes closer and closer to L.

Limit of a function at a point from the right

Similarly, if there is a real number R such that $f(a_n) \to R$ for all sequences a_n such that $a_n \to a$ and $a_n > a$, then we say that the limit of f at a from the right exists and equals R.

We denote this by $\lim_{x \to a^+} f(x) = L$.

If there is no such number R then we say that the limit of f at a from the right does not exist.

An equivalent way of thinking of $\lim_{x \to a^+} f(x) = R$ is that as x comes closer and closer to a from the left, $f(x)$ eventually comes closer and closer to R.

Limit of a function at a point

Let f be a function and a be a point such that $a_n \to a$ where a_n belongs to the domain of definition of f.

Suppose the limit of f at a from both sides (i.e. left and right) exist and are equal i.e. $\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x)$. Then we say that the limit of f at a exists and equals the left (and right) limit.

We denote it by $\lim_{x \to a} f(x)$.

$f(x) = x^2$

$\lim_{x \to a^-} x^2 = \underset{f(a)}{a^2} = \lim_{x \to a^+} x^2 = \lim_{x \to a} x^2 \; : \; f(x) = x^2$

$g(x) = \lfloor x \rfloor \; : \; \lim_{x \to a^-} \lfloor x \rfloor = \lim_{x \to a^+} \lfloor x \rfloor = \lfloor a \rfloor = g(a)$; if $a \notin \mathbb{Z}$

$\lim_{x \to 2^-} \lfloor x \rfloor = 1, \; \lim_{x \to 2^+} \lfloor x \rfloor = 2 \; : \; \lim_{x \to 2} g(x)$ DNE

$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q} \\ 1 & \text{if } x \notin \mathbb{Q} \end{cases}$

$\lim_{x \to a^-} f(x), \; \lim_{x \to a^+} f(x)$ DNE $: \; \lim_{x \to a} f(x)$ DNE

The limit as $x \to (\pm)\infty$

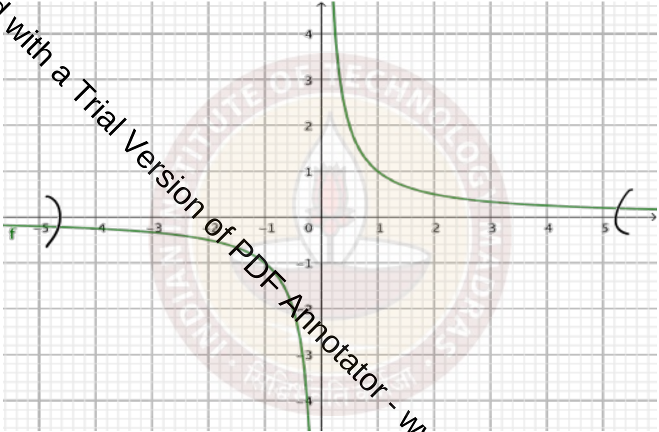
Let f be a function such that there is an M such that it is defined for all $x > M$.

Suppose for all sequences x_n diverging to ∞, there exists L such that $f(x_n)$ converges to L (i.e. as x becomes larger and larger, $f(x)$ eventually gets closer and closer to L). Then we say that $\lim_{x\to\infty} f(x)$ exists and equals L.

Similarly, let f be a function such that there is an N such that it is defined for all $x < N$.

Suppose for all sequences x_n diverging to $-\infty$, there exists L such that $f(x_n)$ converges to L (i.e. as x becomes smaller and smaller, $f(x)$ eventually gets closer and closer to L). Then we say that $\lim_{x\to-\infty} f(x)$ exists and equals L.

The function $\frac{1}{x}$



$$\lim_{x\to\infty}\frac{1}{x}=\lim_{x\to-\infty}\frac{1}{x}=0.$$ Both $\lim_{x\to-0^-}\frac{1}{x}$ and $\lim_{x\to-0^+}\frac{1}{x}$ do not exist.

Some basic examples

1. $\lim_{x \to a} x^k; k \geq 0$

   $= a^k$

2. $\lim_{x \to a} x^x; k < 0, a \neq 0$

   $= a^k$

3. $\lim_{x \to a} e^x$

   $= e^a$

4. $\lim_{x \to a} log_e(x); a > 0$

   $= \log_e(a)$

5. $\lim_{x \to a} sin(x)$

   $= \sin(a)$

6. $\lim_{x \to a} tan(x); a \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

   $= \tan(a)$

Finding limits by substitution : beware

Suppose we want to find the value of the limit of a function $f(x)$ at the point a i.e. $\lim_{x \to a} f(x)$. Often we can substitute the value of a in the expression for $f(x)$ and obtain the limit.

Unfortunately, this does not work when the function gets slightly complicated or the point a does not belong to the domain of definition of $f(x)$.

Examples :

$$\lim_{x \to 2} \frac{x^2 - 5x + 6}{x - 2}.$$

$$\frac{2^2 - 5 \times 2 + 6}{2 - 2} = \frac{0}{0}$$

$$= \frac{(x-2)(x-3)}{(x-2)} = x - 3.$$

$$-1.$$

$$\lim_{x \to 0} \frac{1}{x}$$

DNE

Cannot substitute

Some known limits

1. $\lim\limits_{x\to 0} \dfrac{sin(x)}{x} = 1.$

unit circle

$\ldots \leq \dfrac{\sin(x)}{x} \leq 1.$

2. $\lim\limits_{x\to 0} \dfrac{1 + log_e(x)}{x} = 1.$

3. $\lim\limits_{x\to\infty} \dfrac{a + be^x}{c + de^x} = \dfrac{b}{d}.$

$\hookrightarrow \; = \lim\limits_{x\to\infty} \dfrac{e^{-x}+b}{c\,e^{-x}+d} \quad \dfrac{a\times 0 + b}{c\times 0 + d} = b/d.$

Continuity of a function at a point

Let f be a function and a be a point such that $a_n \to a$ where a_n and a belong to the domain of f.

Then the function f is said to be continuous at the point a if $\lim_{x \to a} f(x) = f(a)$.

i.e. continuity means "the limit at a can be obtained by evaluating the function at a.

$$f(x) = \begin{cases} \frac{\sin(x)}{x} & x \neq 0 \\ 1 & x = 0 \end{cases}$$

Limits and continuity

Sarang S. Sane

Recall

Recall that a sequence $\{a_n\}$ has limit (or tends to) a, if a_n eventually comes closer and closer to a as n increases.

Notation : $\lim a_n = a$ or $a_n \to a$.

Recall that the limit of a function $f(x)$ at a from the left (resp. right) exists if there is a real number M such that $f(a_n) \to M$ for all sequences a_n such that $a_n \to a$ and $a_n < a$ (resp. $a_n > a$).

Statement : the limit of f at a from the left (resp. right) exists and equals M.
Math notation : $\lim_{x \to a^-} f(x) = M$ (resp. $\lim_{x \to a^+} f(x) = M$).

Examples to remember :
$f(x) = x^2$ $\qquad$ $f(x) = \lfloor x \rfloor$

$$f(x) = \frac{1}{x} \qquad f(x) = \begin{cases} 1 & \text{if } x \text{ is a rational number} \\ 0 & \text{if } x \text{ is not a rational number} \end{cases}.$$

Recall (contd.)

Recall that the limit of a function $f(x)$ at a exists if both right and left limits exist and are equal.

In that case, the number (say M) which is the common left and right limit is called the limit of the $f(x)$ at a.

In words : the limit of f at a exists and equals M.
Math notation : $\lim_{x \to a} f(x) = M$.

Defintion : f is continuous at a if the limit of f at a exists and $\lim_{x \to a} f(x) = f(a)$. f is continuous at a is equivalent to $f(a_n) \to f(a)$ whenever $a_n \to a$.

Recall also that we have defined the notion of the limit as x tends to ∞ (resp. $-\infty$) denoted by $\lim_{x \to \infty} f(x)$ (resp. $\lim_{x \to -\infty} f(x)$.

Useful rules regarding continuity of a function at a point

1. If $\lim_{x\to a} f(x) = F$, $\lim_{x\to a} g(x) = G$, then $\lim_{x\to a}(f+g)(x) = F+G$.
2. If $\lim_{x\to a} f(x) = F$ and $c \in \mathbb{R}$, then $\lim_{x\to a}(cf)(x) = cF$.
3. If $\lim_{x\to a} f(x) = F$, $\lim_{x\to a} g(x) = G$, then $\lim_{x\to a}(f-g)(x) = F-G$.
4. If $\lim_{x\to a} f(x) = F$, $\lim_{x\to a} g(x) = G$, then $\lim_{x\to a}(fg)(x) = FG$.
5. If $\lim_{x\to a} f(x) = F$, $\lim_{x\to a} g(x) = G \neq 0$, then the function $\frac{f}{g}$ is defined in at least a small interval around a and $\lim_{x\to a} \frac{f}{g}(x) = \frac{F}{G}$.
6. The sandwich principle : If $\lim_{x\to a} f(x) = L$, $\lim_{x\to a} g(x) = L$, and $h(x)$ is a function such that $f(x) \leq h(x) \leq g(x)$, then $\lim_{x\to a} f(x) = L$.

Examples

1. $f(x) = 5x^3 + 0.45x^2 - 2x + 100$

Note: $\lim_{x\to a} x^n = a^n$

$$\lim_{x\to a} f(x) = 5\lim_{x\to a} x^3 + 0.45\lim_{x\to a} x^2 - 2\lim_{x\to a} x + \lim_{x\to a} 100 = 5a^3 + 0.45a^2 - 2a + 100 = f(a)$$

2. $f(x) = \lim_{x\to 0} \dfrac{5x^3 + 0.45x^2 - 2x + 100}{x^2 - 5x + 6}$

$\lim_{x\to 0} x^2 - 5x + 6 = 0^2 - 5\times 0 + 6 = 6 \neq 0.$

$$= \frac{\lim_{x\to 0} 5x^3 + 0.45x^2 - 2x + 100}{\lim_{x\to 0} x^2 - 5x + 6} = \frac{100}{6} = 16.666\ldots$$

3. $\boxed{\lim_{x\to 0} \dfrac{\sin(x)}{x} = 1.}$

Area of $\Delta OBD = \frac{1}{2} \times 1 \times |BD| = \tan x/2$

Area of sector $= \frac{x}{2}$.

$\pi = 2\pi/2$ is the area of the full circle.

$|OB| = |OC| = 1.$

$|AC| = \sin x.$

$|\text{Arc } BC| = x$

$\&\ \sin x \le x.$

$\frac{|BD|}{|OB|} = \tan x$

$\Rightarrow |BD| = \tan x.$

$\frac{x}{2} \le \frac{\tan x}{2}$

$\Rightarrow x \le \tan x$

$\sin x \le x \le \tan x$

$1 \le \frac{x}{\sin x} \le \frac{1}{\cos x}$

$1 \ge \frac{\sin x}{x} \ge \cos x$

$\lim_{x\to 0} 1 \ge \lim_{x\to 0} \frac{\sin x}{x} \ge \lim_{x\to 0} \cos x = \cos 0 = 1.$

$\lim_{x\to 0} \frac{\sin x}{x} = 1.$

More examples

1. $\lim_{x \to 0} \frac{tan(x)}{x}$ $= \lim_{x\to 0} \frac{\sin x}{x} \times \frac{1}{\cos x} = \lim_{x\to 0} \frac{\sin x}{x} \times \lim_{x\to 0} \frac{1}{\cos x}$

$= 1 \times \frac{1}{\cos 0} = 1 \times \frac{1}{1} = 1.$

2. $\lim_{x \to 0} \frac{1 - cos(x)}{x^2}$ $= \lim_{x\to 0} \frac{2 \sin^2 \frac{x}{2}}{x^2}$

$= \frac{2}{4} \lim_{x\to 0} \frac{\sin^2 \frac{x}{2}}{x^2/4} = \frac{1}{2} \lim_{y\to 0} \left(\frac{\sin y}{y}\right)^2$

$= \frac{1}{2}.$

$\cos 2\theta = \cos^2\theta - \sin^2\theta$
$= 1 - 2\sin^2\theta$
$\Rightarrow 1 - \cos 2\theta = 2\sin^2\theta$

3. $f(x) = \lim_{x \to 0} \frac{log_e(1 + x)}{x}$

Try this.

Continuity of a function

The function f is said to be continuous if it is continuous at all points in its domain i.e. for all points a for which $f(a)$ is defined, $\lim_{x \to a} f(x) = f(a)$.

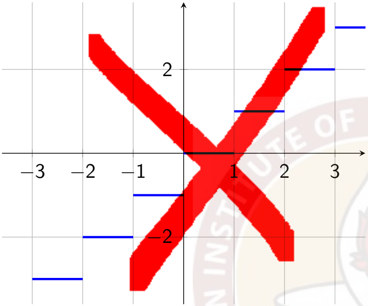
Algebraically this means : if for a sequence of real numbers $\{a_n\}$ the limit $\lim a_n$ exists, then so does the limit $\lim f(a_n)$ and $\lim f(a_n) = f(\lim a_n)$.

We can think of continuity of f as being able to draw the graph of f without lifting our pencils.

Or equivalently that there are no jumps or breaks in the graph of the function.

Examples : Polynomials, rational functions with non-zero denominators, e^x, $log(x); x > 0$, $sin(x)$, $cos(x)$.

Tangents : Examples advising caution : where are we?



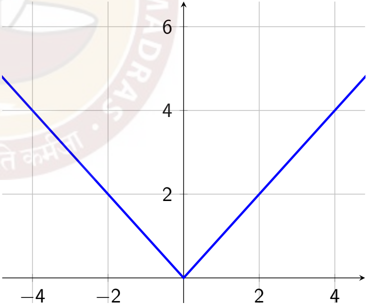
$\Gamma(\lfloor x \rfloor)$

Will not talk about tangents at points of discontinuity

$\lim_{x \to 0^-} |x| = \lim_{x\to 0^-} -x = 0$ " $|0|$.

$\lim_{x \to 0^+} |x| = 0$ " $\lim_{x \to 0} x$

Graph of $y = |x|$



Differentiability and the derivative

Sarang S. Sane

Example : the idea of rate of change

A truck travels the 2900 km distance from Jalandhar, Punjab to Tiruchirappalli (Trichy), Tamil Nadu in about 72 hrs.

What is its speed?

After reaching Tiruchirappalli, the driver tells his friend that he was fined by a traffic constable near Nagpur, Maharashtra for speeding. The friend calculates his speed from the above data and says the constable is wrong and that the fine was unjustified.

Is the friend correct?

The driver then mentions that he covered the 260 km. stretch in Maharashtra in 4 hours with about an hour's break for breakfast. The friend now recalculates and takes back his opinion that the fine is unjustified.

Is the friend correct now?

Example (contd.) and the take-home

Average speed and instantaneous speed are different concepts and one cannot deduce instantaneous speed from average speed.

To calculate instantaneous speed at some time we have to compute the distance travelled in a very short period of time around that time, and divide by that period of time.

e.g. one could take the distance travelled in 1 minute after that time; if it is y km., then the instantaneous speed would be approximately $60y$ km/hr.

Ideally one should take as small a time interval as one can i.e. an infinitesimal time. Thus, we obtain a limit !

$$\text{Infinitesimal speed} = \lim_{\Delta t \to 0} \frac{\text{Distance travelled in time } \Delta t}{\Delta t}$$

(where Δt is measured in seconds and distance in km).

Differentiability

Definition

Let f be a function defined on an open interval around a. Then f is differentiable at a if $\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$ exists.

Examples

$f(x) = x$ is differentiable at a point a.

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{x \to 0} \frac{a+h-a}{h} = \lim_{h \to 0} \frac{h}{h} = \lim_{h \to 0} 1$$
$$= 1.$$

$f(x) = sin(x)$ is differentiable at the point 0.

$$\lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{\sin(0+h) - \sin 0}{h} = \lim_{h \to 0} \frac{\sin h}{h}$$
$$= 1.$$

More examples

$f(x) = |x|$ is NOT differentiable at 0.

$$\lim_{h\to 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0} \frac{|h|-|0|}{h} = \lim_{h\to 0} \frac{|h|}{h} \quad \frac{|h|}{h} \text{ DNE.}$$

$$\frac{|h|}{h} = \begin{cases} 1 & h>0 \\ -1 & h<0 \end{cases}$$

$$\lim_{h\to 0^+} \frac{|h|}{h} = 1. \qquad \lim_{h\to 0^-} \frac{|h|}{h} = -1.$$

$f(x) = x^{\frac{1}{3}}$ is NOT differentiable at 0.

$$\lim_{h\to 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0} \frac{h^{1/3}-0}{h} = \lim_{h\to 0} \frac{h^{1/3}}{h} = \lim_{h\to 0} h^{-2/3} = \lim_{h\to 0} \frac{1}{\sqrt[3]{h^2}}$$

$$\lim_{h\to 0^+} \frac{1}{\sqrt[3]{h^2}} \text{ DNE (diverges to } \infty)$$

$$\lim_{h\to 0^-} \frac{1}{\sqrt[3]{h^2}} \text{ DNE (diverges to } -\infty)$$

$f(x) = \lfloor x \rfloor$ is NOT differentiable at any integer point.

$$\lim_{h\to 0} \frac{\lfloor 0+h \rfloor - \lfloor 0 \rfloor}{h} = \lim_{h\to 0} \frac{\lfloor h \rfloor}{h}$$

$$\frac{\lfloor h \rfloor}{h} = \begin{cases} 0 & \text{if } 0<h<1 \\ -1/h & \text{if } -1<h<0 \end{cases}$$

$$\lim_{h\to 0^+} \frac{\lfloor h \rfloor}{h} = 0 \quad \text{but} \quad \lim_{h\to 0^-} \frac{-1}{h} \text{ DNE (diverges to } +\infty)$$

Differentiability implies continuity

Fact : If f is differentiable at a, then it is continuous at a.

$\lim_{x \to a} f(x)$

We know $\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = L$.

$\lim_{h \to 0} h = 0$.

$\lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{h} \right) \cdot h = L \times 0 = 0$.

$\Rightarrow \lim_{h \to 0} f(a+h) - f(a) = 0 \Rightarrow \lim_{h \to 0} f(a+h) = f(a)$

$\Rightarrow \lim_{x \to a} f(x) = f(a)$.

The derivative function

For a function $f(x)$ its derivative function, $f'(x)$ or $\frac{df}{dx}(x)$ is

$$f'(x) = \frac{df}{dx}(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Its domain consists of those points at which the function $f(x)$ is differentiable.

Examples

For $f(x) = x$, then $f'(x) =$ 1.

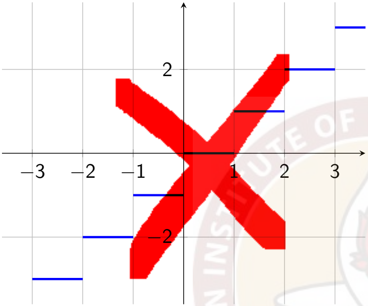
$f(x) = sin(x)$, then $f'(x) =$ $\cos(x)$. $f'(0) = 1$.

$$\lim_{h \to 0} \frac{\sin(a+h) - \sin(a)}{h} = \lim_{h \to 0} \frac{\sin(a)\cos(h) + \cos(a)\sin(h) - \sin(a)}{h}$$

$$= \lim_{h \to 0} \cos(a) \frac{\sin(h)}{h} + \lim_{h \to 0} \sin(a) \frac{\cos(h) - 1}{h}$$

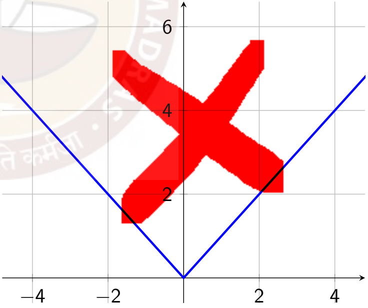
$$= \cos(a) + 0 = \cos(a).$$

Tangents : Examples advising caution : where are we?



$\Gamma\left(\lfloor x \rfloor\right)$

Graph of $y = |x|$



Computing derivatives and L'Hôpital's rule

Sarang S. Sane

Recall

Definition

Let f be a function defined on an open interval around a. Then f is differentiable at a if $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ exists.

Definition

For a function $f(x)$ its derivative function, $f'(x)$ or $\frac{df}{dx}(x)$ is

$$f'(x) = \frac{df}{dx}(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}.$$

Its domain consists of those points at which the function $f(x)$ is differentiable.

Fact

If f is differentiable at a, then it is continuous at a.

Useful rules about derivatives

Linearity :

1. If $f(x)$ and $g(x)$ are differentiable at the point a, then so is $(f+g)(x)$ and $(f+g)'(a) = f'(a) + g'(a)$.
2. If $f(x)$ is differentiable at the point a and $c \in \mathbb{R}$, then $(cf)(x)$ is also differentiable at the point a and $(cf)'(a) = cf'(a)$.
3. If $f(x)$ and $g(x)$ are differentiable at the point a, then so is $(f-g)(x)$ and $(f-g)'(a) = f'(a) - g'(a)$.

The product rule

If $f(x)$ and $g(x)$ are differentiable at a, then so is $(fg)(x)$ and

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

$$\lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} = \lim_{h\to 0} \frac{f(a+h)\left(g(a+h) - g(a)\right) + \left(f(a+h) - f(a)\right)g(a)}{h}$$

Useful rules about derivatives (contd.)

The quotient rule

If $f(x)$ and $g(x)$ are differentiable at the point a and $g(a) \neq 0$, then so is $\frac{f}{g}(x)$ and

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Composition : the chain rule

If $f(x)$ and $g(x)$ are differentiable functions, then so is the function $f(g(x))$ and its derivative is :

$$(f(g))'(x) = f'(g(x))g'(x).$$

Examples

Guess: $f(x) = x^n$
$f'(x) = n x^{n-1}$.

$f(x) = x^2 = x \times x$

$f'(x) = (x)' \times x + x \times (x)'$

$= 1 \times x + x \times 1$

$= x + x = 2x$.

$f(x) = x^3 = x^2 \times x$

$f'(x) = 2x \times x + x^2 \times 1$

$= 2x^2 + x^2$

$= 3x^2$.

$f(x) = x^a$

$f'(x)$

$= a x^{a-1}$.

$f(x) = x^n; n \in \mathbb{N}$

$= x^{n-1} \times x$

$f'(x) = (n-1) x^{n-2} \times x + x^{n-1} \times 1$

$= (n-1) x^{n-1} + x^{n-1} = n x^{n-1}$.

$f(x) = 5x^3 - 17x^2 + \pi x - 0.5$

$f'(x) = 5 \times 3x^2 - 17 \times 2x + \pi \times 1 - 0$

$= 15x^2 - 34x + \pi$

Gen. statement for a polynomial.

$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$

$\Rightarrow f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1$

$f(x) = x^7 sin(x)$

$f'(x) = 7x^6 \times \sin(x) + x^7 \times \cos(x)$

$= 7x^6 \sin(x) + x^7 \cos(x)$.

Examples

$f(x) = tan(x)$ $= \frac{\sin(x)}{\cos(x)}$

$f'(x) = \frac{\cos(x) \times \cos(x) - \sin(x) \times (-\sin(x))}{\cos(x)^2}$

$= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}$

$= \sec^2(x).$

$f(x) = \frac{1}{x^r} \quad ; r > 0$

$= x^{-r}$

$f'(x) = \frac{\frac{d}{dx}(1) \times x^r - 1 \times \frac{d}{dx}(x^r)}{(x^r)^2}$

$= \frac{-r x^{r-1}}{x^{2r}} = -r\, x^{r-1-2r}$

$= -r\, x^{-r-1}$

$f(x) = tan(2x)$ $= \frac{2\tan(x)}{1-\tan^2(x)}$

$= \frac{\sin(2x)}{\cos(2x)}$

$g(x) = 2x. \quad h(x) = \tan(x)$

$f(x) = h(g(x)).$

$f'(x) = h'(g(x))\, g'(x)$

$= \sec^2(2x) \times 2$

$= 2\sec^2(2x)$

Indeterminate limits

Let $f(x)$ and $g(x)$ be functions and suppose $f(x)$ and $g(x)$ are defined on an interval around the point c.

Further suppose $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$ or both limits diverge to ∞ or both limits diverge to $-\infty$.

Suppose we are interested in computing $\lim_{x \to a} \frac{f(x)}{g(x)}$.

Then we cannot use the quotient rule to compute it since the quotient rule yields an *indeterminate form* e.g. $\lim_{x \to 0} \frac{log_e(1+x)}{x} =$

In this situation, we can try and use L'Hôpital's rule.

Indeterminate limits : L'Hôpital's rule

In the situation of the indeterminate form, suppose the following conditions hold :

1. $f'(x)$ and $g'(x)$ exist on this interval (except possibly at c).
2. $g'(x) \neq 0$ on this interval (except possibly at c).
3. $\lim\limits_{x \to a} \dfrac{f'(x)}{g'(x)} = L.$

$$\text{Then} \qquad \lim_{x \to a} \frac{f(x)}{g(x)} = L.$$

e.g. $\lim\limits_{x \to 0} \dfrac{log_e(1+x)}{x} = \lim\limits_{x \to 0} \dfrac{1/1+x}{1} = \lim\limits_{x \to 0} \dfrac{1}{1+x} = 1.$

$\dfrac{d}{dx}\left(\log_e(1+x)\right) = \dfrac{1}{1+x}$, $\dfrac{d}{dx}(x) = 1.$

More examples

$$\lim_{x\to 2}\frac{x^2-5x+6}{x-2} = \lim_{x\to 2}\frac{2x-5}{1} = \lim_{x\to 2}(2x-5) = 2\times 2-5 = -1.$$

$$\lim_{x\to 0}\frac{\sin(x)}{x} = \lim_{x\to 0}\frac{\cos(x)}{1} = \lim_{x\to 0}\cos(x) = \cos(0) = 1.$$

$$\lim_{x\to\infty}\frac{a+be^x}{c+de^x} = \lim_{x\to\infty}\frac{be^x}{de^x} = \lim_{x\to\infty}\frac{b}{d} = \frac{b}{d}.$$

$$\lim_{x\to 0}\frac{1-\cos(x)}{x^2} = \lim_{x\to 0}\frac{\sin(x)}{2x} = \frac{1}{2}\lim_{x\to 0}\frac{\sin(x)}{x} = \frac{1}{2}.$$

Derivatives, tangents and linear approximation

Sarang S. Sane

Recall

Let f be a function defined on an open interval around a. Then f is differentiable at a if $\lim_{h \to 0} \dfrac{f(x+h) - f(x)}{h}$ exists.

$$f'(x) = \frac{df}{dx}(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

A tangent to $f(x)$ at a is a line which represents the *instantaneous* direction in which the graph $\Gamma(f)$ moves at $(a, f(a))$.

Traditionally, the tangent to $f(x)$ at a is thought of as a line which *just touches* $\Gamma(f)$ at $(a, f(a))$.

Tangents as limits of secants

Recall the notion of a tangent to the function f at a point a (i.e. a tangent to the graph $\Gamma(f)$ at $(a, f(a))$).

If it exists, we can think of the tangent as a "limit" of secants joining $(a, f(a))$ and nearby points $(a+h, f(a+h))$.

$$y - f(a) = \frac{f(a+h) - f(a)}{a+h-a}(x-a) = \frac{f(a+h)-f(a)}{h}(x-a)$$

is the equation of the secant.

What happens in the limit to this equation?

Tangents and derivatives

Let f be a function differentiable at the point a. Then the tangent to f at a exists and is given by

$$y = f'(a)(x - a) + f(a).$$

$$y = \frac{f(a+h) - f(a)}{h}(x-a) + f(a)$$

If f is diff. at a, then the limit of

is $f'(a)$

Suppose the tangent to f at a exists and is not vertical (i.e. is not the line $x = a$). Then f is differentiable at a and the equation of the tangent is

$$y = f'(a)(x - a) + f(a).$$

Examples

$f(x) = 5x^3 - 17x^2 + \pi x - 0.5 \quad ; \quad a = 0.$

$f'(x) = 15x^2 - 34x + \pi.$

$f'(0) = \pi.$

$\therefore$ Eqn. of tangent to f at 0 is

$y = \pi(x-0) + f(0) = \pi x - 0.5.$

$f(x) = cos(x) \quad ; \quad a = \frac{\pi}{3}$

$f'(x) = -\sin(x)$

$f'(\pi/3) = -\sin(\pi/3) = -\sqrt{3}/2$

$\therefore \quad y = -\frac{\sqrt{3}}{2}(x - \pi/3) + \cos(\pi/3)$

$= -\frac{\sqrt{3}}{2}(x - \pi/3) + \frac{1}{2}.$

$f(x) = xtan(x) \quad ; \quad a = \frac{\pi}{4}$

$f'(x) = 1 \times \tan(x) + x \times \sec^2(x)$

$= \tan x + x\sec^2(x)$

$f'(\pi/4) = \tan(\pi/4) + \frac{\pi}{4}\sec^2(\pi/4)$

$= 1 + \frac{\pi}{4} \times 2 = 1 + \frac{\pi}{2}.$

$\therefore \quad y = (1 + \frac{\pi}{2})(x - \pi/4) + \frac{\pi}{4}$

Example : $f(x) = x^{\frac{1}{3}}$

$f(x) = x^a$

$f'(x) = a x^{a-1}$.

$$\lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{h^{1/3} - 0}{h} = \lim_{h \to 0} \frac{h^{1/3}}{h} = \lim_{h \to 0} \frac{1}{h^{2/3}}$$

diverges to ∞.

$\therefore$ This limit DNE.

If $x \neq 0$

$$f'(x) = \frac{1}{3} x^{1/3 - 1} = \frac{1}{3} x^{-2/3} = \frac{1}{3x^{2/3}}$$

$y = mx + c$ works only for lines which are not vertical.

$x = 0$.

Linear approximation

Recall that a linear function is a a function of the form $L(x) = c + dx$ (in terms of our notation from linear algebra, this would be called an affine transformation).

Let $f(x)$ be a function and a be a point in the domain of f.

A linear function which takes the value $f(a)$ at a will have the form $L(x) = f(a) + m(x - a)$.

We want to choose a linear function which *best approximates* the function $f(x)$ around the point a.

$$\text{i.e. } f(x) \approx L(x) \quad \forall x \text{ close to } a.$$

Linear approximation (contd.)

If $f(x)$ is differentiable at a, then the best linear approximation is given by $L(x) = f(a) + f'(a)(x - a)$.

Conversely, if there is a best linear approximation for f at a, then f is differentiable at a (and hence the approximation is given by $L(x) = f(a) + f'(a)(x - a)$).

Examples :

$f(x) = x^3 \quad ; \quad a = 1$

$f'(1) = 3\times 1^2 = 3.$ $\quad L(x) = 3(x-1) + 1 = 3x - 2.$

$f(x) = sec(x) \quad ; \quad a = 0$

$f'(0) = \tan(0)\sec(0) = 0.$

$L(x) = 0(x-0) + \sec(0)$

$= 1.$

Local maxima/minima

Recall the following notions from maths 1 for a function $f(x)$:

An interval of increase (respectively decrease) is an interval on which the function $f(x)$ is increasing (respectively decreasing) i.e. if $x_1 < x_2$ are points belonging to that interval, then $f(x_1) < f(x_2)$ (respectively $f(x_1) > f(x_2)$).

A point x is a turning point of $f(x)$ where one of the following phenomena take place :

1. either there is an interval of increase ending at x and an interval of decrease beginning at x
2. or there is an interval of decrease ending at x and an interval of increase beginning at x.

A turning point as in 1 above is called a local maximum and a turning point as in 2 above is called a local minimum.

Example

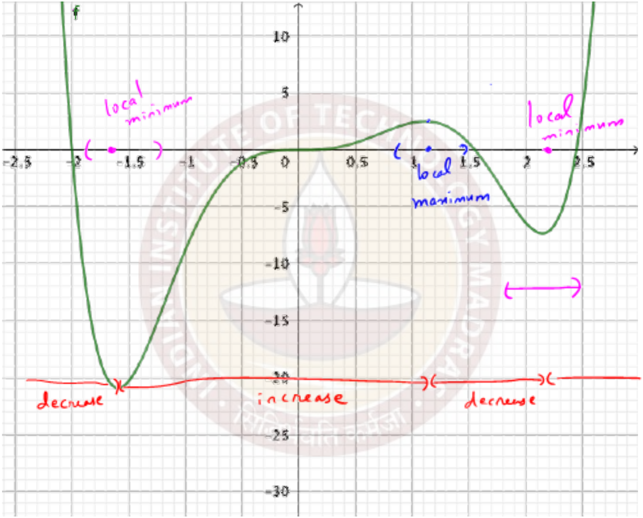


Figure: $\Gamma\left((x^2 - 4x + 3.8)(x+2)x^3\right)$

Example

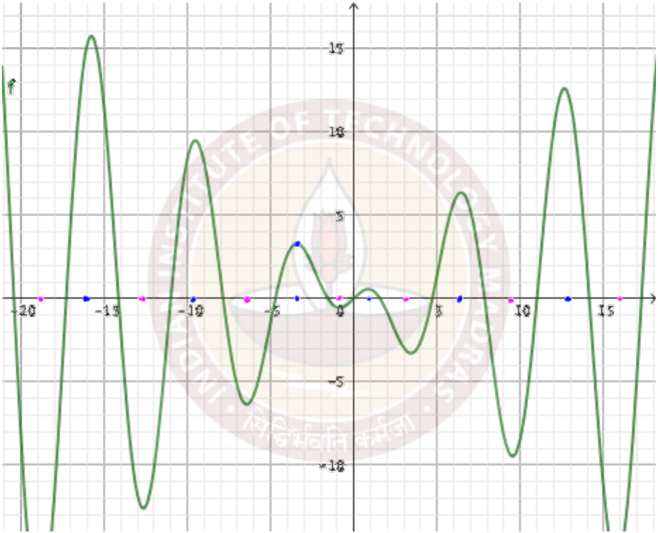


Figure: $\Gamma(x\cos(x))$

Tangents at local maxima/minima

Observe that in both examples, tangents at the turning points are horizontal (i.e. flat or parallel to the X-axis).

This is a general phenomenon that is easy to visualize.

Local maxima/minima and derivatives

How do we prove that the tangent (if it exists) is horizontal at a turning point?

Recall that the tangent to f at a exists (and is not vertical) is equivalent to f being differentiable at a. and its equation is then given by

$$y = f'(a)(x - a) + f(a).$$

Thus, proving that the tangent (if it exists) is horizontal at a turning point is equivalent to showing that $f'(a) = 0$.

a is a turning point & $f'(a)$ exists.

Suppose a is a local maximum. For some $(a-\epsilon, a+\epsilon)$

if $0<h<\epsilon$ then $f(a-h) < f(a)$ & $f(a+h) < f(a)$.

$$f'(a) = \lim_{h\to 0^+} \frac{f(a+h)-f(a)}{h} \leq 0 \quad \Rightarrow f'(a) = 0.$$

$$= \lim_{h\to 0^-} \frac{f(a+h)-f(a)}{h} \geq 0$$

Critical points and saddle points

Thus, if f is differentiable at a turning point a, it satisfies $f'(a) = 0$.

A point a is called a critical point of a function $f(x)$ if either f is not differentiable at a or $f'(a) = 0$.

Every turning point is a critical point.

Thus, in order to find turning points (i.e. local maxima/minima), we will first find critical points.

Suppose f is differentiable. Is every critical point a turning point?

Unfortunately not e.g. 0 for $f(x) = (x^2 - 4x + 3.8)(x + 2)x^3$.

A saddle point is a critical point which is not a local maximum or local minimum.

The second derivative test

Suppose f is differentiable. How do we classify critical points?

Just like the first derivative f' checks for the monotonicity of f, the second derivative f'' checks for the monotonicity of f'.

So if f is twice differentiable, we check f'' at all the critical points.

1. If a is a critical point and $f'(a) > 0$, then a is a local minimum.
2. If a is a critical point and $f'(a) < 0$, then a is a local maximum.
3. If a is a critical point and $f'(a) = 0$, then the test is inconclusive.

Examples

$f(x) = x^3 - 12x$

$f'(x) = 3x^2 - 12$ Setting it to 0, we obtain $3x^2 - 12 = 0 \Rightarrow x^2 = 4 \Rightarrow x = \pm 2$.

Critical points: ± 2. $f''(x) = 6x$.

$f''(2) = 12 > 0$ $\therefore$ 2 is a local minimum

$f''(-2) = -12 < 0$ & -2 is a local maximum

$f(x) = cos(x)$

$f'(x) = -\sin(x)$. Setting it to 0, critical points are $\{k\pi \mid k \in \mathbb{Z}\}$.

$f''(x) = -\cos(x)$.

$f''(k\pi) = -\cos(k\pi) = \begin{cases} -1 & \text{if } k \text{ is an even integer} \\ 1 & \text{if } k \text{ is an odd integer} \end{cases}$

$k\pi$; k is even are local maxima

$k\pi$; k is odd are local minima.

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

$f'(x) = 3x^2 + 2x - 1$. Setting it to 0, we obtain $3x^2 + 2x - 1 = 0 \Rightarrow$ Roots are $\dfrac{-2 \pm \sqrt{4+12}}{2 \times 3} = \dfrac{-2 \pm \sqrt{16}}{6} = \dfrac{-2 \pm 4}{6}$

Roots: $-1, \frac{1}{3}$.

$f''(x) = 6x + 2$.

$f''(-1) = -6 + 2 = -4$: local maximum

$f''(1/3) = 2 + 2 = 4$: local minimum

Local maxima/minima on closed intervals

Sometimes we want to find the local extrema of a function f on a closed interval $I = [a, b]$.

In that case, in addition to finding the extrema via the previously discussed method, it is possible that the end points could also be local extrema and so we have to consider them separately.

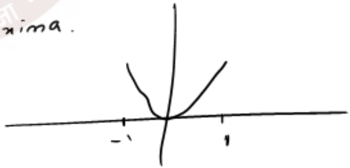
Example : $f(x) = x^2$ on the interval $[-1, 1]$.

$f'(x) = 2x$. Set it to 0.

$f''(x) = 2$. $f''(0) = 2$.

$\therefore$ It is a local minimum.

-1 & 1 are also local maxima.



(Global) maximum/minimum

Sometimes we want to find the maximum or minimum value of a function f on a particular interval I.

In general this may not exist e.g. $f(x) = \frac{1}{x}$ on $I = (0, \infty)$.

Fact : If the interval I is closed and bounded and f is continuous, the maximum and minimum must exist.

Note that the maximum and minimum are in particular local maxima or local minima unless they are on boundary points.

Thus to find the maximum and minimum, we find the critical points and the boundary points and check the value of f on all of them.

We can in fact do this on any function which is defined piecewise continuously with finitely many pieces on a closed and bounded interval.

Example

$$f(x) = \begin{cases} x^3 + x^2 - x + 5 & \text{if } 0 \le x \le 100 \\ x^3 + 2x^2 + x - 5 & \text{if } -100 \le x < 0 \end{cases}$$

$$f'(x) = \begin{cases} 3x^2 + 2x - 1 & \text{if } 0 < x < 100 \\ 3x^2 + 4x + 1 & \text{if } -100 < x < 0 \end{cases}$$

Critical pts: $\dfrac{-2 \pm \sqrt{4+12}}{6}$, $\dfrac{-2 \pm \sqrt{16-12}}{6}$

Bdry pts: $-100, 0, 100$.

$\dfrac{-2 \pm 4}{6} = \begin{cases} 1/3 \\ -1 \end{cases}$, $\dfrac{-2 \pm 2}{6} = \begin{cases} 0 \\ -2/3 \end{cases}$

local min. $f''(1/3) = 4$

local max. $f''(-1) = -4$

$f''(x) = \begin{cases} 6x + 2 \\ 6x + 4 \end{cases}$

$\boxed{f''(0) = 4}$ sep.

$\boxed{f''(-2/3) = 0}$ Inconcl.

Global max. $f(100) = 100^3 + 100^2 - 100 + 5$

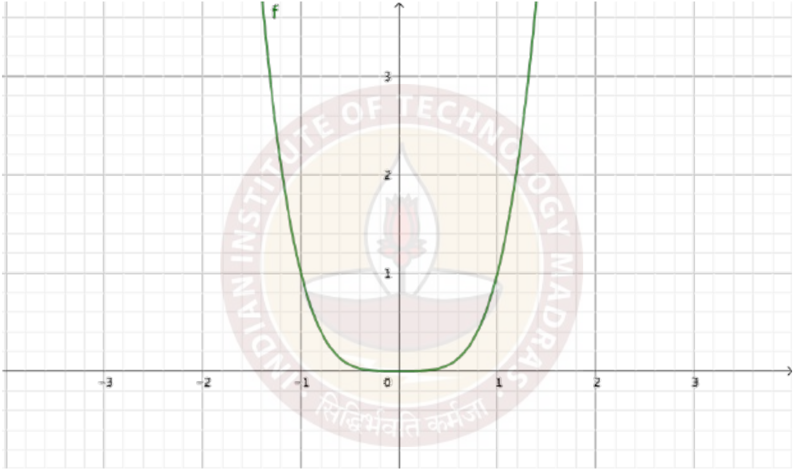
Global min. $f(-100) = (-100)^3 + 2 \times 100^2 - 100 - 5$

$f(-1) = -1 + 2 + (-1) - 5 = -5$

$f(0) = 5$ $\quad f(-2/3) = \dfrac{-8}{27} + \dfrac{8}{9} + \dfrac{(-2)}{3} - 5$

$f(1/3) =$

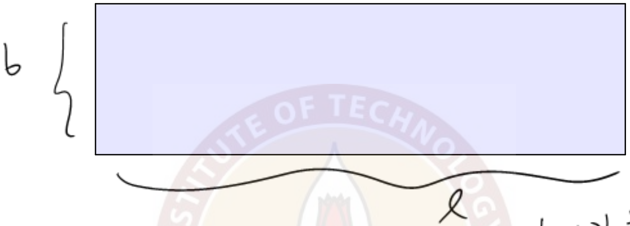
Warning example : $f(x) = x^4$



$f'(x) = 4x^3 \Rightarrow$ Critical point is 0.

$f''(x) = 12x^2 \Rightarrow f''(0) = 0$. Inconclusive.

The area of a rectangle



$A(c\ell, db) = cd\,A(\ell, b).$

$A(\ell, b) = A(\ell \cdot 1, b \cdot 1) = \ell b\, A(1,1)$

$A(1,1)$
$= 1$ sq. unit

Area of a rectangle
$= \ell b\, A(1,1)$
$= \ell b$ sq. units.

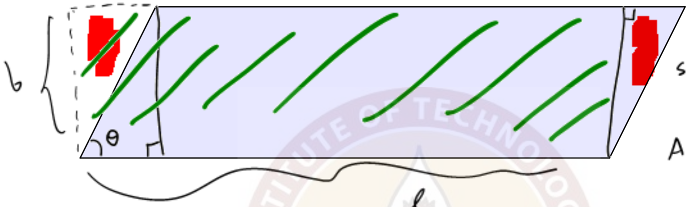
What happens if we double the length?

$A(2\ell, b)$
$= 2A(\ell, b)$

$A(\frac{m}{n}\ell, b)$
$= \frac{m}{n} A(\ell, b)$
$m, n \in \mathbb{N}$

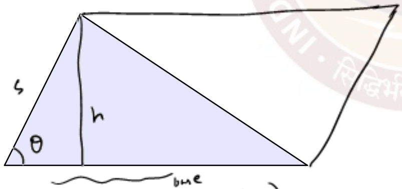
$A(c\ell, db)$
$= cd\,A(\ell, b)$

More areas : parallelograms, triangles, trapeziums

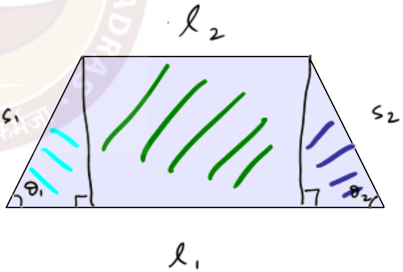


$b = s\sin(\theta)$

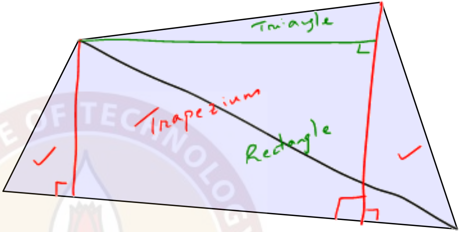
$A(\|^{m}) = l\, s\, \sin(\theta)$

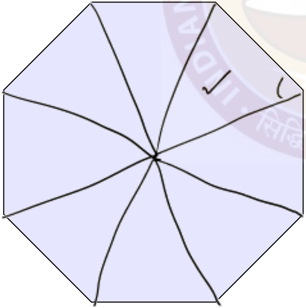


$\text{Area}(\Delta) = \frac{1}{2}\text{Area}(\|^{m})$
$= \frac{1}{2} \times \text{base} \times h$

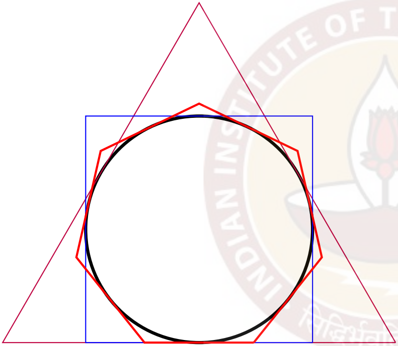


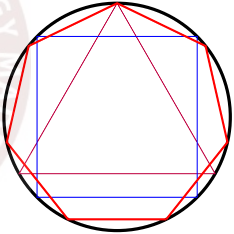
Quadrilaterals and polygons : Divide and conquer





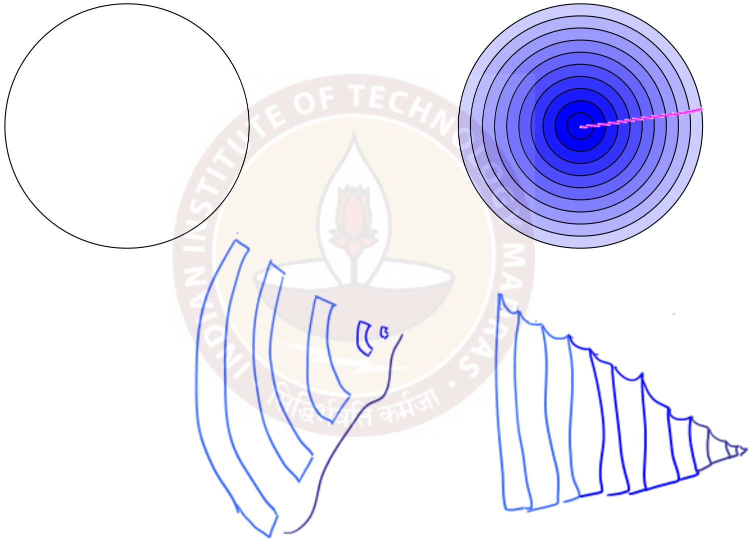
Non-linear shapes : Divide and gradually conquer





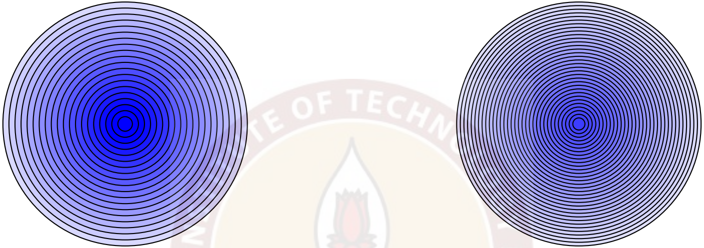
Area of the circle is πr^2.

Non-linear shapes : Divide and gradually conquer



Non-linear shapes : Contd.

Non-linear shapes : Contd.



$$\text{Area} = \frac{1}{2} \times 2\pi r \times r$$
$$= \pi r^2.$$

Strategy to compute areas

Suppose you want to compute the area of a given shape.

1. Break the shape you have into rectangles. Some part of the shape may be left out, some extra part may get included.

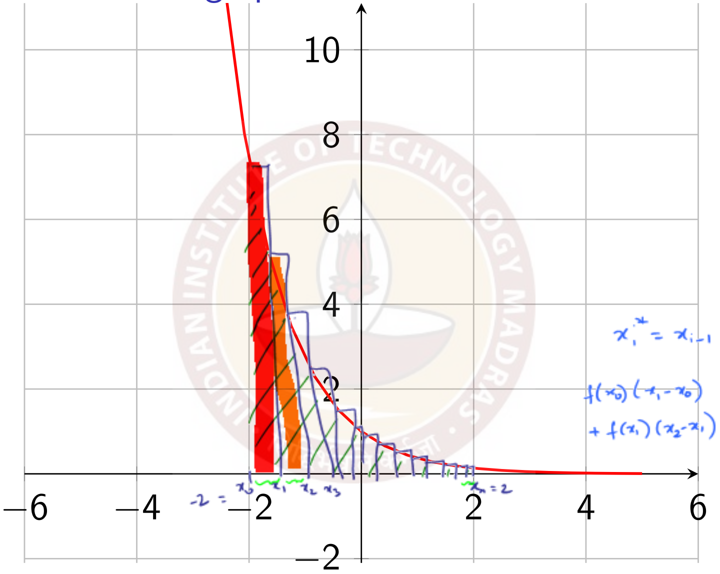
2. Calculate the area covered by the rectangles.

3. Decrease the sizes of the rectangles used. The part of the shape left out and the extra part keep decreasing and recalculate.

4. As the sizes decrease, the area of the shape is better and better approximated.

5. In the limit , you will get the area of the shape.

The area under the graph of a function



Graph of $y = e^{-x}$

Riemann sums

Let f be a function from D to $\mathbb{R}$ for some domain $D \subseteq \mathbb{R}$.

Suppose the interval $[a, b]$ is in the domain D.

Let P consist of the following data :

- a partition of $[a, b]$, i.e. a choice of intermediate points $a = x_0 < x_1 < \ldots x_n = b$
- a choice of $x_i^* \in [x_{i-1}, x_i]$

Define $\Delta x_i = x_i - x_{i-1}$ and $||P|| = \max_i \{\Delta x_i\}$.

The Riemann sum of f w.r.t. the above data is defined as

$$S(P) = \sum_{i=1}^{n} f(x_i^*)\Delta x_i.$$

The definite integral of a function

Let f be a function from D to $\mathbb{R}$ for some domain $D \subseteq \mathbb{R}$. Suppose the interval $[a, b]$ is in the domain D.

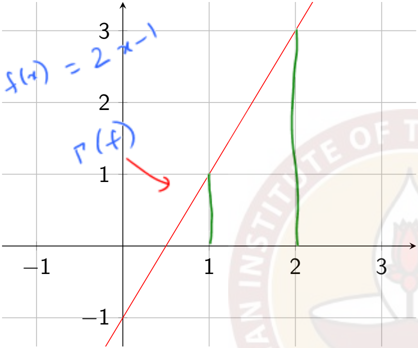
The (definite) integral of f from a to b is defined as

$$\lim_{||P|| \to \infty} S(P) = \lim_{||P|| \to \infty} \sum_{i=1}^{n} f(x_i^*)\Delta x_i.$$

It is denoted by $\displaystyle\int_a^b f(x)dx$.

Let $f \geq 0$ (resp. $f \leq 0$) be piecewise continuous on this interval. Then the area under the graph of the function $\Gamma(f)$ above the interval $[a, b]$ is measured by $\displaystyle\int_a^b f(x)dx$ (resp. $\displaystyle -\int_a^b f(x)dx$).

Example



$$P_n = \{ x_0 = 1,\ x_1 = 1 + \tfrac{1}{n},\ x_2 = 1 + \tfrac{2}{n},$$

$$x_i = 1 + i/n.$$

$$1 = x_0 < x_1 < x_2 < \cdots < x_n = 2.$$

$$\Delta x_i = \frac{1}{n}. \qquad \|P_n\| = \frac{1}{n}.$$

$$x_i^* = x_i = 1 + i/n$$

$$S(P_n) = \sum_{i=1}^{n} f(x_i^*)\, \Delta x_i$$

$$= \sum_{i=1}^{n} \left\{ 2\left(1 + \frac{i}{n}\right) - 1 \right\} \frac{1}{n}$$

$$= \sum_{i=1}^{n} \left(1 + \frac{2i}{n}\right) \frac{1}{n}$$

$$= \sum_{i=1}^{n} \frac{1}{n} + \sum_{i=1}^{n} \frac{2i}{n^2}$$

$$= 1 + \frac{2}{n^2} \sum_{i=1}^{n} i$$

$$= 1 + \frac{2}{n^2} \cdot \frac{n(n+1)}{2}$$

$$= 1 + \frac{n+1}{n} = 2 + \frac{1}{n}.$$

$$\lim_{\|P_n\| \to 0} S(P_n) = \lim_{n \to \infty} 2 + \frac{1}{n} = 2.$$

Recall : Riemann sums and the definite integral

Let f be a function defined on a domain containing the interval $[a, b]$ to $\mathbb{R}$.

Let P consist of (i) a partition $a = x_0 < x_1 < \ldots x_n = b$ of $[a, b]$ and (ii) a choice of $x_i^* \in [x_{i-1}, x_i]$.

Define $\Delta x_i = x_i - x_{i-1}$ and $||P|| = \max_i \{\Delta x_i\}$.

The Riemann sum of f w.r.t. the above data is defined as

$$S(P) = \sum_{i=1}^{n} f(x_i^*)\Delta x_i.$$

The (definite) integral of f from a to b is defined as

$$\int_a^b f(x)dx = \lim_{||P|| \to \infty} S(P) = \lim_{||P|| \to 0} \sum_{i=1}^{n} f(x_i^*)\Delta x_i.$$

Anti-derivatives aka (indefinite) integrals

Let f be a function defined from a domain D in $\mathbb{R}$ to $\mathbb{R}$.

An anti-derivative of the function f is a function F defined on the domain D such that $F'(x) = f(x)$ for all $x \in D$.

Example : An anti-derivative of $f(x) = x^7 + 2x^6 - \pi x^5 + 0.5x^4 - 9$ is $F(x) = \frac{x^8}{8} + 2\frac{x^7}{7} - \pi\frac{x^6}{6} + 0.5\frac{x^5}{5} - 9x + C$

Fact : If F is an anti-derivative of f, then so is $F_1(x) = F(x) + c$ where c is any constant (i.e. a real number). In fact, every anti-derivative has this form.

Because of this, We will use the instead of an for the anti-derivative, since any two anti-derivatives only differ by a constant.

Anti-derivatives and integrals

The anti-derivative for a function f is more often called the (indefinite) integral of f and denoted by $\int f(x)dx$.

This is explained by :

The fundamental theorem of calculus

Suppose f is continuous on the domain D which includes the interval $[a, b]$. Then an anti-derivative for f on (a, b) is given by

$$F(x) = \int_a^x f(t)dt.$$

Conversely, if f is continuous on the domain D which includes the interval $[a, b]$ and F is the (indefinite) integral of f, then the (definite) integral from a to b of f can be computed by

$$\int_a^b f(x)dx = F(b) - F(a).$$

Tables of derivatives and integrals

Function	Derivative	Integral
1	0	x
x^r	rx^{r-1}	$\frac{x^{r+1}}{r+1} \quad r \neq 1$
		$\frac{1}{ln \mid x \mid} \quad r = 1$
$sin(x)$	$cos(x)$	$-cos(x)$
$cos(x)$	$-sin(x)$	$sin(x)$
$tan(x)$	$sec^2(x)$	$ln \mid sec(x) \mid$
$sec(x)$	$sec(x)tan(x)$	$ln \mid sec(x) + tan(x) \mid$
$cot(x)$	$sec^2(x)$	$ln \mid sin(x) \mid$
$cosec(x)$	$sec^2(x)$	$ln \mid cosec(x) - cot(x) \mid$

can be obtained

Tables (contd.)

Function	Derivative	Integral
$e^{\lambda x}$	$\lambda e^{\lambda x}$	$\dfrac{e^{\lambda x}}{\lambda}$
a^x	$a^x ln(a)$	$\dfrac{a^x}{ln(a)}$
$ln(x)$	$\dfrac{1}{x}$	$xln(x) - x$
$\dfrac{1}{\sqrt{a^2 - x^2}}$		$sin^{-1}\left(\dfrac{x}{a}\right)$
$\dfrac{1}{\sqrt{a^2 + x^2}}$		$tan^{-1}\left(\dfrac{x}{a}\right)$
$\dfrac{1}{a^2 - x^2}$		$\dfrac{1}{2a} ln\left\lvert\dfrac{x+a}{x-a}\right\rvert$

Examples

$$\int_1^2 x dx = \left[\frac{x^2}{2}\right]_1^2 = \frac{2^2}{2} - \frac{1^2}{2}$$

$$= \frac{4}{2} - \frac{1}{2} = \frac{3}{2}.$$

$f(x) = x$

$F(x) = x^2/2$

$F(2) - F(1)$

$$\int_0^\pi sin(x)dx = \left[-\cos(x)\right]_0^\pi = (-\cos(\pi)) - (-\cos(0))$$

$$= (-(-1)) - (-1)$$

$$= 1 + 1 = 2.$$

$$\int_0^\infty e^{-2x} dx = \lim_{b \to \infty} \int_0^b e^{-2x} dx$$

$$= \lim_{b \to \infty} \left[\frac{e^{-2x}}{-2}\right]_0^b = \lim_{b \to \infty} \frac{-1}{2}\left(e^{-bx} - e^0\right)$$

$$= \lim_{b \to \infty} \frac{1}{2} - e^{-bx} = \frac{1}{2}.$$

$$\int_a^\infty f(x)\,dx = \lim_{b \to \infty} \int_a^b f(x)\,dx$$

Recall : Integrals and Newton's theorem

The integral $\int f(x)dx$ of the function f defined on a domain D is a function F defined on the domain D such that $F'(x) = f(x)$ for all $x \in D$.

If f is continuous on the domain D which includes the interval $[a, b]$ and F is the integral of f, then the (definite) integral from a to b of f can be computed by

$$\int_a^b f(x)dx = F(b) - F(a).$$

Upshot : If one knows the integral of a continuous function, then one can use it to compute the area between the graph of the function and an interval on the X-axis.

Basic properties of integrals

- $\int cf(x)dx = c\int f(x)dx$
- $\int (f+g)(x)dx = \int f(x)dx + \int g(x)dx$
- $\int (f-g)(x)dx = \int f(x)dx - \int g(x)dx$
- Integration by parts : $\int (fg')(x)dx = (fg)(x) - \int (f'g)(x)dx$

$$\frac{d}{dx}(fg)(x) = f'(x)\,g(x) + g'(x)\,f(x) = (f'g)(x) + (g'f)(x)$$

$$\therefore \int (f'g + g'f)(x)\,dx = (fg)(x)$$

$$\therefore \int (f'g)(x)\,dx + \int (g'f)(x)\,dx = (fg)(x)$$

$$\therefore \int (g'f)(x)\,dx = (fg)(x) - \int (f'g)(x)\,dx.$$

Basic properties of definite integrals

- $\int_a^b cf(x)dx = c\int_a^b f(x)dx$
- $\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$
- $\int_a^b (f-g)(x)dx = \int_a^b f(x)dx - \int_a^b g(x)dx$
- $\int_b^a f(x)dx = -\int_a^b f(x)dx$
- For any $c \in \mathbb{R}$, $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$
- If $f(x) \geq g(x)$ for all but finitely many points on the interval $[a, b]$, then $\int_a^b f(x)dx \geq \int_a^b g(x)dx$

Examples

$$\int_1^3 x^2 - 4x + 2 \quad dx$$

$$= \int_1^3 x^2\,dx - 4\int_1^3 x\,dx + 2\int_1^3 1\,dx$$

$$= \left[\frac{x^3}{3}\right]_1^3 - 4\left[\frac{x^2}{2}\right]_1^3 + 2\left[x\right]_1^3$$

$$= \frac{1}{3}(3^3-1^3) - \frac{4}{2}(3^2-1^2) + 2(3-1)$$

$$= \frac{26}{3} - 16 + 4 = \frac{26}{3} - 12 = \frac{-10}{3}.$$

$$\int_{-2}^2 x^2 sin(x)dx$$

$$= \boxed{\int_{-2}^0 x^2 \sin(x)\,dx} + \int_0^2 x^2 \sin(x)\,dx$$

$$f(x) = x^2 \sin(x)$$

$$f(-x) = (-x)^2 \sin(-x)$$

$$= -x^2 \sin(x)$$

$$= \int_2^0 x^2 \sin(x)\,dx + \int_0^2 x^2 \sin(x)\,dx$$

$$= -\int_0^2 x^2 \sin(x)\,dx + \int_0^2 x^2 \sin(x)\,dx = 0.$$

Integration of piecewise defined functions

If f is defined piecewise on subintervals of $[a, b]$ then its definite integral from a to b can be computed by computing the definite integrals on each subinterval and adding them up.

Example : $f(x) = \begin{cases} x & \text{if } 0 \leq x \leq 1 \\ 3 - x & \text{if } 1 \leq x \leq 2 \end{cases}$

What is $\int_0^2 f(x)dx$?

$$\int_0^2 f(x)\,dx = \int_0^1 f(x)\,dx + \int_1^2 f(x)\,dx$$

$$= \int_0^1 x\,dx + \int_1^2 (3-x)\,dx$$

$$= \left[\frac{x^2}{2}\right]_0^1 + \left[3x - \frac{x^2}{2}\right]_1^2 = \frac{1}{2} + 3(2-1) - \frac{1}{2}(4-1)$$

$$= 2.$$

Integration by parts

$$\int_a^b (fg')(x)dx = \underline{(fg)(b) - (fg)(a)} - \int_a^b (f'g)(x)dx$$

Example : $\int_0^\infty 3xe^{-3x}dx = \lim_{b\to\infty} \underbrace{-be^{-3b}}_{\to 0} - \frac{1}{3}e^{-3b} + \frac{1}{3} = \frac{1}{3}.$

$f(x) = x$

$g'(x) = 3e^{-3x}$

$\Rightarrow g(x) = -e^{-3x}$

$$\int_0^b 3xe^{-3x}dx$$

$$= \underline{\left[-xe^{-3x}\right]_0^b} - \int_0^b 1 \times (-e^{-3x})dx$$

$$= -be^{-3b} - \left[\frac{e^{-3x}}{3}\right]_0^b = -be^{-3b} - \left(\frac{e^{-3b}}{3} - \frac{1}{3}\right)$$

$$= -be^{-3b} - \frac{1}{3}e^{-3b} + \frac{1}{3}.$$

Substitution

$$\int_a^b (f(g(x))g'(x)dx = \int_g (a)^g (b) f(u) du$$

$\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1.$

Example : $\int_0^a \sqrt{a^2 - u^2 du}$ $= \pi a^2/4.$

$u = a \sin(x)$ $g(x)$

Limits are : 0 to $\pi/2$.

$g(0) = 0$

$g(\pi/2) = a.$

$f(u) = \sqrt{a^2 - u^2}$

$f(g(x)) = \sqrt{a^2 - g(x)^2} = \sqrt{a^2 - a^2 \sin^2 x}$

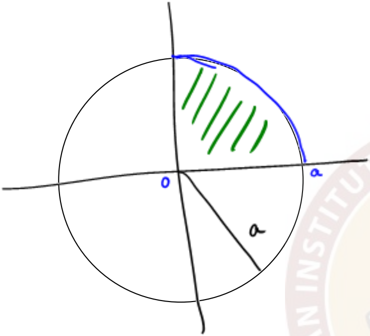
$$\int_0^{\pi/2} \sqrt{a^2 - a^2 \sin^2(x)} \; a\cos(x)\, dx$$

$$= \int_0^{\pi/2} a\cos(x)\; a\cos(x)\, dx = a^2 \int_0^{\pi/2} \cos^2(x)\, dx$$

$$= \int_0^{\pi/2} a^2 \cos^2(x)\, dx = a^2 \int_0^{\pi/2} \frac{1 + \cos(2x)}{2} dx$$

$$= a^2/2 \left[x + \frac{\sin(2x)}{2} \right]_0^{\pi/2} = \frac{a^2}{2}\left(\frac{\pi}{2} + \frac{\sin(\pi)}{2} \right)$$

Back to computing areas



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

$y=\sqrt{a^2-x^2}.$

$f(x)=\sqrt{a^2-x^2}$

$$\int_0^a f(x)\,dx = \frac{\pi a^2}{4}.$$

Area of the circle $= 4 \times \dfrac{\pi a^2}{4}$

$= \pi a^2.$

Graphs

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Mathematics for Data Science 1
Week 10

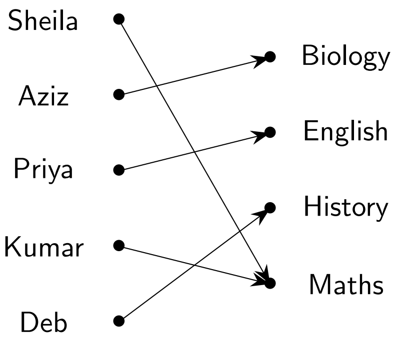
Visualizing relations

- Cartesian product $A \times B$

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

- A relation is a subset of $A \times B$
- Teachers and courses
  - T, set of teachers in a college
    C, set of courses being offered
  - $A \subseteq T \times C$ describes the allocation of teachers to courses
  - $A = \{(t, c) \mid (t, c) \in T \times C, t \text{ teaches } c\}$
- Introduce graphs formally

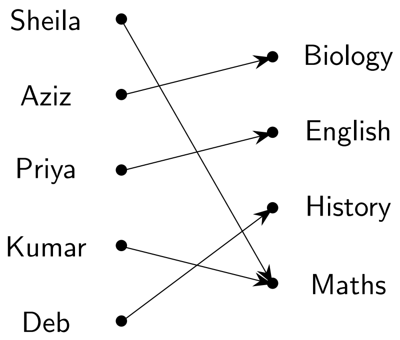
A relation as a graph



Graphs

- Graph: $G = (V, E)$
  - V is a set of vertices or nodes
    - One vertex, many vertices
  - E is a set of edges
  - $E \subseteq V \times V$ — binary relation
- Directed graph
  - $(v, v') \in E$ does not imply $(v', v) \in E$
  - The teacher-course graph is directed

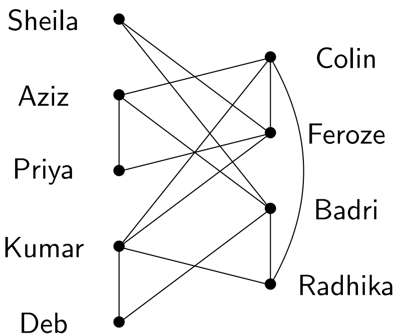
A relation as a graph



Graphs

- Graph: $G = (V, E)$
  - V is a set of vertices or nodes
    - One vertex, many vertices
  - E is a set of edges
  - $E \subseteq V \times V$ — binary relation
- Directed graph
  - $(v, v') \in E$ does not imply $(v', v) \in E$
  - The teacher-course graph is directed
- Undirected graph
  - $(v, v') \in E$ iff $(v', v) \in E$
  - Effectively (v, v'), (v', v) are the same edge
  - Friendship relation

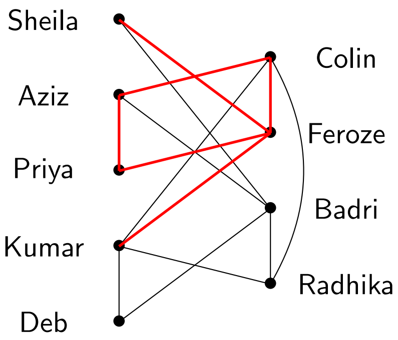
Friendship as a graph



Paths

- Priya needs some help that Radhika can provide. How will Priya come to know about this?
- Priya — Aziz — Badri — Radhika
- Priya — Feroze — Kumar — Radhika
- A path is a sequence of vertices $v_1, v_2, \ldots, v_k$ connected by edges
  - For $1 \leq i < k$, $(v_i, v_{i+1}) \in E$
- Normally, a path does not visit a vertex twice
  - Kumar — Feroze — Colin — Aziz — Priya — Feroze — Sheila
  - Such a sequence is usually called a walk

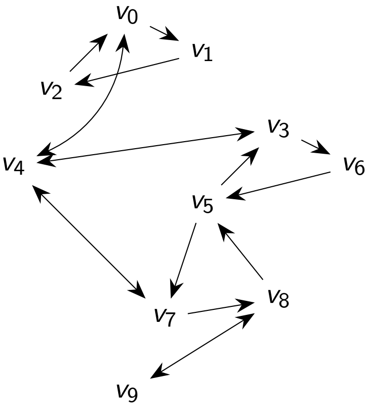
Friendship as a graph



Reachability

- Paths in directed graphs
- How can I fly from Madurai to Delhi?
  - Find a path from v_9 to v_0
- Vertex v is reachable from vertex u if there is a path from u to v
- Typical questions
  - Is v reachable from u?
  - What is the shortest path from u to v?
  - What are the vertices reachable from u?
  - Is the graph connected? Are all vertices reachable from each other?

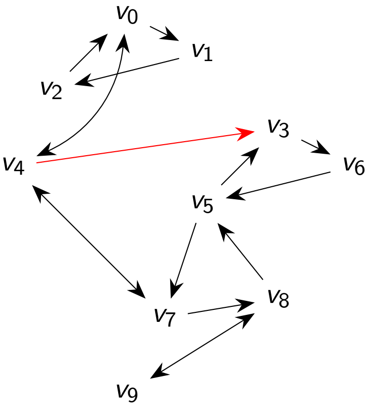
Airline routes



Reachability

- Paths in directed graphs
- How can I fly from Madurai to Delhi?
  - Find a path from v_9 to v_0
- Vertex v is reachable from vertex u if there is a path from u to v
- Typical questions
  - Is v reachable from u?
  - What is the shortest path from u to v?
  - What are the vertices reachable from u?
  - Is the graph connected? Are all vertices reachable from each other?

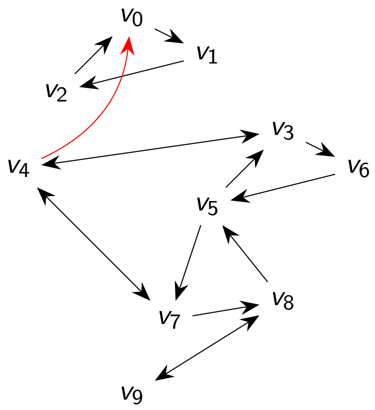
Airline routes



Reachability

- Paths in directed graphs
- How can I fly from Madurai to Delhi?
  - Find a path from v_9 to v_0
- Vertex v is reachable from vertex u if there is a path from u to v
- Typical questions
  - Is v reachable from u?
  - What is the shortest path from u to v?
  - What are the vertices reachable from u?
  - Is the graph connected? Are all vertices reachable from each other?

Airline routes



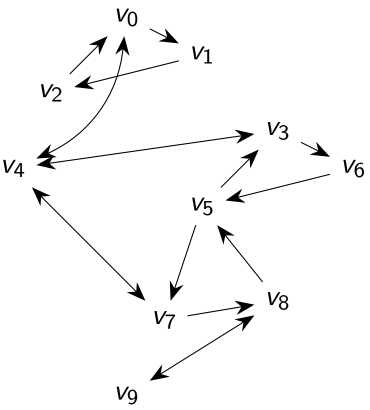
Summary

- A graph represents relationships between entities
  - Entities are vertices/nodes
  - Relationships are edges
- A graph may be directed or undirected
  - A is a parent of B — directed
  - A is a friend of B — undirected
- Paths are sequences of connected edges
- Reachability: is there a path from u to v?

Graphs

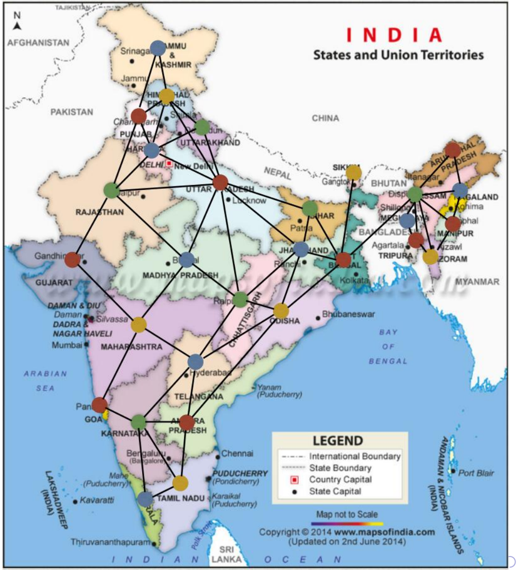
- Graph $G = (V, E)$
  - V — set of vertices
  - $E \subseteq V \times V$ — set of edges
- A path is a sequence of vertices $v_1, v_2, \ldots, v_k$ connected by edges
  - For $1 \leq i < k$, $(v_i, v_{i+1}) \in E$
- Vertex v is reachable from vertex u if there is a path from u to v
- What more can we do with graphs?

Airline routes



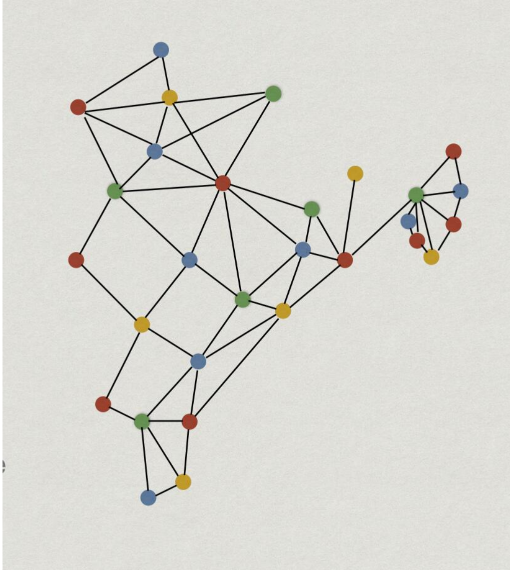
Map colouring

- Assign each state a colour
- States that share a border should be coloured differently
- How many colours do we need?
- Create a graph
  - Each state is a vertex
  - Connect states that share a border
- Assign colours to nodes so that endpoints of an edge have different colours



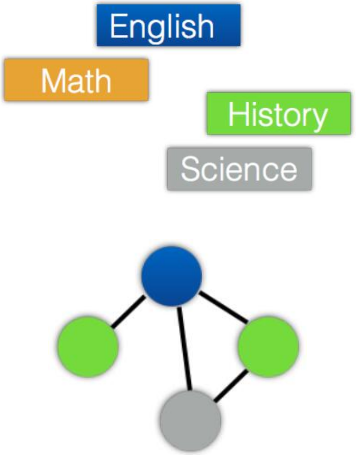
Map colouring

- Assign each state a colour
- States that share a border should be coloured differently
- How many colours do we need?
- Create a graph
  - Each state is a vertex
  - Connect states that share a border
- Assign colours to nodes so that endpoints of an edge have different colours
- Only need the underlying graph
- Abstraction: if we distort the graph, problem is unchanged



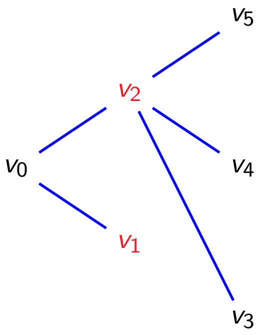
Graph colouring

- Graph $G = (V, E)$, set of colours C
- Colouring is a function $c : V \to C$ such that $(u, v) \in E \Rightarrow c(u) \neq c(v)$
- Given $G = (V, E)$, what is the smallest set of colours need to colour G
  - Four Colour Theorem For graphs derived from geographical maps, 4 colours suffice
  - Not all graphs are planar. General case? Why do we care?
- How many classrooms do we need?
  - Courses and timetable slots
  - Graph: Edges are overlaps in slots
  - Colours are classrooms



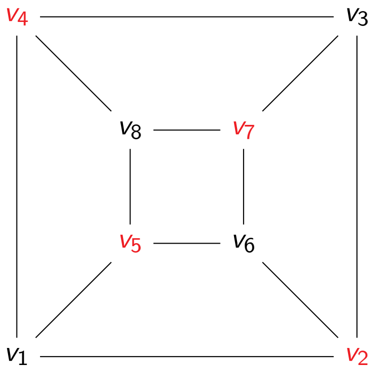
Vertex cover

- A hotel wants to install security cameras
  - All corridors are straight lines
  - Camera at the intersection of corridors can monitor all those corridor.
- Minimum number of cameras needed?
- Represent the floor plan as a graph
  - V — intersections of corridors
  - E — corridor segments connecting intersections
- Vertex cover
  - Marking v covers all edges from v
  - Mark smallest subset of V to cover all edges



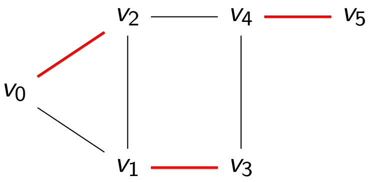
Independent set

- A dance school puts up group dances
  - Each dance has a set of dancers
  - Sets of dancers may overlap across dances
- Organizing a cultural programme
  - Each dancer performs at most once
  - Maximum number of dances possible?
- Represent the dances as a graph
  - V — dances
  - E — sets of dancers overlap
- Independent set
  - Subset of vertices such that no two are connected by an edge



Matching

- Class project can be done by one or two people
  - If two people, they must be friends
- Assume we have a graph describing friendships
- Find a good allocation of groups
- Matching
  - $G = (V, E)$, an undirected graph
  - A matching is a subset $M \subseteq E$ of mutually disjoint edges
- Find a maximal matching in G
- Is there a perfect matching, covering all vertices?



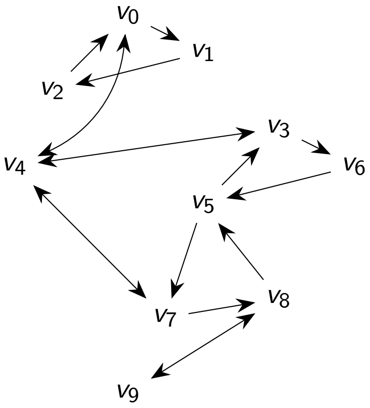
Summary

- Graphs are useful abstract representations for a wide range of problems
- Reachability and connectedness are not the only interesting problems we can solve on graphs
  - Graph colouring
  - Vertex cover
  - Independent set
  - Matching
  - ...

Working with graphs

- Graph $G = (V, E)$
  - V — set of vertices
  - $E \subseteq V \times V$ — set of edges
- A path is a sequence of vertices $v_1, v_2, \ldots, v_k$ connected by edges
  - For $1 \leq i < k$, $(v_i, v_{i+1}) \in E$
- Vertex v is reachable from vertex u if there is a path from u to v
- Looking at the picture of G, we can "see" that v_0 is reachable from v_9
- How do we represent this picture so that we can compute reachability?

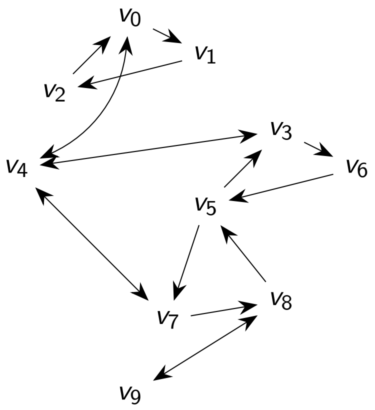
Airline routes



Adjacency matrix

- Let $|V| = n$
  - Assume $V = \{0, 1, \ldots, n-1\}$
  - Use a table to map actual vertex "names" to this set
- Edges are now pairs (i, j), where $0 \leq i, j < n$
  - Usually assume $i \neq j$, no self loops
- Adjacency matrix
  - Rows and columns numbered $\{0, 1, \ldots, n-1\}$
  - $A[i, j] = 1$ if $(i, j) \in E$

Airline routes

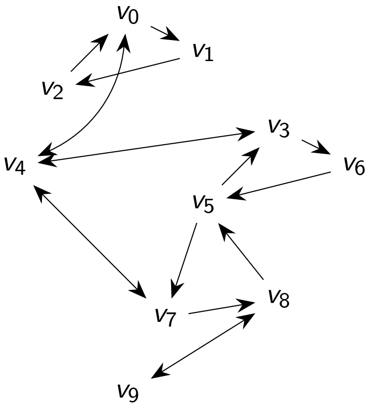


Adjacency matrix

- Adjacency matrix
  - Rows and columns numbered $\{0, 1, \ldots, n-1\}$
  - $A[i,j] = 1$ if $(i,j) \in E$

	0	1	2	3	4	5	6	7	8	9
0	0	1	0	0	1	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	1	0	0	0	1	0
8	0	0	0	0	0	1	0	0	0	1
9	0	0	0	0	0	0	0	0	1	0

Airline routes

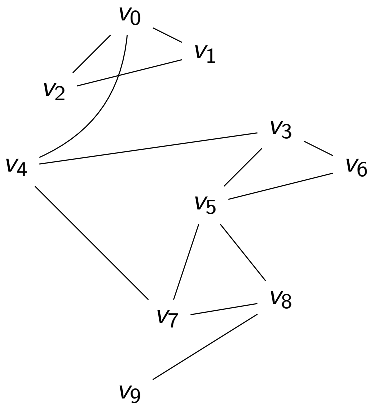


Adjacency matrix

- Undirected graph
  - $A[i,j] = 1$ iff $A[j,i] = 1$
  - Symmetric across main diagonal

	0	1	2	3	4	5	6	7	8	9
0	0	1	1	0	1	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
3	0	0	0	0	1	1	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	1	1	1	0
6	0	0	0	1	0	1	0	0	0	0
7	0	0	0	0	1	1	0	0	1	0
8	0	0	0	0	0	1	0	1	0	1
9	0	0	0	0	0	0	0	0	1	0

Airline routes, all routes bidirectional



Computing with the adjacency matrix

- Neighbours of i — column j with entry 1
  - Scan row i to identify neighbours of i
  - Neighbours of 6 are 3 and 5
- Directed graph
  - Rows represent outgoing edges
  - Columns represent incoming edges
- Degree of a vertex i
  - Number of edges incident on i

    $\text{degree}(6) = 2$
  - For directed graphs, outdegree and indegree

    $\text{indegree}(6) = 1$, $\text{outdegree}(6) = 1$

Directed airline routes

	0	1	2	3	4	5	6	7	8	9
0	0	1	0	0	1	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	1	0	0	0	1	0
8	0	0	0	0	0	1	0	0	0	1
9	0	0	0	0	0	0	0	0	1	0

Checking reachability

- Is Delhi (0) reachable from Madurai (9)?
- Mark 9 as reachable
- Mark each neighbour of 9 as reachable
- Systematically mark neighbours of marked vertices
- Stop when 0 becomes marked
- If marking process stops without target becoming marked, the target is unreachable

Undirected airline routes

	0	1	2	3	4	5	6	7	8	9
0	0	1	1	0	1	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
3	0	0	0	0	1	1	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	1	1	1	0
6	0	0	0	1	0	1	0	0	0	0
7	0	0	0	0	1	1	0	0	1	0
8	0	0	0	0	0	1	0	1	0	1
9	0	0	0	0	0	0	0	0	1	0

Checking reachability

- Mark source vertex as reachable
- Systematically mark neighbours of marked vertices
- Stop when target becomes marked
- Need a strategy to systematically explore marked neighbours
- Two primary strategies
  - Breadth first — propagate marks in "layers"
  - Depth first — explore a path till it dies out, then backtrack

Undirected airline routes

	0	1	2	3	4	5	6	7	8	9
0	0	1	1	0	1	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
3	0	0	0	0	1	1	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	1	1	1	0
6	0	0	0	1	0	1	0	0	0	0
7	0	0	0	0	1	1	0	0	1	0
8	0	0	0	0	0	1	0	1	0	1
9	0	0	0	0	0	0	0	0	1	0

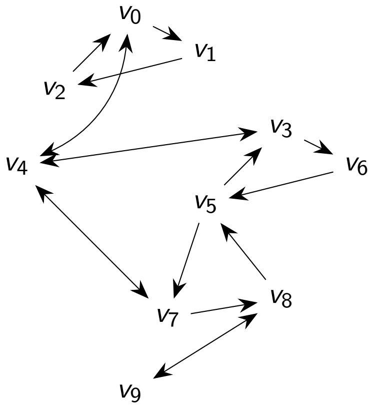
Adjacency lists

- Adjacency matrix has many 0's
  - Size is n^2, regardless of number of edges
  - Undirected graph: $|E| \leq n(n-1)/2$
  - Directed graph: $|E| \leq n(n-1)$
  - Typically $|E|$ much less than n^2
- Adjacency list
  - List of neighbours for each vertex

0	{1,4}
1	{2}
2	{0}
3	{4,6}
4	{0,3,7}

5	{3,7}
6	{5}
7	{4,8}
8	{5,9}
9	{8}

Airline routes



Comparing representations

- Adjacency list typically requires less space
- Is j a neighbour of i?
  - Check if $A[i,j] = 1$ in adjacency matrix
  - Scan all neighbours of i in adjacency list
- Which are the neighbours of i?
  - Scan all n entries in row i in adjacency matrix
  - Takes time proportional to (out)degree of i in adjacency list
- Choose representation depending on requirement

	0	1	2	3	4	5	6	7	8	9
0	0	1	0	0	1	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	1	0	0	0	1	0
8	0	0	0	0	0	1	0	0	0	1
9	0	0	0	0	0	0	0	0	1	0

0	{1,4}
1	{2}
2	{0}
3	{4,6}
4	{0,3,7}

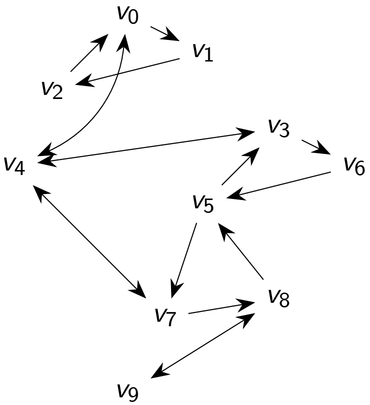
5	{3,7}
6	{5}
7	{4,8}
8	{5,9}
9	{8}

Summary

- To operate on graphs, we need to represent them
- Adjacency matrix
  - $n \times n$ matrix, $A[i,j] = 1$ iff $(i,j) \in E$
- Adjacency list
  - For each vertex i, list of neighbours of i
- Can systematically explore a graph using these representations
  - For reachability, propagate marking to all reachable vertices

Reachability in a graph

- Mark source vertex as reachable
- Systematically mark neighbours of marked vertices
- Stop when target becomes marked



Reachability in a graph

- Mark source vertex as reachable
- Systematically mark neighbours of marked vertices
- Stop when target becomes marked
- Choose an appropriate representation
  - Adjacency matrix
  - Adjacency list
- Strategies for systematic exploration
  - Breadth first — propagate marks in "layers"
  - Depth first — explore a path till it dies out, then backtrack

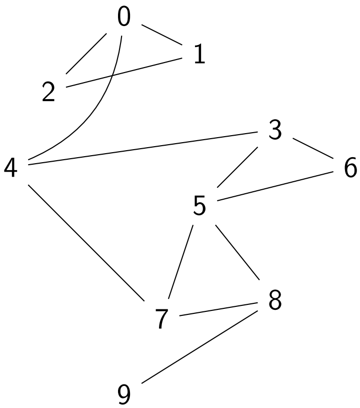
	0	1	2	3	4	5	6	7	8	9
0	0	1	0	0	1	0	0	0	0	0
1	0	0	1	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	1	0	0	0
4	1	0	0	1	0	0	0	1	0	0
5	0	0	0	1	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	1	0	0	0	1	0
8	0	0	0	0	0	1	0	0	0	1
9	0	0	0	0	0	0	0	0	1	0

0	{1,4}
1	{2}
2	{0}
3	{4,6}
4	{0,3,7}

5	{3,7}
6	{5}
7	{4,8}
8	{5,9}
9	{8}

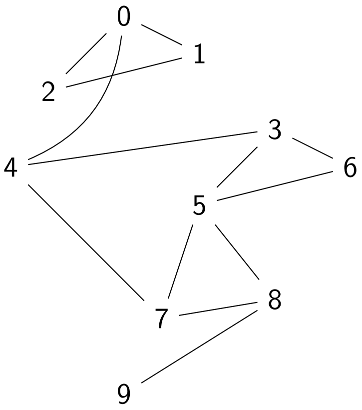
Breadth first search (BFS)

- Explore the graph level by level
  - First visit vertices one step away
  - Then two steps away
  - ...
- Each visited vertex has to be explored
  - Extend the search to its neighbours
  - Do this only once for each vertex!
- Maintain information about vertices
  - Which vertices have been visited already
  - Among these, which are yet to be explored



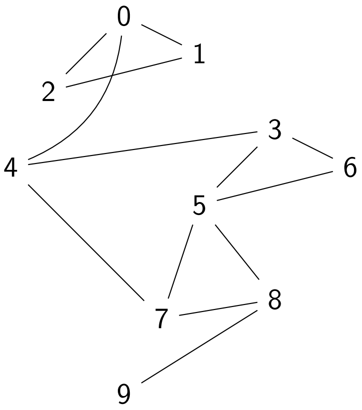
Breadth first search (BFS) ...

- Assume $V = \{0, 1, \dots, n-1\}$
- $\text{visited} : V \to \{\text{True}, \text{False}\}$ tells us whether $v \in V$ has been visited
  - Initially, $\text{visited}(v) = \text{False}$ for all $v \in V$
- Maintain a sequence of visited vertices yet be explored
  - A queue — first in, first out
  - Initially empty
- Exploring a vertex i
  - For each edge (i, j), if $\text{visited}(j)$ is False,
    - Set $\text{visited}(j)$ to True
    - Append j to the queue



Breadth first search (BFS) ...

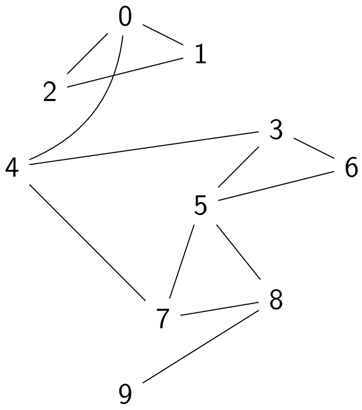
- Initially
  - $\text{visited}(v) = \text{False}$ for all $v \in V$
  - Queue of vertices to be explored is empty
- Start BFS from vertex j
  - Set $\text{visited}(j) = \text{True}$
  - Add j to the queue
- Remove and explore vertex i at head of queue
  - For each edge (i, j), if $\text{visited}(j)$ is False,
    - Set $\text{visited}(j)$ to True
    - Append j to the queue
- Stop when queue is empty



BFS from vertex 7

Visited	
0	False
1	False
2	False
3	False
4	False
5	False
6	False
7	False
8	False
9	False

To explore queue									

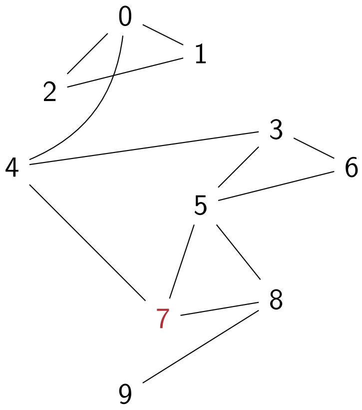


BFS from vertex 7

Visited	
0	False
1	False
2	False
3	False
4	False
5	False
6	False
7	True
8	False
9	False

To explore queue									
7									

- Mark 7 and add to queue

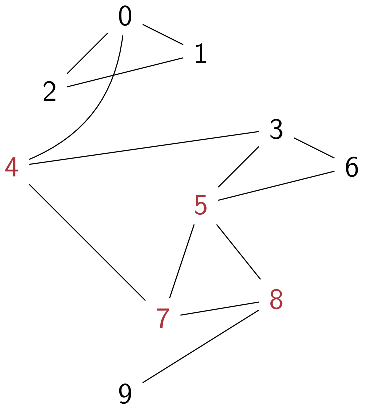


BFS from vertex 7

Visited	
0	False
1	False
2	False
3	False
4	True
5	True
6	False
7	True
8	True
9	False

To explore queue									
4	5	8							

- Mark 7 and add to queue
- Explore 7, visit {4,5,8}

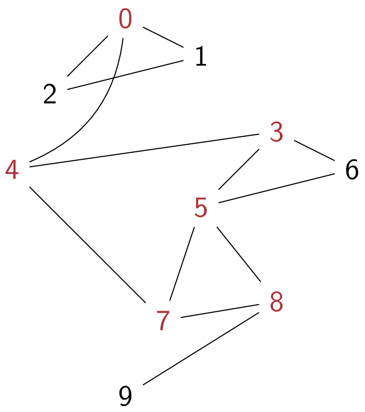


BFS from vertex 7

Visited	
0	True
1	False
2	False
3	True
4	True
5	True
6	False
7	True
8	True
9	False

To explore queue									
5	8	0	3						

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$

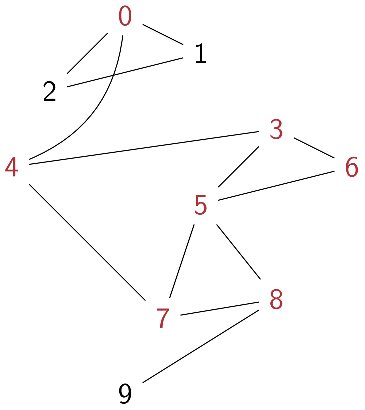


BFS from vertex 7

Visited	
0	True
1	False
2	False
3	True
4	True
5	True
6	True
7	True
8	True
9	False

To explore queue										
8	0	3	6							

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$

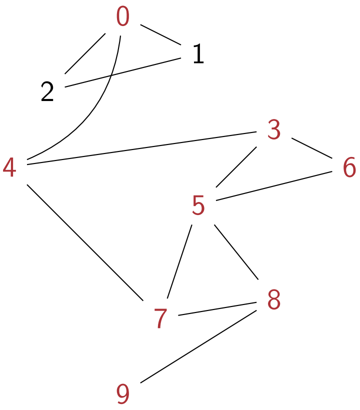


BFS from vertex 7

Visited	
0	True
1	False
2	False
3	True
4	True
5	True
6	True
7	True
8	True
9	True

To explore queue									
0	3	6	9						

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$

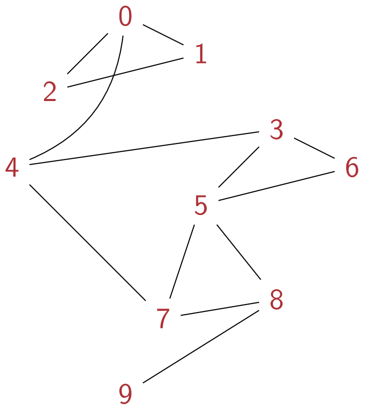


BFS from vertex 7

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

To explore queue									
3	6	9	1	2					

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$

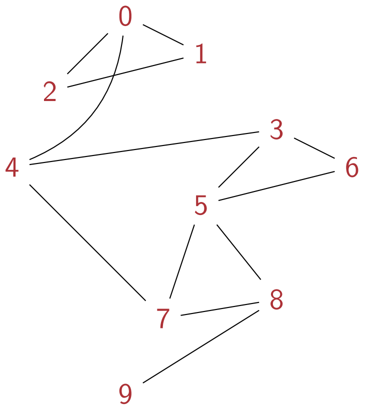


BFS from vertex 7

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

To explore queue									
6	9	1	2						

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3

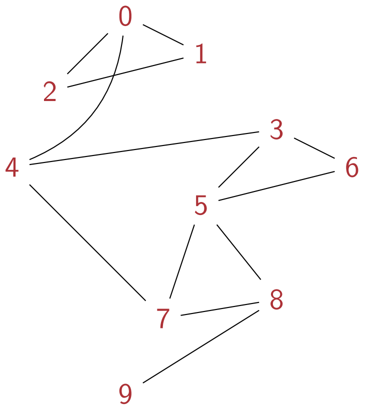


BFS from vertex 7

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

To explore queue										
9	1	2								

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6

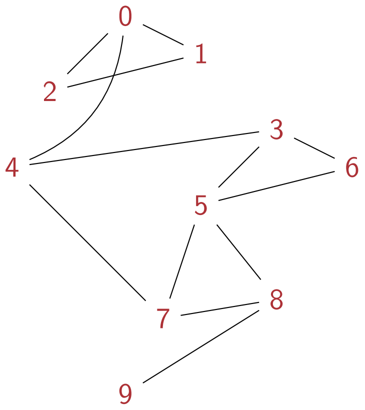


BFS from vertex 7

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

To explore queue										
1	2									

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9

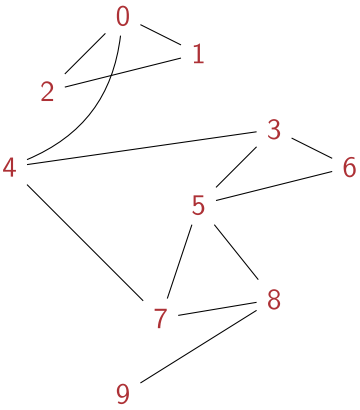


BFS from vertex 7

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

To explore queue									
2									

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9
- Explore 1

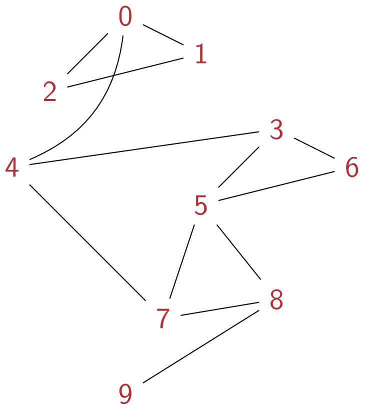


BFS from vertex 7

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

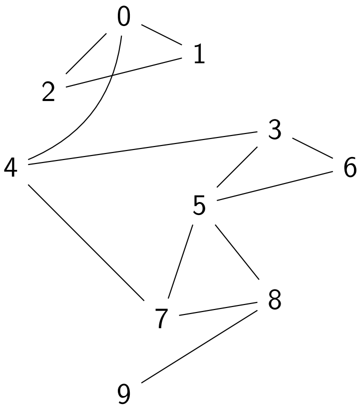
To explore queue									

- Mark 7 and add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9
- Explore 1
- Explore 2



Enhancing BFS to record paths

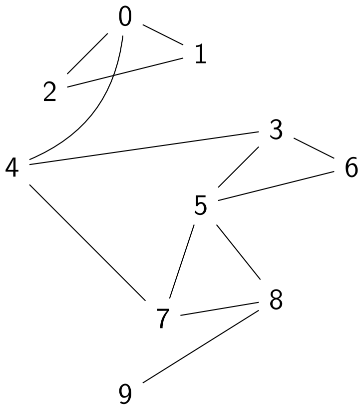
- If BFS from i sets $\text{visited}(j) = \text{True}$, we know that j is reachable from i
- How do we recover a path from i to j?
- $\text{visited}(j)$ was set to True when exploring some vertex k
- Record $\text{parent}(j) = k$
- From j, follow parent links to trace back a path to i



BFS from vertex 7 with parent information

	Visited	Parent
0	False	
1	False	
2	False	
3	False	
4	False	
5	False	
6	False	
7	False	
8	False	
9	False	

To explore queue									

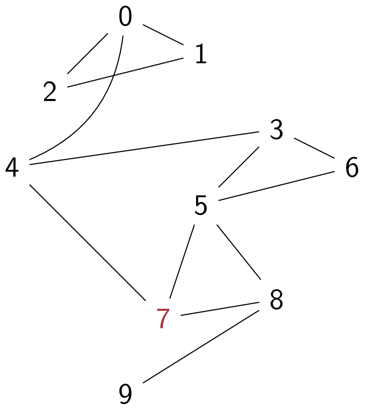


BFS from vertex 7 with parent information

	Visited	Parent
0	False	
1	False	
2	False	
3	False	
4	False	
5	False	
6	False	
7	True	
8	False	
9	False	

To explore queue										
7										

- Mark 7, add to queue

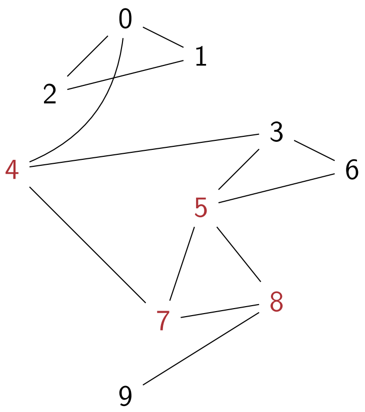


BFS from vertex 7 with parent information

	Visited	Parent
0	False	
1	False	
2	False	
3	False	
4	True	7
5	True	7
6	False	
7	True	
8	True	7
9	False	

To explore queue										
4	5	8								

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$

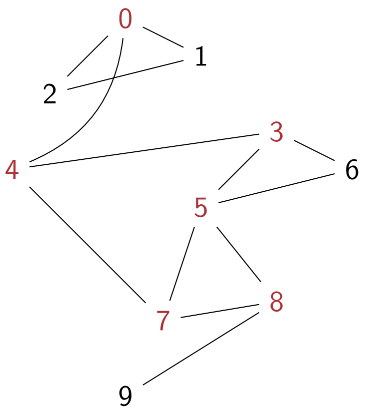


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	False	
2	False	
3	True	4
4	True	7
5	True	7
6	False	
7	True	
8	True	7
9	False	

To explore queue									
5	8	0	3						

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}

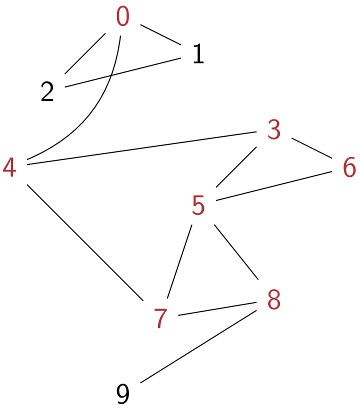


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	False	
2	False	
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	False	

To explore queue									
8	0	3	6						

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}

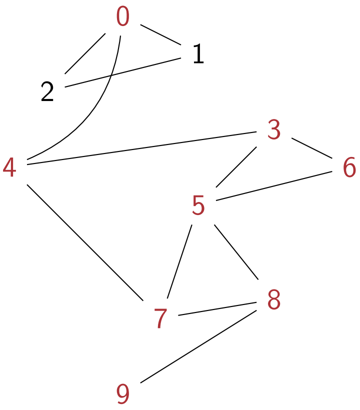


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	False	
2	False	
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

To explore queue									
0	3	6	9						

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}

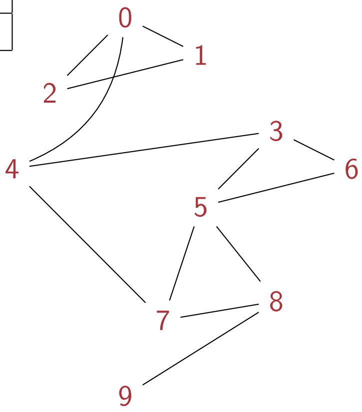


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

To explore queue									
3	6	9	1	2					

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}

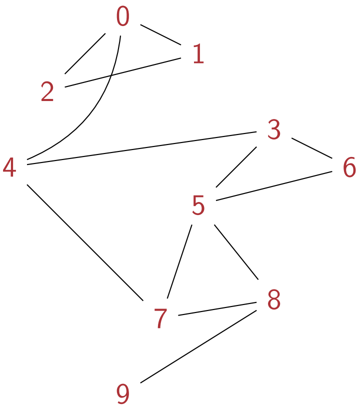


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

To explore queue									
6	9	1	2						

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3

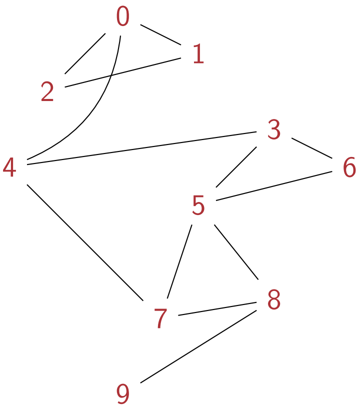


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

To explore queue										
9	1	2								

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3
- Explore 6

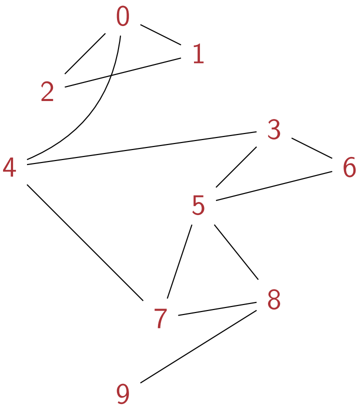


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

To explore queue										
1	2									

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3
- Explore 6
- Explore 9

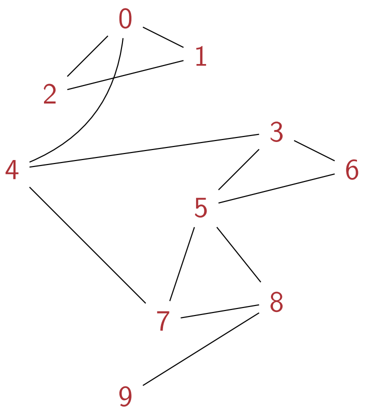


BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

To explore queue									
2									

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3
- Explore 6
- Explore 9
- Explore 1



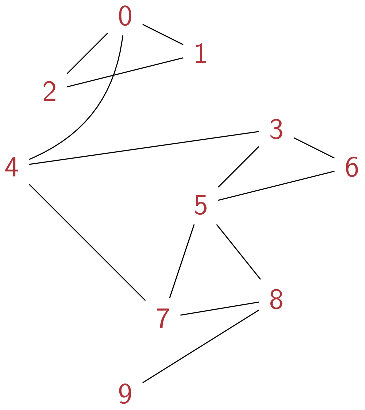
BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

Path from 7 to 6 is 7–5–6

To explore queue									

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9
- Explore 1
- Explore 2



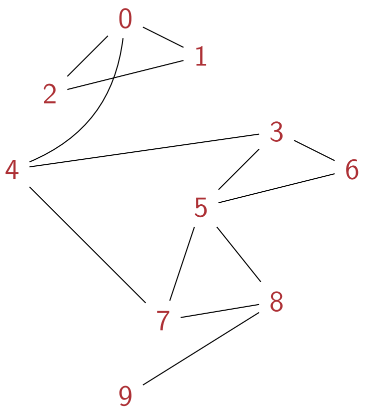
BFS from vertex 7 with parent information

	Visited	Parent
0	True	4
1	True	0
2	True	0
3	True	4
4	True	7
5	True	7
6	True	5
7	True	
8	True	7
9	True	8

Path from 7 to 2 is 7–4–0–2

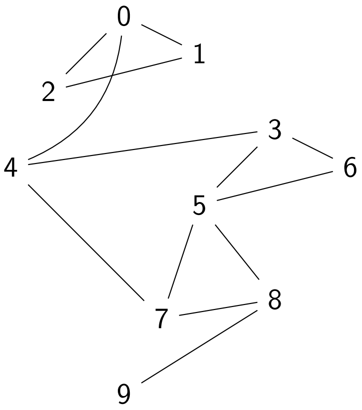
To explore queue									

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9
- Explore 1
- Explore 2



Enhancing BFS to record distance

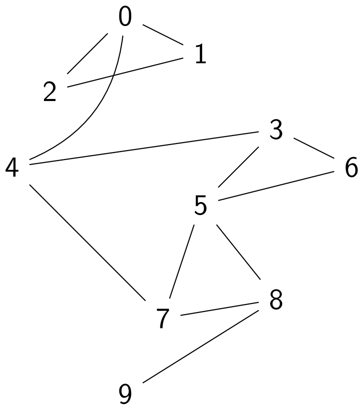
- BFS explores neighbours level by level
- By recording the level at which a vertex is visited, we get its distance from the source vertex
- Instead of $\text{visited}(j)$, maintain $\text{level})(j)$
- Initalize $\text{level}(j) = -1$ for all j
- Set $\text{level}(i) = 0$ for source vertex
- If we visit j from k, set $\text{level}(j)$ to $\text{level}(k) + 1$
- $\text{level}(j)$ is the length of the shortest path from the source vertex, in number of edges



BFS from vertex 7 with parent and distance information

	Level	Parent
0	-1	
1	-1	
2	-1	
3	-1	
4	-1	
5	-1	
6	-1	
7	-1	
8	-1	
9	-1	

To explore queue									

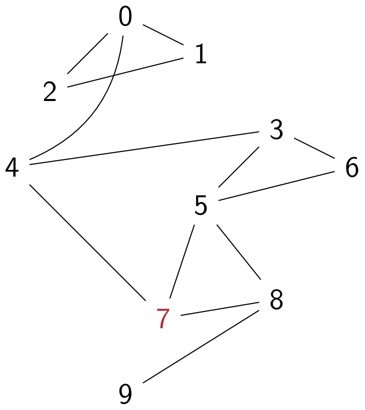


BFS from vertex 7 with parent and distance information

	Level	Parent
0	-1	
1	-1	
2	-1	
3	-1	
4	-1	
5	-1	
6	-1	
7	0	
8	-1	
9	-1	

To explore queue									
7									

- Mark 7, add to queue

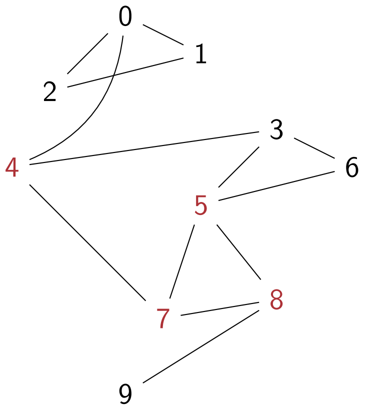


BFS from vertex 7 with parent and distance information

	Level	Parent
0	-1	
1	-1	
2	-1	
3	-1	
4	1	7
5	1	7
6	-1	
7	0	
8	1	7
9	-1	

To explore queue										
4	5	8								

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$

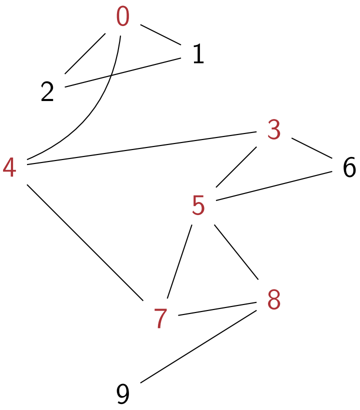


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	-1	
2	-1	
3	2	4
4	1	7
5	1	7
6	-1	
7	0	
8	1	7
9	-1	

To explore queue									
5	8	0	3						

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$

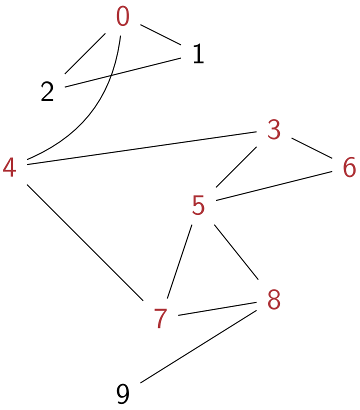


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	−1	
2	−1	
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	−1	

To explore queue									
8	0	3	6						

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$

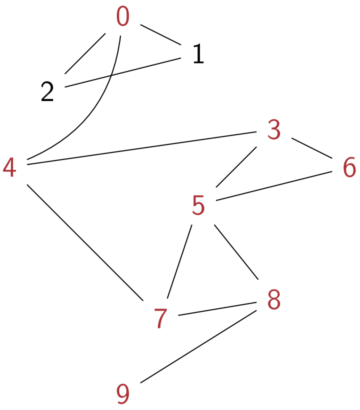


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	-1	
2	-1	
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue									
0	3	6	9						

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$

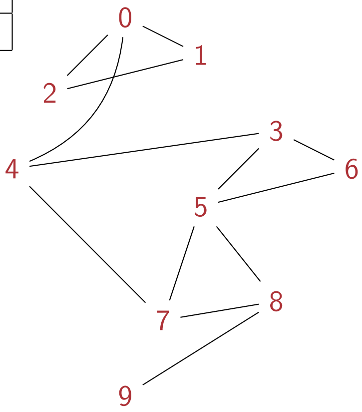


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	3	0
2	3	0
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue									
3	6	9	1	2					

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}

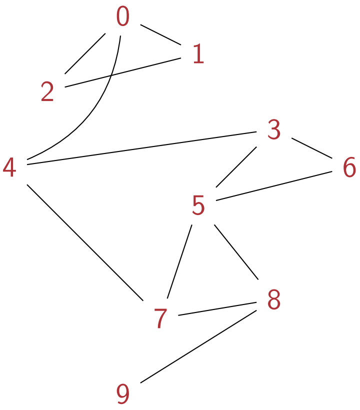


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	3	0
2	3	0
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue									
6	9	1	2						

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3

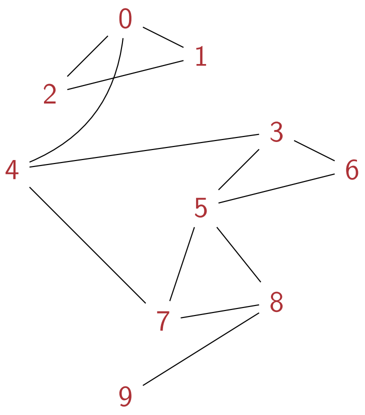


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	3	0
2	3	0
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue										
9	1	2								

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3
- Explore 6

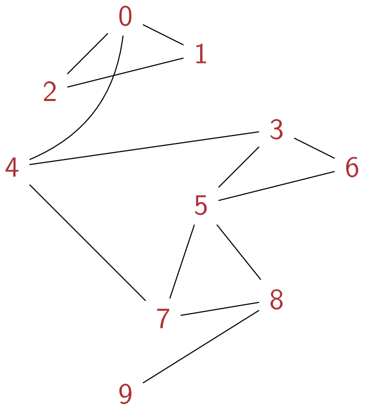


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	3	0
2	3	0
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue										
1	2									

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9

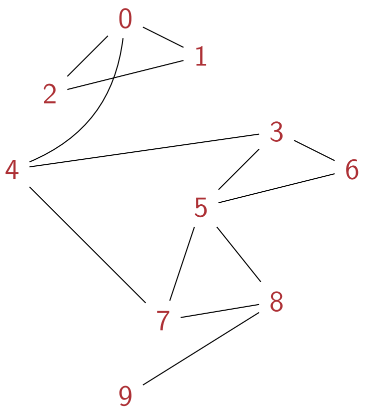


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	3	0
2	3	0
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue									
2									

- Mark 7, add to queue
- Explore 7, visit {4,5,8}
- Explore 4, visit {0,3}
- Explore 5, visit {6}
- Explore 8, visit {9}
- Explore 0, visit {1,2}
- Explore 3
- Explore 6
- Explore 9
- Explore 1

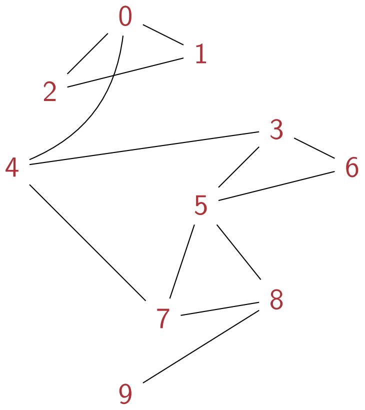


BFS from vertex 7 with parent and distance information

	Level	Parent
0	2	4
1	3	0
2	3	0
3	2	4
4	1	7
5	1	7
6	2	5
7	0	
8	1	7
9	2	8

To explore queue									

- Mark 7, add to queue
- Explore 7, visit $\{4,5,8\}$
- Explore 4, visit $\{0,3\}$
- Explore 5, visit $\{6\}$
- Explore 8, visit $\{9\}$
- Explore 0, visit $\{1,2\}$
- Explore 3
- Explore 6
- Explore 9
- Explore 1
- Explore 2

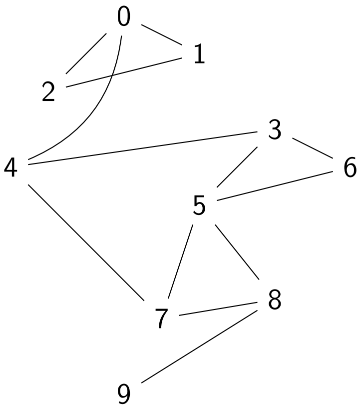


Summary

- Breadth first search is a systematic strategy to explore a graph, level by level
- Record which vertices have been visited
- Maintain visited but unexplored vertices in a queue
- Add parent information to recover the path to each reachable vertex
- Maintain level information to record length of the shortest path, in terms of number of edges
- In general, edges are labelled with a cost (distance, time, ticket price, ...)
- Will look at weighted graphs, where shortest paths are in terms of cost, not number of edges

Depth first search (DFS)

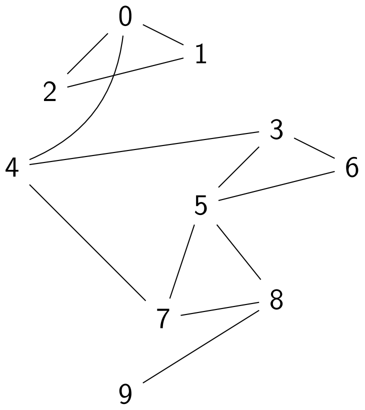
- Start from i, visit an unexplored neighbour j
- Suspend the exploration of i and explore j instead
- Continue till you reach a vertex with no unexplored neighbours
- Backtrack to nearest suspended vertex that still has an unexplored neighbour
- Suspended vertices are stored in a stack
  - Last in, first out
  - Most recently suspended is checked first



DFS from vertex 4

Visited	
0	False
1	False
2	False
3	False
4	False
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices									

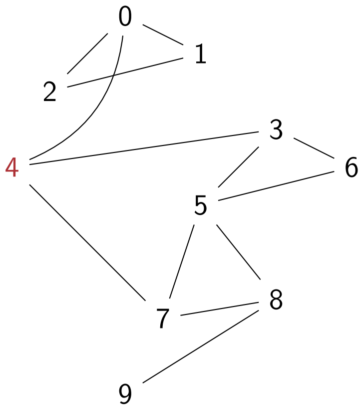


DFS from vertex 4

Visited	
0	False
1	False
2	False
3	False
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices									

- Mark 4,

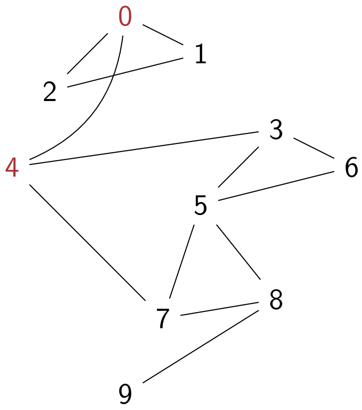


DFS from vertex 4

Visited	
0	True
1	False
2	False
3	False
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices									
4									

- Mark 4, Suspend 4, explore 0

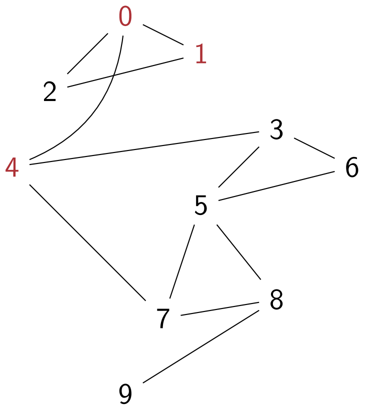


DFS from vertex 4

Visited	
0	True
1	True
2	False
3	False
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices								
4	0							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1

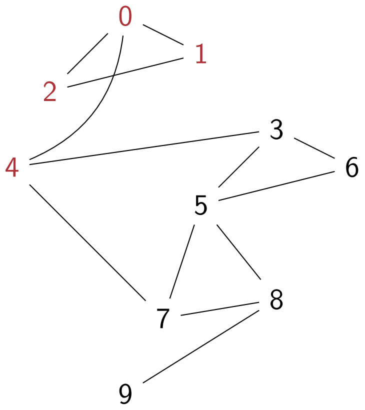


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	False
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices									
4	0	1							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2

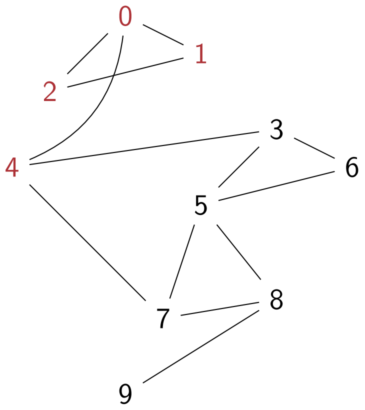


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	False
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices								
4	0							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1,

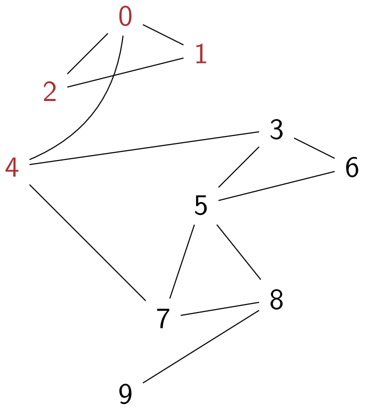


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	False
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices									
4									

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0,

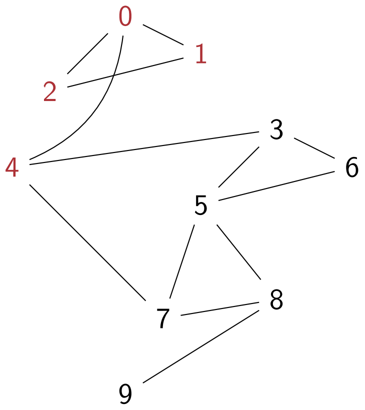


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4

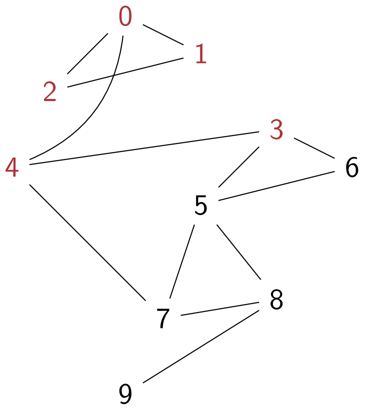


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	False
6	False
7	False
8	False
9	False

Stack of suspended vertices									
4									

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3

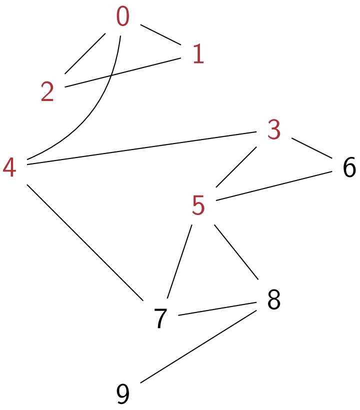


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	False
7	False
8	False
9	False

Stack of suspended vertices									
4	3								

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5

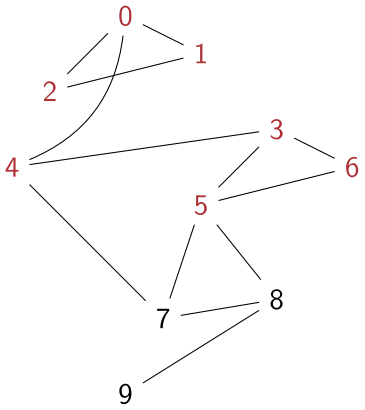


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	False
8	False
9	False

Stack of suspended vertices									
4	3	5							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6

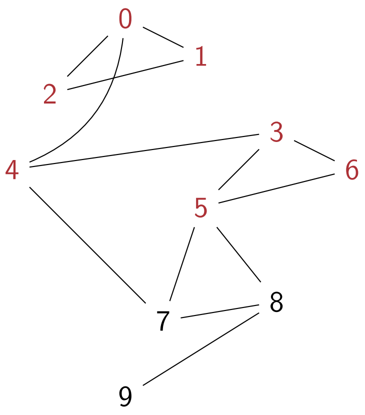


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	False
8	False
9	False

Stack of suspended vertices									
4	3								

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5,

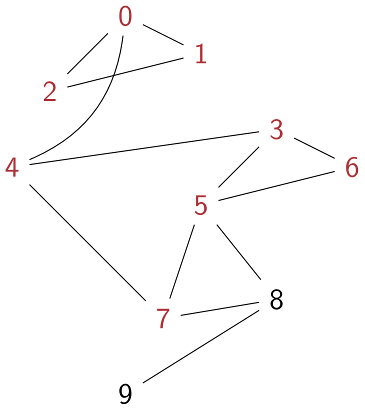


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	False
9	False

Stack of suspended vertices									
4	3	5							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7

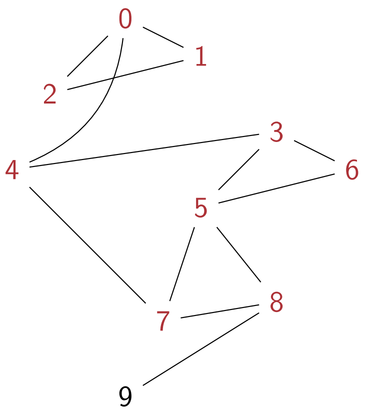


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	False

Stack of suspended vertices									
4	3	5	7						

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8

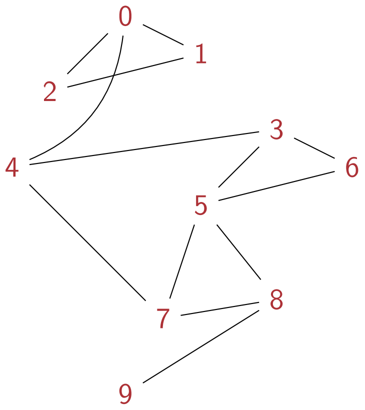


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

Stack of suspended vertices									
4	3	5	7	8					

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8
- Suspend 8, explore 9

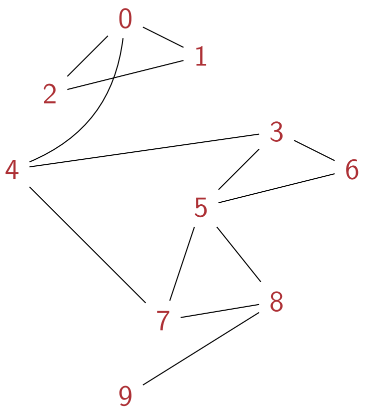


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

Stack of suspended vertices									
4	3	5	7						

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8
- Suspend 8, explore 9
- Backtrack to 8,

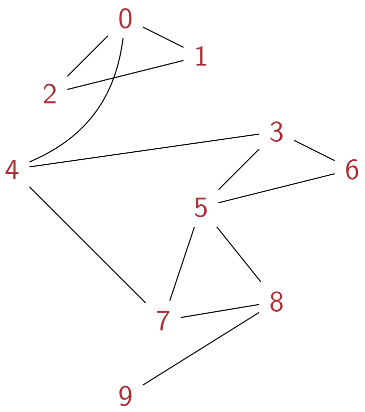


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

Stack of suspended vertices									
4	3	5							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8
- Suspend 8, explore 9
- Backtrack to 8, 7,

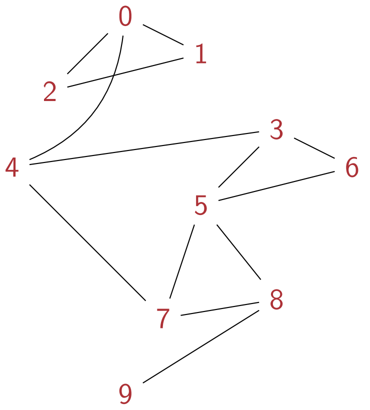


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

Stack of suspended vertices								
4	3							

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8
- Suspend 8, explore 9
- Backtrack to 8, 7, 5,

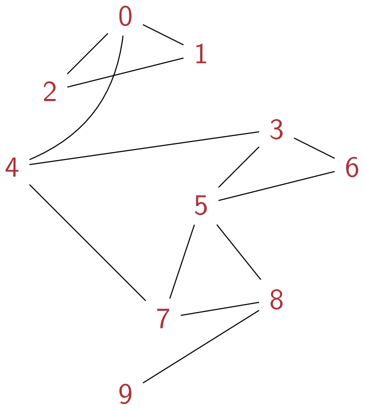


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

Stack of suspended vertices									
4									

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8
- Suspend 8, explore 9
- Backtrack to 8, 7, 5, 3,

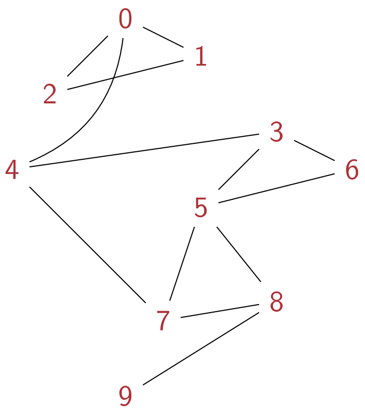


DFS from vertex 4

Visited	
0	True
1	True
2	True
3	True
4	True
5	True
6	True
7	True
8	True
9	True

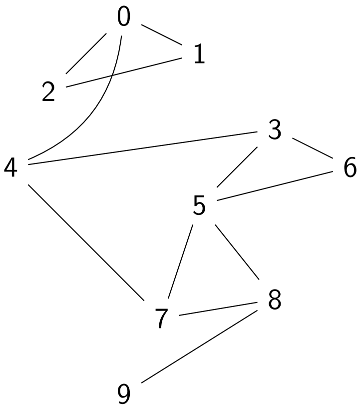
Stack of suspended vertices

- Mark 4, Suspend 4, explore 0
- suspend 0, explore 1
- Suspend 1, explore 2
- Backtrack to 1, 0, 4
- Suspend 4, explore 3
- Suspend 3, explore 5
- Suspend 5, explore 6
- Backtrack to 5, suspend 5, explore 7
- Suspend 7, explore 8
- Suspend 8, explore 9
- Backtrack to 8, 7, 5, 3, 4



Depth first search (DFS)

- Paths discovered by BFS are not shortest paths, unlike BFS
- Useful features can be found by recording the order in which DFS visits vertices
- DFS numbering — maintain a counter
  - Increment and record counter value each time you start and finish exploring a vertex
- DFS numbering can be used to
  - Find cut vertices (deleting vertex disconnects graph)
  - Find bridges (deleting edge disconnects graph)

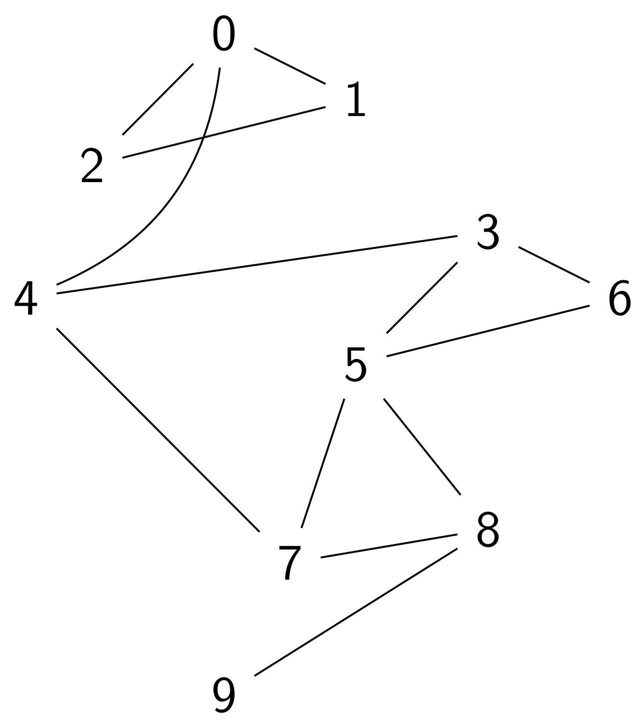


Summary

- DFS is another systematic strategy to explore a graph
- DFS uses a stack to suspend exploration and move to unexplored neighbours
- DFS numbering can be used to discover many facts about graphs

BFS and DFS

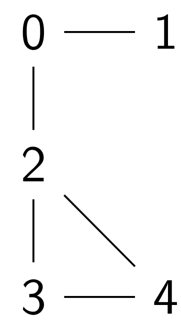
- BFS and DFS systematically compute reachability in graphs
- BFS works level by level
  - Discovers shortest paths in terms of number of edges
- DFS explores a vertex as soon as it is visited neighbours
  - Suspend a vertex while exploring its neighbours
  - DFS numbering describes the order in which vertices are explored
- Beyond reachability, what can we find out about a graph using BFS/DFS?



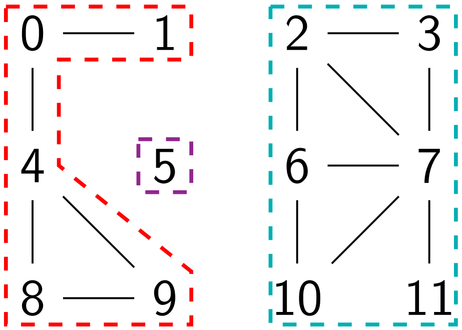
Connectivity

- An undirected graph is connected if every vertex is reachable from every other vertex
- In a disconnected graph, we can identify the connected components
  - Maximal subsets of vertices that are connected
  - Isolated vertices are trivial components

Connected Graph



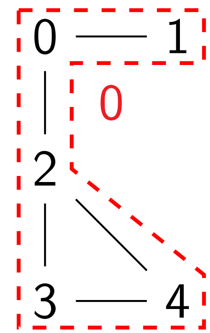
Disconnected Graph



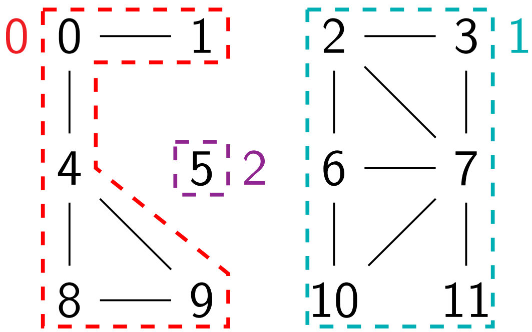
Identifying connected components

- Assign each vertex a component number
- Start BFS/DFS from vertex 0
  - Initialize component number to 0
  - All visited nodes form a connected component
  - Assign each visited node component number 0
- Pick smallest unvisited node j
  - Increment component number to 1
  - Run BFS/DFS from node j
  - Assign each visited node component number 1
- Repeat until all nodes are visited

Connected Graph



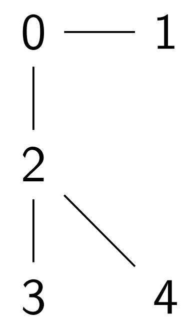
Disconnected Graph



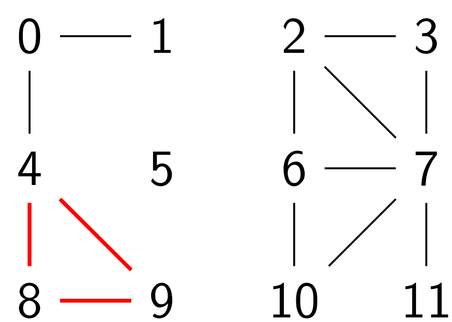
Detecting cycles

- A cycle is a path (technically, a walk) that starts and ends at the same vertex
  - $4 - 8 - 9 - 4$ is a cycle
  - Cycle may repeat a vertex: $2 - 3 - 7 - 10 - 6 - 7 - 2$
  - Cycle should not repeat edges: $i - j - i$ is not a cycle, e.g., $2 - 4 - 2$
  - Simple cycle — only repeated vertices are start and end
- A graph is acyclic if it has no cycles

Acyclic Graph



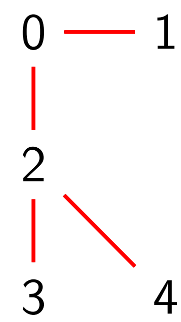
Graph with cycles



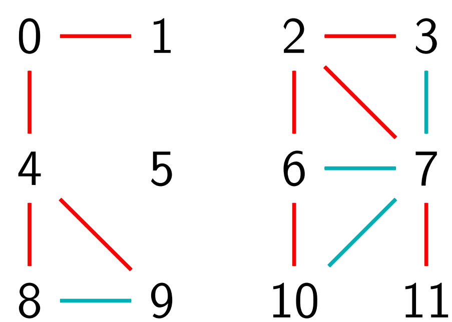
BFS tree

- A tree is a minimally connected graph
- Edges explored by BFS form a tree
  - Technically, one tree per component
  - Collection of trees is a forest
- Facts about trees
  - A tree on n vertices has $n-1$ edges
  - A tree is acyclic
- Any non-tree edge creates a cycle
  - Detect cycles by searching for non-tree edges

Acyclic Graph

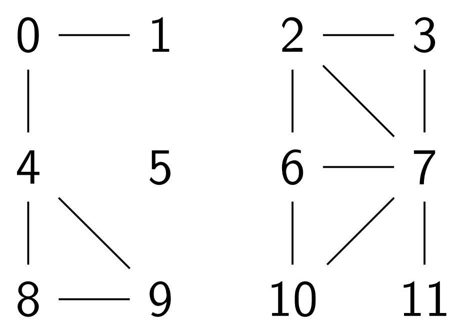


Graph with cycles

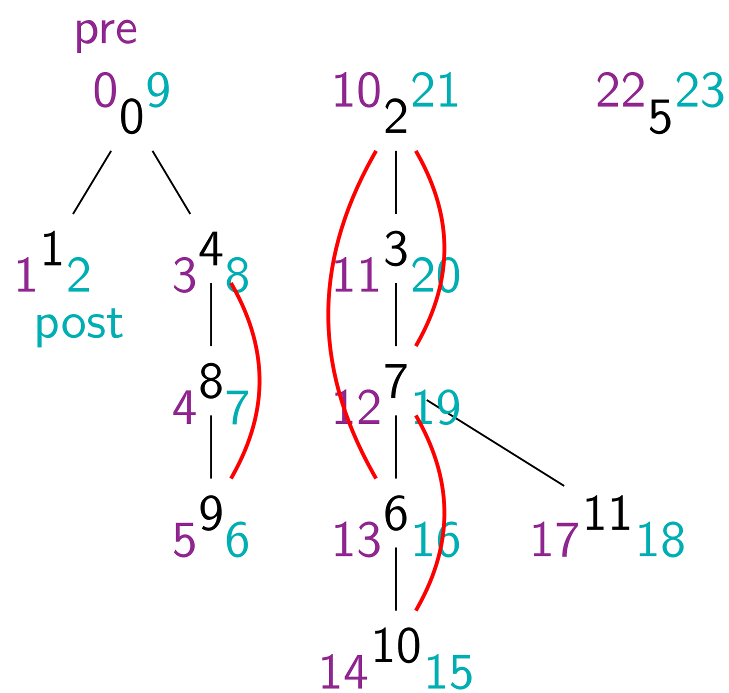


DFS tree

- Maintain a DFS counter, initially 0
- Increment counter each time we start and finish exploring a node
- Each vertex is assigned an entry number (pre) and exit number (post)

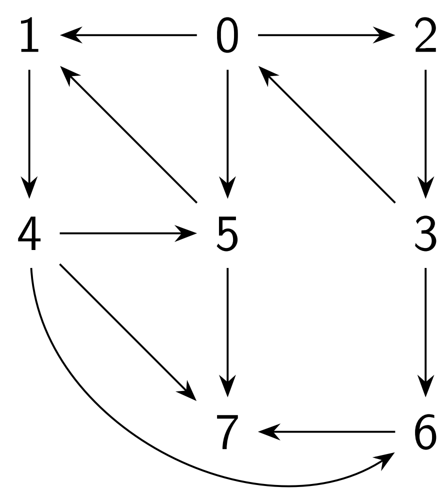


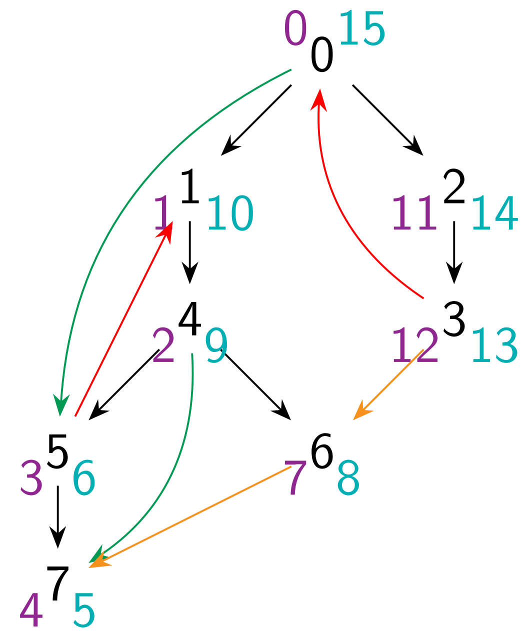
- As before, non-tree edges generate cycles



Directed cycles

- In a directed graph, a cycle must follow same direction
  - $0 \to 2 \to 3 \to 0$ is a cycle
  - $0 \to 5 \to 1 \leftarrow 0$ is not
- Tree edges
- Different types of non-tree edges
  - Forward edges
  - Back edges
  - Cross edges





- Only back edges correspond to cycles

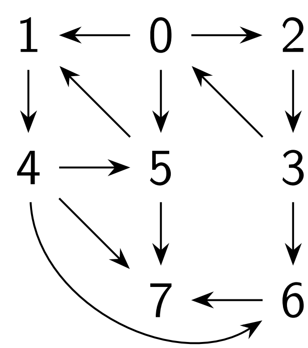
Classifying non-tree edges in directed graphs

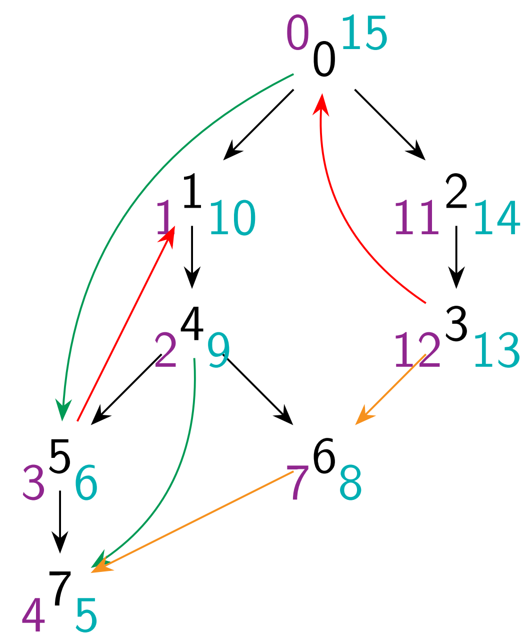
- Use pre/post numbers
- Tree edge/forward edge (u, v)

  Interval $[\text{pre}(u), \text{post}(u)]$ contains $[\text{pre}(v), \text{post}(v)]$
- Back edge (u, v)

  Interval $[\text{pre}(v), \text{post}(v)]$ contains $[\text{pre}(u), \text{post}(u)]$
- Cross edge (u, v)

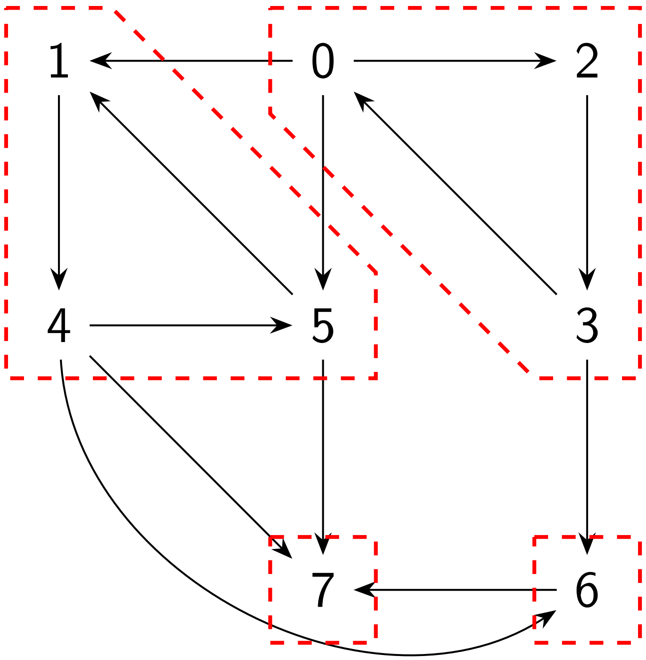
  Intervals $[\text{pre}(u), \text{post}(u)]$ and $[\text{pre}(v), \text{post}(v)]$ are disjoint





Connectivity in directed graphs

- Take directions into account
- Vertices i and j are strongly connected if there is a path from i to j and a path from j to i
- Directed graphs can be decomposed into strongly connected components (SCCs)
  - Within an SCC, each pair of vertices is strongly connected
- DFS numbering can be used to compute SCCs

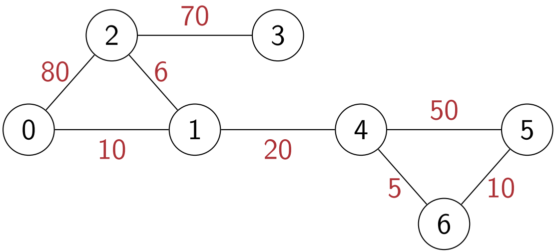


Summary

- BFS and DFS can be used to identify connected components in an undirected graph
  - BFS and DFS identify an underlying tree, non-tree edges generate cycles
- In a directed graph, non-tree edges can be forward / back / cross
  - Only back edges generate cycles
  - Classify non-tree edges using DFS numbering
- Directed graphs decompose into strongly connected components
  - DFS numbering can be used to compute SCC decomposition
- DFS numbering can also be used to identify other features such as articulation points (cut vertices) and bridges (cut edges)
- Directed acyclic graphs are useful for representing dependencies
  - Given course prerequisites, find a valid sequence to complete a programme

Shortest paths

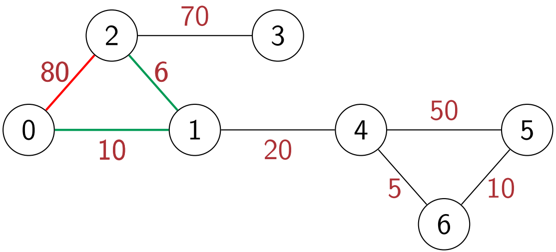
- Recall that BFS explores a graph level by level
- BFS computes shortest path, in terms of number of edges, to every reachable vertex
- May assign values to edges
  - Cost, time, distance, . . .
  - Weighted graph
- $G = (V, E)$, $W : E \to \mathbb{R}$
- Adjacency matrix: record the weight wherever there is an edge, 0 if no edge



	0	1	2	3	4	5	6
0	0	10	80	0	0	0	0
1	10	0	6	0	20	0	0
2	80	6	0	70	0	0	0
3	0	0	70	0	0	0	0
4	0	20	0	0	0	50	5
5	0	0	0	0	50	0	10
6	0	0	0	0	5	10	0

Shortest paths in weighted graphs

- BFS computes shortest path, in terms of number of edges, to every reachable vertex
- In a weighted graph, add up the weights along a path
- Weighted shortest path need not have minimum number of edges
  - Shortest path from 0 to 2 is via 1



Shortest path problems

Single source shortest paths

- Find shortest paths from a fixed vertex to every other vertex
- Transport finished product from factory (single source) to all retail outlets
- Courier company delivers items from distribution centre (single source) to addressees

All pairs shortest paths

- Find shortest paths between every pair of vertices i and j
- Optimal airline, railway, road routes between cities

Negative edge weights

Negative edge weights

- Can negative edge weights be meaningful?
- Taxi driver trying to head home at the end of the day
  - Roads with few customers, drive empty (positive weight)
  - Roads with many customers, make profit (negative weight)
  - Find a route toward home that minimizes the cost

Negative cycles

- A negative cycle is one whose weight is negative
  - Sum of the weights of edges that make up the cycle
- By repeatedly traversing a negative cycle, total cost keeps decreasing
- If a graph has a negative cycle, shortest paths are not defined
- Without negative cycles, we can compute shortest paths even if some weights are negative

Summary

- In a weighted graph, each edge has a cost
  - Entries in adjacency matrix capture edge weights
- Length of a path is the sum of the weights
  - Shortest path in a weighted graph need not be minimum in terms of number of edges
- Different shortest path problems
  - Single source — from one designated vertex to all others
  - All-pairs — between every pair of vertices
- Negative edge weights
  - Should not have negative cycles
  - Without negative cycles, shortest paths still well defined

Single Source Shortest Paths

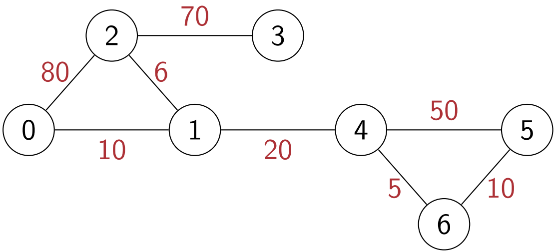
Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Mathematics for Data Science 1
Week 12

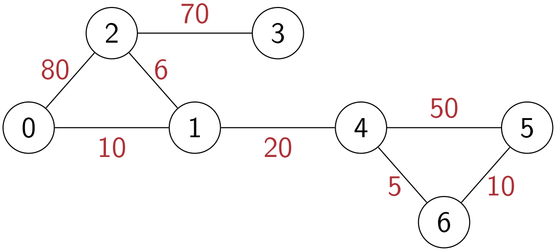
Single source shortest paths

- Weighted graph:
  - $G = (V, E)$
  - $W : E \to \mathbb{R}$
- Single source shortest paths
  - Find shortest paths from a fixed vertex to every other vertex
- Assume, for now, that edge weights are all non-negative

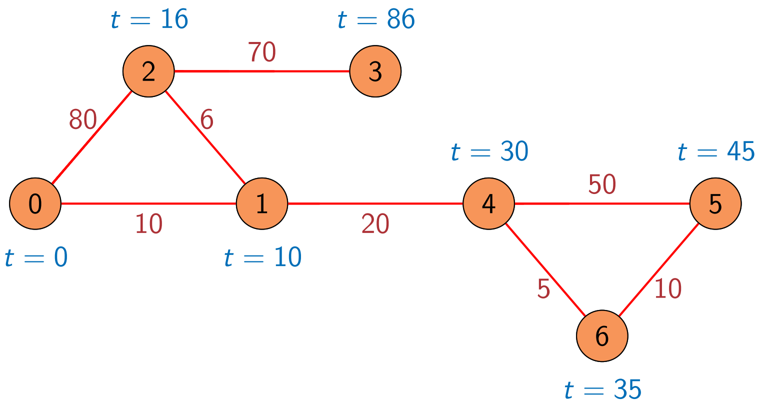


Single source shortest paths

- Compute shortest paths from 0 to all other vertices
- Imagine vertices are oil depots, edges are pipelines
- Set fire to oil depot at vertex 0
- Fire travels at uniform speed along each pipeline
- First oil depot to catch fire after 0 is nearest vertex
- Next oil depot is second nearest vertex
- ...

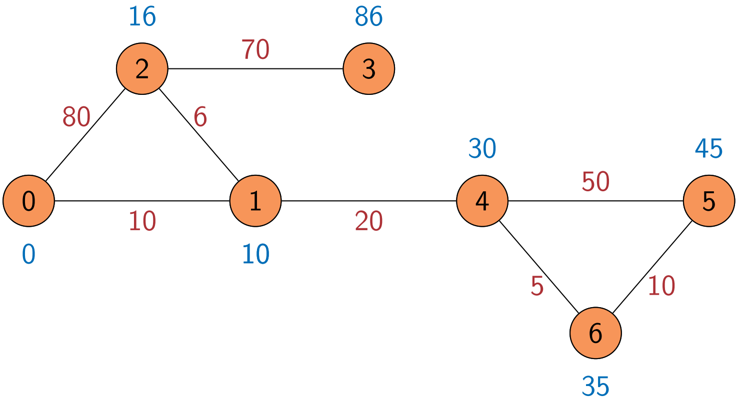


- Set fire to oil depot at vertex 0
- Fire travels at uniform speed along each pipeline
- First oil depot to catch fire after 0 is nearest vertex
- Next oil depot is second nearest vertex
- ...



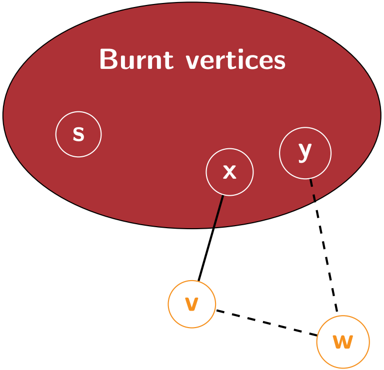
Single source shortest paths

- Compute expected burn time for each vertex
- Each time a new vertex burns, update the expected burn times of its neighbours
- Algorithm due to Edsger W Dikjstra



Dijkstra's algorithm: Proof of correctness

- Each new shortest path we discover extends an earlier one
- By induction, assume we have found shortest paths to all vertices already burnt
- Next vertex to burn is **v**, via **x**
- Cannot find a shorter path later from **y** to **v** via **w**
  - Burn time of **w** $\geq$ burn time of **v**
  - Edge from **w** to **v** has weight ≥ 0
- This argument breaks down if edge (**w**,**v**) can have negative weight
  - Can't use Dijkstra's algorithm with negative edge weights



Summary

- Dijkstra's algorithm computes single source shortest paths
- Use fire analogy
  - Keep track of expected burn times for each vertex
  - Update burn times of neighbours each time a vertex burns
- Correctness requires edge weights to be non-negative

Single Source Shortest Paths with Negative Weights

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Mathematics for Data Science 1
Week 12

Dijkstra's algorithm

- Recall the burning pipeline analogy
- We keep track of the following
  - The vertices that have been burnt
  - The expected burn time of vertices
- Initially
  - No vertex is burnt
  - Expected burn time of source vertex is 0
  - Expected burn time of rest is ∞
- While there are vertices yet to burn
  - Pick unburnt vertex with minimum expected burn time, mark it as burnt
  - Update the expected burn time of its neighbours

Initialization (assume source vertex 0)

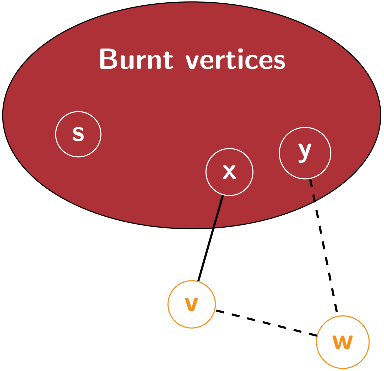
- $B(i) = \text{False}$, for $0 \leq i < n$
  - $UB = \{k \mid B(k) = \text{False}\}$
- $EBT(i) = \begin{cases} 0, & \text{if } i = 0 \\ \infty, & \text{otherwise} \end{cases}$

Update, if $UB \neq \emptyset$

- Let $j \in UB$ such that $EBT(j) \leq EBT(k)$ for all $k \in UB$
- Update $B(j) = \text{True}$, $UB = UB \setminus \{j\}$
- For each $(j, k) \in E$ such that $k \in UB$, $EBT(k) = \min(\; EBT(k),\; EBT(j) + W(j, k)\;)$

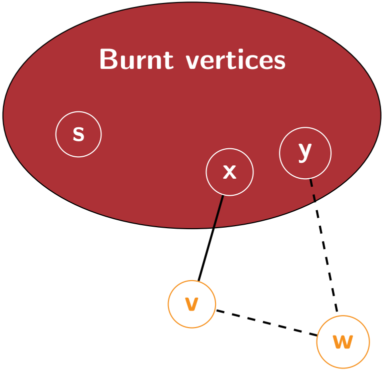
Correctness requires non-negative edge weights

- Each new shortest path we discover extends an earlier one
- By induction, assume we have found shortest paths to all vertices already burnt
- Next vertex to burn is **v**, via **x**
- Cannot find a shorter path later from **y** to **v** via **w**
  - Burn time of **w** $\geq$ burn time of **v**
  - Edge from **w** to **v** has weight ≥ 0
- This argument breaks down if edge (**w**,**v**) can have negative weight



Extending to negative edge weights

- The difficulty with negative edge weights is that we stop updating the burn time once a vertex is burnt
- What if we allow updates even after a vertex is burnt?
- Recall, negative edge weights are allowed, but no negative cycles
- Going around a cycle can only add to the length
- Shortest route to every vertex is a path, no loops



Extending to negative edge weights

- Suppose minimum weight path from 0 to k is

  $0 \xrightarrow{w_1} j_1 \xrightarrow{w_2} j_2 \xrightarrow{w_3} \cdots \xrightarrow{w_{\ell-1}} j_{\ell-1} \xrightarrow{w_\ell} k$

  - Need not be minimum in terms of number of edges
- Every prefix of this path must itself be a minimum weight path
  - $0 \xrightarrow{w_1} j_1$
  - $0 \xrightarrow{w_1} j_1 \xrightarrow{w_2} j_2$
  - ...
  - $0 \xrightarrow{w_1} j_1 \xrightarrow{w_2} j_2 \xrightarrow{w_3} \cdots \xrightarrow{w_{\ell-1}} j_{\ell-1}$
- Once we discover shortest path to $j_{\ell-1}$, next update will fix shortest path to k
- Repeatedly update shortest distance to each vertex based on shortest distance to its neighbours
  - Update cannot push this distance below actual shortest distance
- After ℓ updates, all shortest paths using $\leq \ell$ edges have stabilized
  - Minimum weight path to any node has at most $n-1$ edges
  - After $n-1$ updates, all shortest paths have stabilized

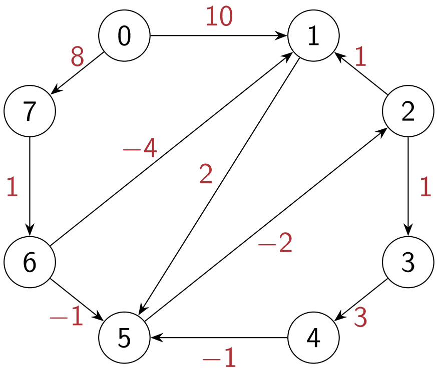
Bellman-Ford Algorithm

Initialization (source vertex 0)

- $D(j)$: minimum distance known so far to vertex j
- $D(j) = \begin{cases} 0, & \text{if } j = 0 \\ \infty, & \text{otherwise} \end{cases}$

Repeat $n{-}1$ times

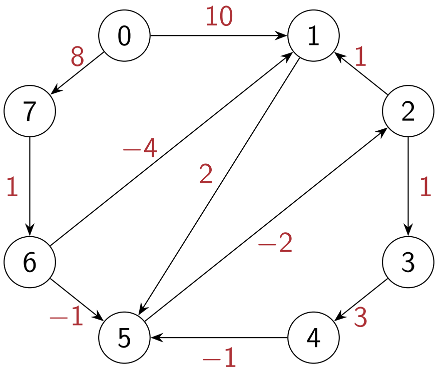
- For each vertex $j \in \{0, 1, \ldots, n{-}1\}$, for each edge $(j, k) \in E$, $D(k) = \min(\ D(k), D(j) + W(j,k)\ )$



Works for directed and undirected graphs

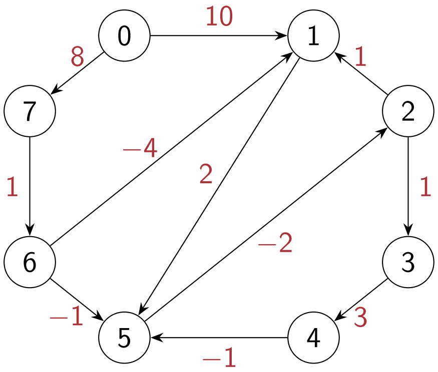
Bellman-Ford Algorithm

v	$D(v)$							
0								
1								
2								
3								
4								
5								
6								
7								



Bellman-Ford Algorithm

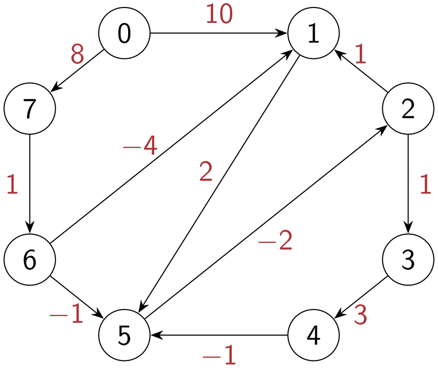
v	$D(v)$							
0	0							
1	∞							
2	∞							
3	∞							
4	∞							
5	∞							
6	∞							
7	∞							



- Initialize $D(0) = 0$

Bellman-Ford Algorithm

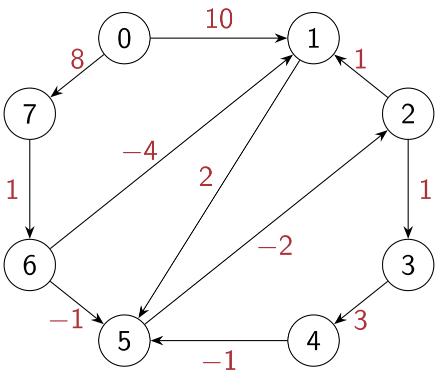
v	$D(v)$							
0	0	0						
1	∞	10						
2	∞	∞						
3	∞	∞						
4	∞	∞						
5	∞	∞						
6	∞	∞						
7	∞	8						



- Initialize $D(0) = 0$
- For each $(j, k) \in E$, update
  $D(k) = \min(\ D(k),$
  $D(j) + W(j, k)\ )$

Bellman-Ford Algorithm

v	$D(v)$							
0	0	0	0					
1	∞	10	10					
2	∞	∞	∞					
3	∞	∞	∞					
4	∞	∞	∞					
5	∞	∞	12					
6	∞	∞	9					
7	∞	8	8					

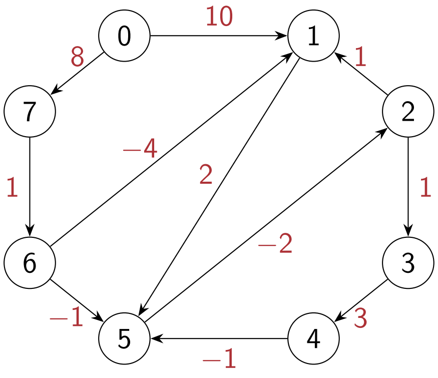


- Initialize $D(0) = 0$
- For each $(j,k) \in E$, update

$$D(k) = \min(\ D(k),\ D(j) + W(j,k)\ )$$

Bellman-Ford Algorithm

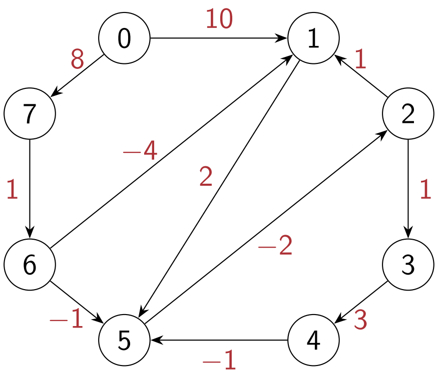
v	$D(v)$							
0	0	0	0	0				
1	∞	10	10	5				
2	∞	∞	∞	10				
3	∞	∞	∞	∞				
4	∞	∞	∞	∞				
5	∞	∞	12	8				
6	∞	∞	9	9				
7	∞	8	8	8				



- Initialize $D(0) = 0$
- For each $(j, k) \in E$, update
  $$D(k) = \min(\ D(k),\ D(j) + W(j,k)\ )$$

Bellman-Ford Algorithm

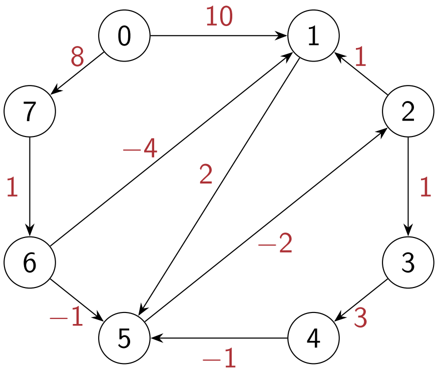
v	$D(v)$							
0	0	0	0	0	0			
1	∞	10	10	5	5			
2	∞	∞	∞	10	6			
3	∞	∞	∞	∞	11			
4	∞	∞	∞	∞	∞			
5	∞	∞	12	8	7			
6	∞	∞	9	9	9			
7	∞	8	8	8	8			



- Initialize $D(0) = 0$
- For each $(j, k) \in E$, update
  $D(k) = \min(\ D(k),$
  $\qquad D(j) + W(j, k)\ )$

Bellman-Ford Algorithm

v	$D(v)$							
0	0	0	0	0	0	0		
1	∞	10	10	5	5	5		
2	∞	∞	∞	10	6	5		
3	∞	∞	∞	∞	11	7		
4	∞	∞	∞	∞	∞	14		
5	∞	∞	12	8	7	7		
6	∞	∞	9	9	9	9		
7	∞	8	8	8	8	8		

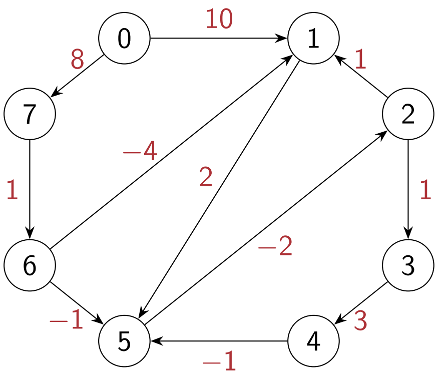


- Initialize $D(0) = 0$
- For each $(j, k) \in E$, update

  $D(k) = \min(D(k), D(j) + W(j, k))$

Bellman-Ford Algorithm

v	$D(v)$							
0	0	0	0	0	0	0	0	
1	∞	10	10	5	5	5	5	
2	∞	∞	∞	10	6	5	5	
3	∞	∞	∞	∞	11	7	6	
4	∞	∞	∞	∞	∞	14	10	
5	∞	∞	12	8	7	7	7	
6	∞	∞	9	9	9	9	9	
7	∞	8	8	8	8	8	8	

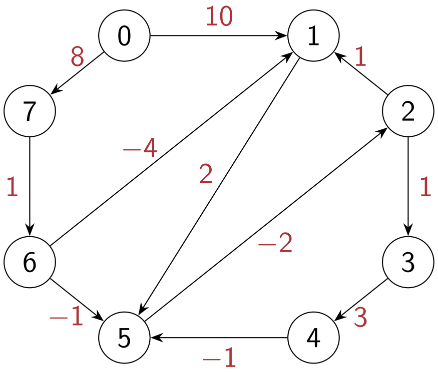


- Initialize $D(0) = 0$
- For each $(j, k) \in E$, update

$$D(k) = \min(\ D(k),\ D(j) + W(j, k)\ )$$

Bellman-Ford Algorithm

v	$D(v)$							
0	0	0	0	0	0	0	0	0
1	∞	10	10	5	5	5	5	5
2	∞	∞	∞	10	6	5	5	5
3	∞	∞	∞	∞	11	7	6	6
4	∞	∞	∞	∞	∞	14	10	9
5	∞	∞	12	8	7	7	7	7
6	∞	∞	9	9	9	9	9	9
7	∞	8	8	8	8	8	8	8

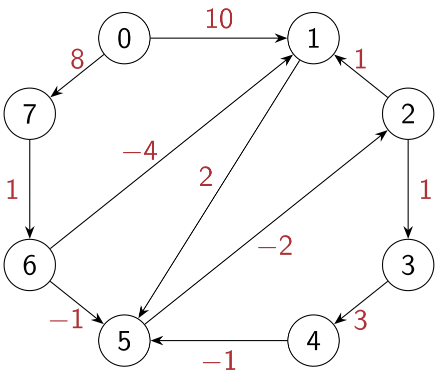


- Initialize $D(0) = 0$
- For each $(j, k) \in E$, update

  $D(k) = \min(D(k),$
  $D(j) + W(j, k))$

Bellman-Ford Algorithm

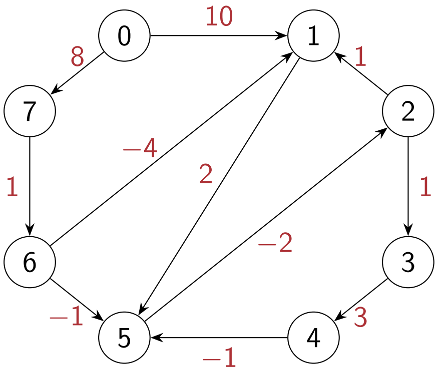
v	$D(v)$							
0	0	0	0	0	0	0	0	0
1	∞	10	10	5	5	5	5	5
2	∞	∞	∞	10	6	5	5	5
3	∞	∞	∞	∞	11	7	6	6
4	∞	∞	∞	∞	∞	14	10	9
5	∞	∞	12	8	7	7	7	7
6	∞	∞	9	9	9	9	9	9
7	∞	8	8	8	8	8	8	8



- What if there was a negative cycle?
- Distance would continue to decrease

Bellman-Ford Algorithm

v	$D(v)$							
0	0	0	0	0	0	0	0	0
1	∞	10	10	5	5	5	5	5
2	∞	∞	∞	10	6	5	5	5
3	∞	∞	∞	∞	11	7	6	6
4	∞	∞	∞	∞	∞	14	10	9
5	∞	∞	12	8	7	7	7	7
6	∞	∞	9	9	9	9	9	9
7	∞	8	8	8	8	8	8	8



- What if there was a negative cycle?
- Distance would continue to decrease
- Check if update n reduces any $D(v)$

Summary

- Dijkstra's algorithm assumes non-negative edge weights
  - Final distance is frozen each time a vertex "burns"
  - Should not encounter a shorter route discovered later
- Without negative cycles, every shortest route is a path
- Every prefix of a shortest path is also a shortest path
- Iteratively find shortest paths of length $1, 2, \ldots, n-1$
- Update distance to each vertex with every iteration — Bellman-Ford algorithm
- If Bellman-Ford algorithm does not converge after $n - 1$ iterations, there is a negative cycle

All-Pairs Shortest Paths

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Mathematics for Data Science 1
Week 12

Shortest paths in weighted graphs

Two types of shortest path problems of interest

Single source shortest paths

- Find shortest paths from a fixed vertex to every other vertex
- Transport finished product from factory (single source) to all retail outlets
- Courier company delivers items from distribution centre (single source) to addressees
- Dijkstra's algorithm (non-negative weights), Bellman-Ford algorithm (allows negative weights)

All pairs shortest paths

- Find shortest paths between every pair of vertices i and j
- Optimal airline, railway, road routes between cities
- Run Dijkstra or Bellman-Ford from each vertex
- Is there is another way?

Transitive closure

- Recall transitive closure algorithm
- Adjacency matrix A represents paths of length 1
- Matrix multiplication, $A^2 = A \times A$
  - $A^2[i,j] = 1$ if there is a path of length 2 from i to j
  - For some k, $A[i,k] = A[k,j] = 1$
- In general, $A^{\ell+1} = A^{\ell} \times A$,
  - $A^{\ell+1}[i,j] = 1$ if there is a path of length $\ell+1$ from i to j
  - For some k, $A^{\ell}[i,k] = 1$, $A[k,j] = 1$
- $A^+ = A + A^2 + \cdots + A^{n-1}$

An alternative approach

- $B^k[i,j] = 1$ if there is path from i to j via vertices $\{0, 1, \ldots, k-1\}$
  - Constraint applies only to intermediate vertices i to j
  - $B^0[i,j] = 1$ if there is a direct edge
  - $B^0 = A$
- $B^{k+1}[i,j] = 1$ if
  - $B^k[i,j] = 1$ — can already reach j from i via $\{0, 1, \ldots, k-1\}$
  - $B^k[i,k] = 1$ and $B^k[k,j] = 1$ — use $\{0, 1, \ldots, k-1\}$ to go from i to k and then from k to j

Warshall's Algorithm

- The algorithm on the right also computes transitive closure — Warshall's algorithm
- $B^n[i,j] = 1$ if there is some path from i to j with intermediate vertices in $\{0, 1, \ldots, n-1\}$
- $B^n = A^+$
- We adapt Warshall's algorithm to compute all-pairs shortest paths

Computing transitive closure

- $B^k[i,j] = 1$ if there is path from i to j via vertices $\{0, 1, \ldots, k-1\}$
- $B^0[i,j] = A[i,j]$
  - Direct edges, no intermediate vertices
- $B^{k+1}[i,j] = 1$ if
  - $B^k[i,j] = 1$, or
  - $B^k[i,k] = 1$ and $B^k[k,j] = 1$

Floyd-Warshall Algorithm

- Let $SP^k[i,j]$ be the length of the shortest path from i to j via vertices $\{0, 1, \ldots, k-1\}$
- $SP^0[i,j] = W[i,j]$
  - No intermediate vertices, shortest path is weight of direct edge
  - Assume $W[i,j] = \infty$ if $(i,j) \notin E$
- $SP^{k+1}[i,j]$ is the minimum of
  - $SP^k[i,j]$

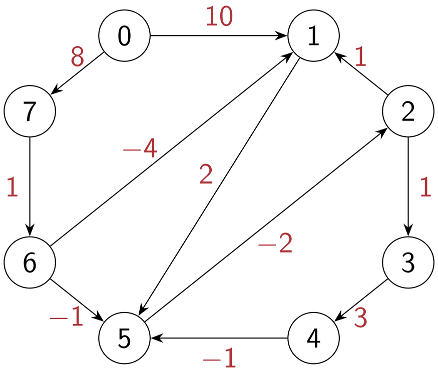
    Shortest path using only $\{0, 1, \ldots, k-1\}$
  - $SP^k[i,k] + SP^k[k,j]$

    Combine shortest path from i to k and k to j
- $SP^n[i,j] = 1$ is the length of the shortest path overall from i to j
  - Intermediate vertices lie in $\{0, 1, \ldots, n-1\}$

Floyd-Warshall Algorithm

SP^0	0	1	2	3	4	5	6	7
0	∞	10	∞	∞	∞	∞	∞	8
1	∞	∞	∞	∞	∞	2	∞	∞
2	∞	1	∞	1	∞	∞	∞	∞
3	∞	∞	∞	∞	3	∞	∞	∞
4	∞	∞	∞	∞	∞	-1	∞	∞
5	∞	∞	-2	∞	∞	∞	∞	∞
6	∞	-4	∞	∞	∞	-1	∞	∞
7	∞	∞	∞	∞	∞	∞	1	∞

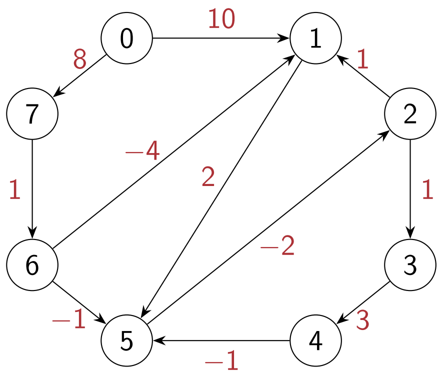
SP^1	0	1	2	3	4	5	6	7
0	∞	10	∞	∞	∞	∞	∞	8
1	∞	∞	∞	∞	∞	2	∞	∞
2	∞	1	∞	1	∞	∞	∞	∞
3	∞	∞	∞	∞	3	∞	∞	∞
4	∞	∞	∞	∞	∞	-1	∞	∞
5	∞	∞	-2	∞	∞	∞	∞	∞
6	∞	-4	∞	∞	∞	-1	∞	∞
7	∞	∞	∞	∞	∞	∞	1	∞



Floyd-Warshall Algorithm

SP^1	0	1	2	3	4	5	6	7
0	∞	10	∞	∞	∞	∞	∞	8
1	∞	∞	∞	∞	∞	2	∞	∞
2	∞	1	∞	1	∞	∞	∞	∞
3	∞	∞	∞	∞	3	∞	∞	∞
4	∞	∞	∞	∞	∞	-1	∞	∞
5	∞	∞	-2	∞	∞	∞	∞	∞
6	∞	-4	∞	∞	∞	-1	∞	∞
7	∞	∞	∞	∞	∞	∞	1	∞

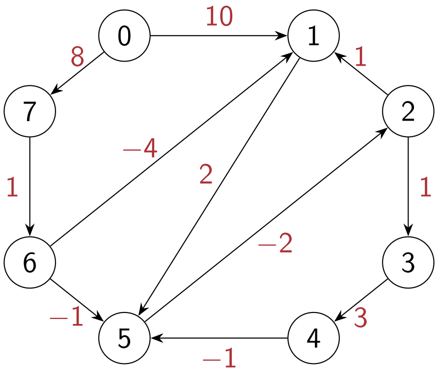
SP^2	0	1	2	3	4	5	6	7
0	∞	10	∞	∞	∞	12	∞	8
1	∞	∞	∞	∞	∞	2	∞	∞
2	∞	1	∞	1	∞	3	∞	∞
3	∞	∞	∞	∞	3	∞	∞	∞
4	∞	∞	∞	∞	∞	-1	∞	∞
5	∞	∞	-2	∞	∞	∞	∞	∞
6	∞	-4	∞	∞	∞	-2	∞	∞
7	∞	∞	∞	∞	∞	∞	1	∞



Floyd-Warshall Algorithm

SP^2	0	1	2	3	4	5	6	7
0	∞	10	∞	∞	∞	12	∞	8
1	∞	∞	∞	∞	∞	2	∞	∞
2	∞	1	∞	1	∞	3	∞	∞
3	∞	∞	∞	∞	3	∞	∞	∞
4	∞	∞	∞	∞	∞	-1	∞	∞
5	∞	∞	-2	∞	∞	∞	∞	∞
6	∞	-4	∞	∞	∞	-2	∞	∞
7	∞	∞	∞	∞	∞	∞	1	∞

SP^3	0	1	2	3	4	5	6	7
0	∞	10	∞	∞	∞	12	∞	8
1	∞	∞	∞	∞	∞	2	∞	∞
2	∞	1	∞	1	∞	3	∞	∞
3	∞	∞	∞	∞	3	∞	∞	∞
4	∞	∞	∞	∞	∞	-1	∞	∞
5	∞	-1	-2	-1	∞	1	∞	∞
6	∞	-4	∞	∞	∞	-2	∞	∞
7	∞	∞	∞	∞	∞	∞	1	∞



Summary

- Warshall's algorithm is an alternative way to compute transitive closure
  - $B^k[i,j] = 1$ if we can reach j from i using vertices in $\{0, 1, \ldots, k-1\}$
- Adapt Warshall's algorithm to compute all pairs shortest paths
  - $SP^k[i,j]$ is the length of the shortest path from i to j using vertices in $\{0, 1, \ldots, k-1\}$
  - $SP^n[i,j]$ is the length of the overall shortest path
  - Floyd-Warshall algorithm
- Works with negative edge weights, assuming no negative cycles

Minimum Cost Spanning Trees

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

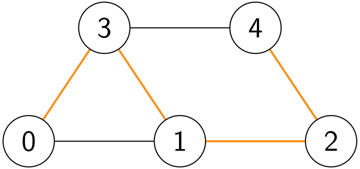
Mathematics for Data Science 1
Week 12

Examples

Roads

- District hit by cyclone, roads are damaged
- Government sets to work to restore roads
- Priority is to ensure that all parts of the district can be reached
- What set of roads should be restored first?

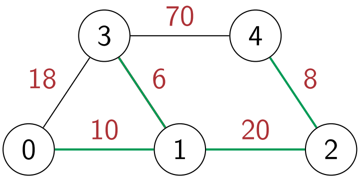
Spanning trees



- Retain a minimal set of edges so that graph remains connected
- Recall that a minimally connected graph is a tree
  - Adding an edge to a tree creates a loop
  - Removing an edge disconnects the graph
- Want a tree that connects all the vertices — spanning tree
- More than one spanning tree, in general

Spanning trees with costs

- Restoring a road or laying a fibre optic cable has a cost
- Minimum cost spanning tree
  - Add the cost of all the edges in the tree
  - Among the different spanning trees, choose one with minimum cost
- Example
  - Spanning tree, Cost is 114 — not minimum cost spanning tree
  - Anonther spanning tree, Cost is 44 — minimum cost spanning tree



Some facts about trees

Definition A tree is a connected acylic graph.

Fact 1

A tree on n vertices has exactly $n-1$ edges

- Initially, one single component
- Deleting edge (i,j) must split component
  - Otherwise, there is still a path from i to j, combine with (i,j) to form cycle
- Each edge deletion creates one more component
- Deleting $n-1$ edges creates n components, each an isolated vertex

Fact 2

Adding an edge to a tree must create a cycle.

- Suppose we add an edge (i,j)
- Tree is connected, so there is already a path from i to j
- The new edge (i,j) combined with this path from i to j forms a cycle

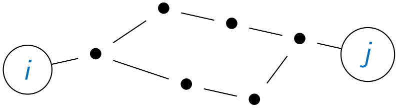
Some facts about trees

Definition A tree is a connected acylic graph.

Fact 3

In a tree, every pair of vertices is connected by a unique path.

- If there are two paths from i to j, there must be a cycle



Observation

Any two of the following facts about a graph G implies the third

- G is connected
- G is acyclic
- G has $n-1$ edges

Building minimum cost spanning trees

- We will use these facts about trees to build minimum cost spanning trees
- Two natural strategies
- Start with the smallest edge and "grow" a tree
  - Prim's algorithm
- Scan the edges in ascending order of weight to connect components without forming cycles
  - Kruskal's algorithm

Minimum Cost Spanning Trees: Prim's Algorithm

Madhavan Mukund

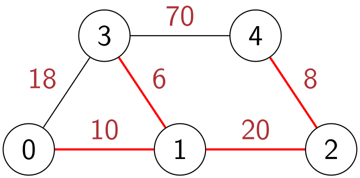
https://www.cmi.ac.in/~madhavan

Mathematics for Data Science 1
Week 12

Minimum cost spanning tree (MCST)

- Weighted undirected graph, $G = (V, E), W : E \to \mathbb{R}$
  - G assumed to be connected
- Find a minimum cost spanning tree
  - Tree connecting all vertices in V
- Strategy
  - Incrementally grow the minimum cost spanning tree
  - Start with a smallest weight edge overall
  - Extend the current tree by adding the smallest edge from the tree to a vertex not yet in the tree

Example

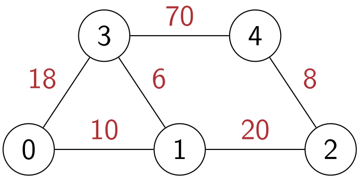


- Start with smallest edge, $(1, 3)$
- Extend the tree with $(1, 0)$
- Can't add $(0, 3)$, forms a cycle
- Instead, extend the tree with $(1, 2)$
- Extend the tree with $(2, 4)$

Prim's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Incrementally build an MCST
  - $TV \subseteq V$: tree vertices, already added to MCST
  - $TE \subseteq E$: tree edges, already added to MCST
- Initially, $TV = TE = \emptyset$
- Choose minimum weight edge $e = (i, j)$
  - Set $TV = \{i, j\}$, $TE = \{e\}$ MCST
- Repeat $n - 2$ times
  - Choose minimum weight edge $f = (u, v)$ such that $u \in TV$, $v \notin TV$
  - Add v to TV, f to TE

Example



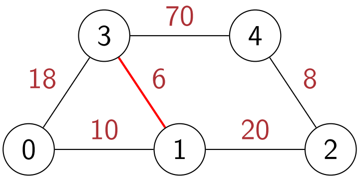
$TV = \emptyset$

$TE = \emptyset$

Prim's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Incrementally build an MCST
  - $TV \subseteq V$: tree vertices, already added to MCST
  - $TE \subseteq E$: tree edges, already added to MCST
- Initially, $TV = TE = \emptyset$
- Choose minimum weight edge $e = (i, j)$
  - Set $TV = \{i, j\}$, $TE = \{e\}$ MCST
- Repeat $n - 2$ times
  - Choose minimum weight edge $f = (u, v)$ such that $u \in TV$, $v \notin TV$
  - Add v to TV, f to TE

Example



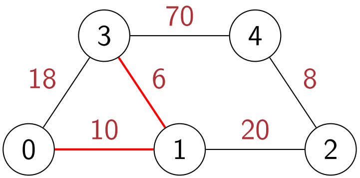
$TV = \{1, 3\}$

$TE = \{(1, 3)\}$

Prim's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Incrementally build an MCST
  - $TV \subseteq V$: tree vertices, already added to MCST
  - $TE \subseteq E$: tree edges, already added to MCST
- Initially, $TV = TE = \emptyset$
- Choose minimum weight edge $e = (i, j)$
  - Set $TV = \{i, j\}$, $TE = \{e\}$ MCST
- Repeat $n - 2$ times
  - Choose minimum weight edge $f = (u, v)$ such that $u \in TV$, $v \notin TV$
  - Add v to TV, f to TE

Example



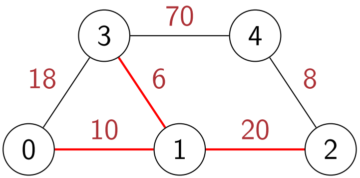
$TV = \{1, 3, 0\}$

$TE = \{(1, 3), (1, 0)\}$

Prim's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Incrementally build an MCST
  - $TV \subseteq V$: tree vertices, already added to MCST
  - $TE \subseteq E$: tree edges, already added to MCST
- Initially, $TV = TE = \emptyset$
- Choose minimum weight edge $e = (i, j)$
  - Set $TV = \{i, j\}$, $TE = \{e\}$ MCST
- Repeat $n - 2$ times
  - Choose minimum weight edge $f = (u, v)$ such that $u \in TV$, $v \notin TV$
  - Add v to TV, f to TE

Example



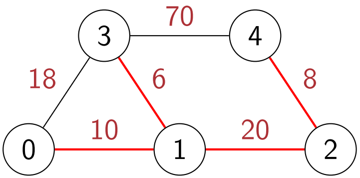
$TV = \{1, 3, 0, 2\}$

$TE = \{(1, 3), (1, 0), (1, 2)\}$

Prim's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Incrementally build an MCST
  - $TV \subseteq V$: tree vertices, already added to MCST
  - $TE \subseteq E$: tree edges, already added to MCST
- Initially, $TV = TE = \emptyset$
- Choose minimum weight edge $e = (i, j)$
  - Set $TV = \{i, j\}$, $TE = \{e\}$ MCST
- Repeat $n - 2$ times
  - Choose minimum weight edge $f = (u, v)$ such that $u \in TV$, $v \notin TV$
  - Add v to TV, f to TE

Example



$TV = \{1, 3, 0, 2, 4\}$

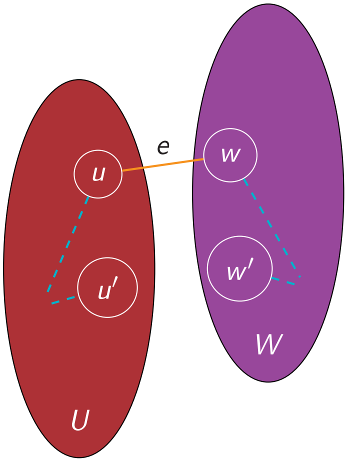
$TE = \{(1, 3), (1, 0), (1, 2), (2, 4)\}$

Correctness of Prim's algorithm

Minimum Separator Lemma

- Let V be partitioned into two non-empty sets U and $W = V \setminus U$
- Let $e = (u, w)$ be the minimum cost edge with $u \in U$, $w \in W$
- Every MCST must include e

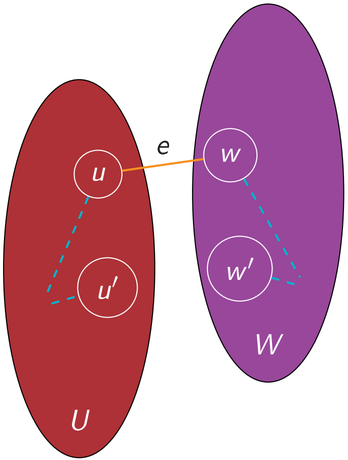
- Assume for now, all edge weights distinct
- Let T be an MCST, $e \notin T$
- T contains a path p from u to
  - p starts in U, ends in W
  - Let $f = (u', w')$ be the first edge on p crossing from U to W
  - Drop f, add e to get a cheaper spanning tree



Correctness of Prim's algorithm

Minimum Separator Lemma

- Let V be partitioned into two non-empty sets U and $W = V \setminus U$
- Let $e = (u, w)$ be the minimum cost edge with $u \in U$, $w \in W$
- Every MCST must include e



- Assume for now, all edge weights distinct
- What if two edges have the same weight?
- Assign each edge a unique index from 0 to $m - 1$
- Define $(e, i) < (f, j)$ if $W(e) < W(j)$ or $W(e) = W(j)$ and $i < j$

Correctness of Prim's algorithm

Minimum Separator Lemma

- Let V be partitioned into two non-empty sets U and $W = V \setminus U$
- Let $e = (u, w)$ be the minimum cost edge with $u \in U$, $w \in W$
- Every MCST must include e

- In Prim's algorithm, TV and $W = V \setminus TV$ partition V
- Algorithm picks smallest edge connecting TV and W, which must belong to every MCST
- In fact, for any $v \in V$, $\{v\}$ and $V \setminus \{v\}$ form a partition
- The smallest weight edge leaving any vertex must belong to every MCST
- We started with overall minimum cost edge
- Instead, can start at any vertex v, with $TV = \{v\}$ and $TE = \emptyset$
- First iteration will pick minimum cost edge from v

Summary

- Prim's algorithm grows an MCST starting with any vertex
- At each step, connect one more vertex to the tree using minimum cost edge from inside the tree to outside the tree
- Correctness follows from Minimum Separator Lemma

Minimum Cost Spanning Trees: Kruskal's Algorithm

Madhavan Mukund

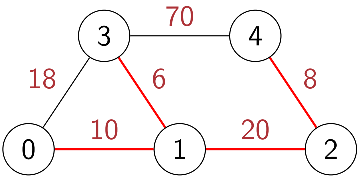
https://www.cmi.ac.in/~madhavan

Mathematics for Data Science 1
Week 12

Minimum cost spanning tree (MCST)

- Weighted undirected graph, $G = (V, E), W : E \to \mathbb{R}$
  - G assumed to be connected
- Find a minimum cost spanning tree
  - Tree connecting all vertices in V
- Strategy 2
  - Start with n components, each a single vertex
  - Process edges in ascending order of cost
  - Include edge if it does not create a cycle

Example

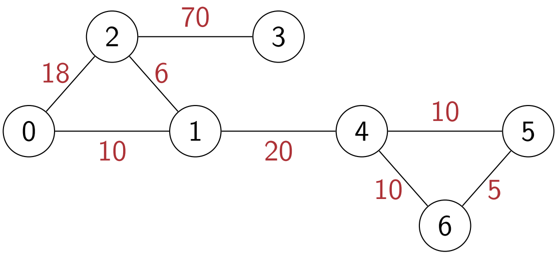


- Start with smallest edge, $(1, 3)$
- Add next smallest edge, $(2, 4)$
- Add next smallest edge, $(0, 1)$
- Can't add $(0, 3)$, forms a cycle
- Add next smallest edge, $(1, 2)$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



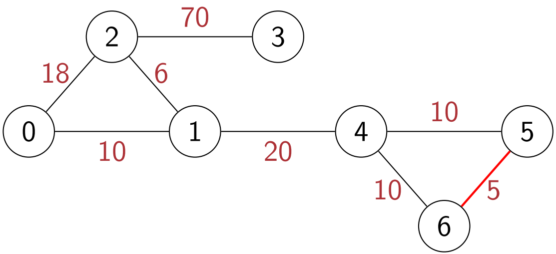
Sort E as
$\{(5,6), (1,2), (0,1), (4,5), (4,6), (0,2), (1,4), (2,3)\}$

Set $TE = \emptyset$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as
$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

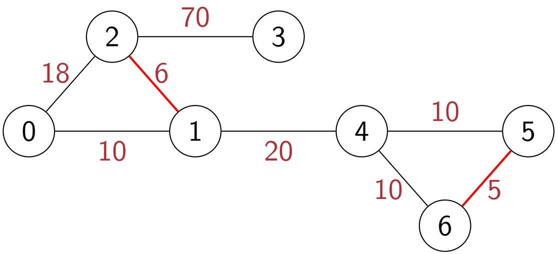
Add $(5,6)$

Set $TE = \{(5,6)\}$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as
$\{(5,6), (1,2), (0,1), (4,5), (4,6), (0,2), (1,4), (2,3)\}$

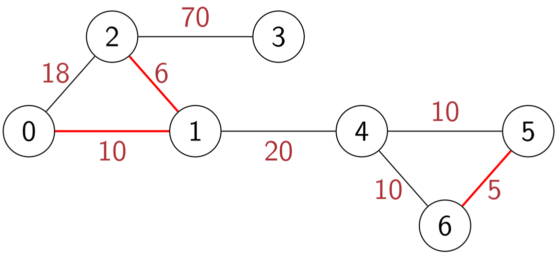
Add $(1,2)$

Set $TE = \{(5,6), (1,2)\}$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as

$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

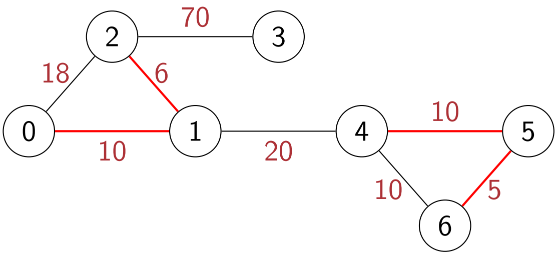
Add $(0,1)$

Set $TE = \{(5,6),(1,2),(0,1)\}$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as

$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

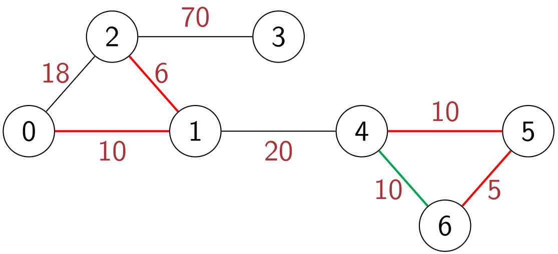
Add $(4,5)$

Set $TE = \{(5,6),(1,2),(0,1),(4,5)\}$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as
$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

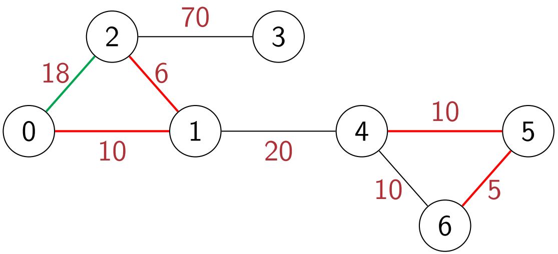
Skip $(4,6)$

Set $TE = \{(5,6),(1,2),(0,1),(4,5)\}$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as

$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

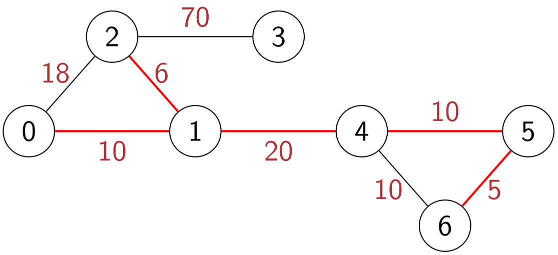
Skip $(0,2)$

Set $TE = \{(5,6),(1,2),(0,1),(4,5)\}$

Kruskal's algorithm

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



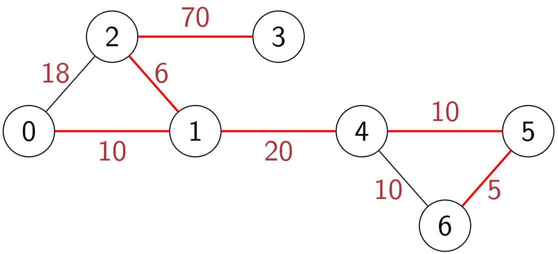
Sort E as
$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

Add $(1,4)$

Set $TE = \{(5,6),(1,2),(0,1),(4,5),(1,4)\}$

- $G = (V, E)$, $W : E \to \mathbb{R}$
- Let $E = \{e_0, e_1, \ldots, e_{m-1}\}$ be edges sorted in ascending order by weight
- Let $TE \subseteq E$ be the set of tree edges already added to MCST
- Initially, $TE = \emptyset$
- Scan E from e_0 to e_{m-1}
  - If adding e_i to TE creates a loop, skip it
  - Otherwise, add e_i to TE

Example



Sort E as

$\{(5,6),(1,2),(0,1),(4,5),(4,6),(0,2),(1,4),(2,3)\}$

Add $(2,3)$

Set $TE = \{(5,6),(1,2),(0,1),(4,5),(1,4),(2,3)\}$

Correctness of Krusksal's algorithm

Minimum Separator Lemma

- Let V be partitioned into two non-empty sets U and $W = V \setminus U$
- Let $e = (u, w)$ be the minimum cost edge with $u \in U$, $w \in W$
- Every MCST must include e

- Edges in TE partition vertices into connected components
- Initially each vertex is a separate component
- Adding $e = (u, w)$ merges components of u and w
  - If u and w are in the same component, e forms a cycle and is discarded
  - Otherwise, let U be component of u, W be $V \setminus U$
  - U, W form a partition of V with $u \in U$ and $w \in W$
  - Since we are scanning edges in ascending order of cost, e is minimum cost edge connecting U and W, so it must be part of any MCST

Summary

- Kruskal's algorithm builds an MCST bottom up
  - Start with n components, each an isolated vertex
  - Scan edges in ascending order of cost
  - Whenever an edge merges disjoint components, add it to the MCST
- Correctness follows from Minimum Separator Lemma
- If edge weights repeat, MCST is not unique
- "Choose minimum cost edge" will allow choices
  - Consider a triangle on 3 vertices with all edges equal
- Different choices lead to different spanning trees
- In general, there may be a very large number of minimum cost spanning trees