

# Ve203 Discrete Mathematics

Runze Cai

University of Michigan - Shanghai Jiao Tong University  
Joint Institute

Fall 2021



**JOINT INSTITUTE**  
                  
**交大密西根学院**

## Part I

# Basic Set Theory and Applications

# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Sets

## Definition

A set is an unordered collection of distinct objects, called **elements** or **members** of the set. A set is said to contain its elements. We write

- ▶  $a \in A$  if  $a$  is an element of the set  $A$ .
- ▶  $a \notin A$  if  $a$  is not an element of the set  $A$ .

## Examples

- ▶ The set  $P$  of primes less than 10:  $P = \{2, 3, 5, 7\}$ .
- ▶ The set  $V$  of all vowels in the English alphabet:  $V = \{a, e, i, o, u\}$ .
- ▶ The set  $S$  of all suits  $S = \{\heartsuit, \diamondsuit, \clubsuit, \spadesuit\}$ .
- ▶ Different kinds of objects:  $S = \{\text{USA}, \text{USB}, \text{UCSB}, 0.0, \{\text{CAT}\}\}$ .
- ▶ The empty set  $S = \{\} = \emptyset = \emptyset$ .

# Multisets

## Caution

- ▶ Elements in a set are *distinct* and *unordered*.

## Example

- ▶  $\{1, 0, 0, 0\} = \{0, 0, 1, 1\} = \{0, 1\} = \{1, 0\}$
- ▶  $\{\pm 1\} = \{(-1)^n \mid n \in \mathbb{N}\} = \{1, -1, 1, -1, \dots\}$

A multiset does allow repeated objects. (but order still does not matter)

## Example

- ▶ Roots of a polynomial.
- ▶ Eigenvalues of a square matrix.
- ▶ Stock of drinks.

# Set Notation

## Number Systems

- ▶  $\mathbb{N}$ , the natural numbers
- ▶  $\mathbb{Z}$ , the integers
- ▶  $\mathbb{Q}$ , the rational numbers
- ▶  $\mathbb{R}$ , the real numbers
- ▶  $\mathbb{C}$ , the complex numbers

## Cardinality

The **size** of a set  $A$  is called its **cardinality**, denoted by  $|A|$ ,  $\#A$ , or  $\text{card } A$ .

- ▶  $|A| = n \in \mathbb{N}$  if  $A$  is a finite set;
- ▶  $|A| = \infty$  otherwise. (Question: infinities?)

# Set Operations

Let  $A, B$  be sets.

## Inclusion

- ▶  $A$  is a subset of  $B$ , denoted by  $A \subset B$ , if every element of  $A$  is an element of  $B$ .
- ▶  $B$  is called a superset of  $A$ , denoted by  $B \supset A$ .

## Remark

Unlike  $<$  and  $\leq$ ,  $A \subset B$  is the same as  $A \subseteq B$ . Similarly for  $\supset$  and  $\supseteq$ .

## Proper Subset/Superset

$A$  is a proper subset of  $B$  if  $A \subset B$  and  $A \neq B$ , denoted by  $A \subsetneq B$  or  $A \subsetneqq B$ . Similarly for proper superset.

## Remark

$A = B$  if and only if  $A \subset B$  and  $B \subset A$ . (cf.,  $x = y$  iff  $x \leq y$  and  $y \leq x$ .)

## Set Operations

Let  $A, B$  be sets.

### Union

The **union** of  $A$  and  $B$  is the set of elements in either  $A$  or  $B$ , denoted by

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

### Intersection

The **intersection** of  $A$  and  $B$  is the set of elements in both  $A$  and  $B$ , denoted by

$$A \cap B := \{x \mid x \in A \text{ and } x \in B\}$$

### Questions

- ▶  $\bigcap \emptyset = ?$
- ▶  $\bigcup \emptyset = ?$
- ▶  $\{A, B\} = A \cup B ?$

# Set Operations

Let  $A, B$  be sets.

## Set Difference

The **set difference** of  $A$  and  $B$ , denoted by  $A - B$ , or  $A \setminus B$ , is the set of elements in  $A$  but not in  $B$ , that is,

$$\begin{aligned} A - B &:= \{x \mid x \in A \text{ and } x \notin B\} \\ &= \{x \in A \mid x \notin B\} \end{aligned}$$

## Symmetric Difference

The **symmetric difference** of  $A$  and  $B$  is the set of elements that are in **exclusively** one of  $A$  and  $B$ , but not the other.

$$A \triangle B = (A - B) \cup (B - A)$$

# Set Operations

## Power Set

The power set of a set  $A$  is the set of all subsets of  $A$ , denoted by  $\mathcal{P}(A)$  or  $2^A$ .

## Example

- ▶  $\mathcal{P}(\{j, i\}) = \{\emptyset, \{j\}, \{i\}, \{j, i\}\}$ .
- ▶  $\mathcal{P}(\{0\}) = \{\emptyset, \{0\}\}$ .
- ▶  $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$ .
- ▶  $\mathcal{P}(\emptyset) = \{\emptyset\}$ .

## Cardinality of Power sets

Given a finite set  $A$ ,

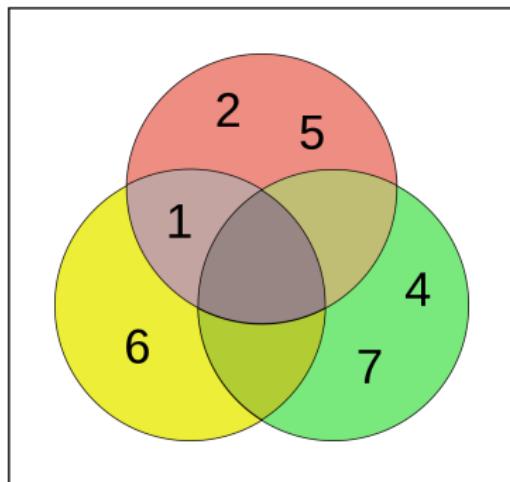
$$|2^A| = |\mathcal{P}(A)| = 2^{|A|}$$

# Venn Diagram vs Euler Diagram

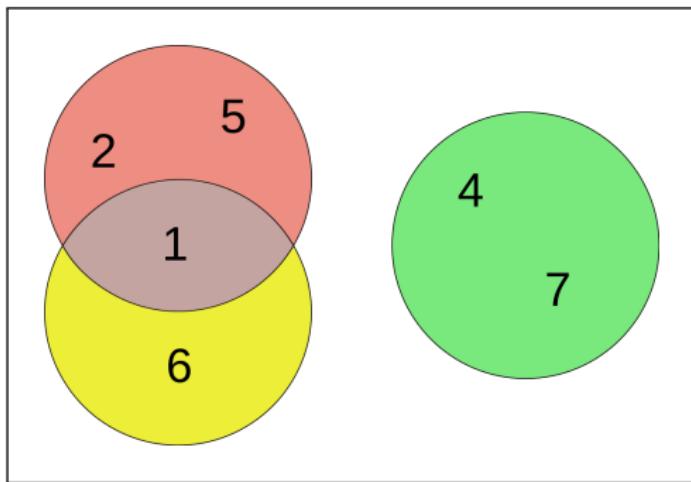
Given

- ▶  $A = \{1, 2, 5\}$
- ▶  $B = \{1, 6\}$
- ▶  $C = \{4, 7\}$

Venn Diagram



Euler Diagram



## Set Algebras

Let  $A, B, C$  be sets.

- ▶ Commutative Laws
  - ▶  $A \cup B = B \cup A$
  - ▶  $A \cap B = B \cap A$
- ▶ Associative Laws
  - ▶  $(A \cup B) \cup C = A \cup (B \cup C)$
  - ▶  $(A \cap B) \cap C = A \cap (B \cap C)$
- ▶ (Left) Distributive Laws
  - ▶  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  - ▶  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- ▶ De Morgan's Laws
  - ▶  $C - (A \cup B) = (C - A) \cap (C - B)$
  - ▶  $C - (A \cap B) = (C - A) \cup (C - B)$

# Cartesian Product

## Definition

The Cartesian product of sets  $A$  and  $B$  is the set of **ordered pairs**, such that

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

## Definition (Kuratowski)

An ordered pair  $(a, b)$  is given by

$$(a, b) := \{\{a\}, \{a, b\}\}$$

## Theorem

If  $a \in C$  and  $b \in C$ , then  $(a, b) \in \mathcal{P}(\mathcal{P}(C))$ .

## Proof.

- ▶  $a \in C \Rightarrow \{a\} \in \mathcal{P}(C); a, b \in C \Rightarrow \{a, b\} \in \mathcal{P}(C).$
- ▶ Hence  $(a, b) = \{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(C)).$



# Cartesian Product

## Theorem

$(x, y) = (a, b)$  iff  $x = a$  and  $y = b$ .

## Proof.

- ▶ ( $\Leftarrow$ ). Trivial.
- ▶ ( $\Rightarrow$ ). By definition, we need to show that

$$\underbrace{\{\{x\}, \{x, y\}\}}_U = \underbrace{\{\{a\}, \{a, b\}\}}_V \Rightarrow x = a \text{ and } y = b$$

- ▶ If  $x \neq y$ , then  $|U| = |V| = 2$  (b/c  $\{x\} \neq \{x, y\}$ ), hence  $U = V$ .  
By matching sizes we have  $\{x\} = \{a\}$  and  $\{x, y\} = \{a, b\}$ ,  
therefore  $x = a$  and  $y = b$ .
- ▶ If  $x = y$ , then  $|U| = |V| = 1$ . Similarly we have  
 $x = y = a = b$ .

□

## Cartesian Product of Sets

In this manner, we can define Cartesian product of three sets as the set of **ordered triples**, e.g.,

$$A \times B \times C := \{(a, b, c) \mid a \in A, b \in B, c \in C\}.$$

More generally, the  $n$ -fold Cartesian product  $A_1 \times \cdots \times A_n$  of sets  $A_k$ ,  $k = 1, \dots, n$ , as the set of ordered  **$n$ -tuple**  $(a_1, \dots, a_n)$ .

If we take the cartesian product of a set with itself, we may abbreviate it using exponents, e.g.,

$$A^2 := A \times A, \quad A^3 := A \times A \times A, \dots$$

$$\mathbb{N}^2 := \mathbb{N} \times \mathbb{N}, \quad \mathbb{N}^3 := \mathbb{N} \times \mathbb{N} \times \mathbb{N}, \dots$$

## Associative Set Operations

Let  $A_1, A_2, \dots, A_n$  be sets, then

- ▶  $A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^n A_i$
- ▶  $A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^n A_i$
- ▶  $A_1 \times A_2 \times \cdots \times A_n = \bigtimes_{i=1}^n A_i$
- ▶  $A_1 \triangle A_2 \triangle \cdots \triangle A_n = \bigtriangleup_{i=1}^n A_i$

### Remark

Brackets are permitted everywhere but not required anywhere.

### Question

How many ways to put the brackets?

# Simple Graphs

## $k$ -element subsets

Let  $X$  be a finite set. For a positive integer  $k$ , let  $\binom{X}{k}$  denote the set of all  $k$ -element subsets. Note that  $|\binom{X}{k}| = \binom{|X|}{k}$ .

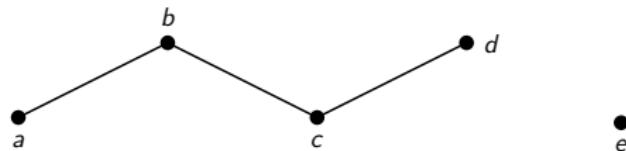
## Definition

A finite simple **graph**  $G$  is a pair  $(V, E)$  where  $V$  is a non-empty finite set and  $E$  is a set of 2-element subsets of  $V$ , i.e.,  $E \subset \binom{V}{2}$ . Elements of  $V$  are called **vertices** and elements of  $E$  are called **edges**. We also call  $V$  the **vertex set** of  $G$ , denoted  $V(G)$ , and  $E$  the **edge set** of  $G$ , denoted  $E(G)$ .

## Example

Consider the following simple graph  $G = (V, E)$ , where

- ▶  $V(G) = \{a, b, c, d, e\}$ ,
- ▶  $E(G) = \{\{a, b\}, \{b, c\}, \{c, d\}\}$ .



# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Propositional Logic

## Definition

A ***proposition*** or ***statement*** is a declarative sentence that is either **true** or **false**, but not both.

## Examples

- ▶ Washington, D.C., is the capital of the United States of America.
- ▶ Toronto is the capital of Canada.
- ▶  $1 + 1 = 2$ .
- ▶  $2 + 2 = 3$ .

## Non-examples

- ▶ What time is it?
- ▶ Read this carefully.
- ▶  $x + 1 = 2$ .
- ▶  $x + y = z$ .

# Notation

## Propositional/Logical Variables

Denoted by  $p, q, r, \dots$

## True or False

- ▶ True: T, 1,  $\top$
- ▶ False: F, 0,  $\perp$

## Connectives

- ▶  $\neg$ , negation/not
- ▶  $\wedge$ , and
- ▶  $\vee$ , or (inclusive or)
- ▶  $\rightarrow$ , implies
- ▶  $\leftrightarrow$ , if and only if (iff)

# Conjunction and disjunction

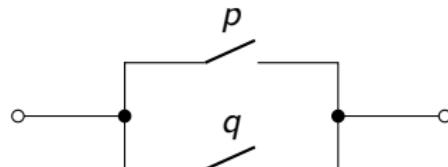
AND

| $p$ | $q$ | $p \wedge q$ |
|-----|-----|--------------|
| 0   | 0   | 0            |
| 0   | 1   | 0            |
| 1   | 0   | 0            |
| 1   | 1   | 1            |



OR (inclusive or)

| $p$ | $q$ | $p \vee q$ |
|-----|-----|------------|
| 0   | 0   | 0          |
| 0   | 1   | 1          |
| 1   | 0   | 1          |
| 1   | 1   | 1          |



Remark

- ▶ In the conjunction  $p \wedge q$ , the proposition  $p$  and  $q$  are called **conjuncts**.
- ▶ In the disjunction  $p \vee q$ , the proposition  $p$  and  $q$  are called **disjuncts**.

# Negation and Exclusive OR

NOT

| $p$ | $\neg p$ |
|-----|----------|
| 0   | 1        |
| 1   | 0        |

XOR (exclusive or)

| $p$ | $q$ | $p \oplus q$ |
|-----|-----|--------------|
| 0   | 0   | 0            |
| 0   | 1   | 1            |
| 1   | 0   | 1            |
| 1   | 1   | 0            |

# CNF and DNF

## Conjunctive Normal Form (CNF)

A proposition is in **Conjunctive Normal Form (CNF)** if it is a conjunction of one or more clauses, where a clause is a disjunction of literals; i.e., it is a **product of sums** or an **AND of ORs**.

## Disjunctive Normal Form (DNF)

A proposition is in **Disjunctive Normal Form (CNF)** if it is a Disjunction of one or more clauses, where a clause is a conjunction of literals; i.e., it is a **sum of products** or an **OR of ANDS**.

### Remark

A **literal** is a Boolean variable, (i.e., an atomic proposition) or its negation.

### Example

- ▶ CNF:  $(\neg p \vee q \vee r) \wedge (\neg q \vee \neg r) \wedge (r)$
- ▶ DNF:  $(\neg p \wedge q \wedge r) \vee (\neg q \wedge \neg r) \vee (r)$

# Conditional Statements

## Conditional statement (implication)

- ▶  $p$ : hypothesis/antecedent/premise
- ▶  $q$ : thesis/conclusion/consequence

| $p$ | $q$ | $p \rightarrow q$ |
|-----|-----|-------------------|
| 0   | 0   | 1                 |
| 0   | 1   | 1                 |
| 1   | 0   | 0                 |
| 1   | 1   | 1                 |

## Equivalent forms

- ▶ if  $p$ , then  $q$
- ▶ if  $p$ ,  $q$
- ▶  $p$  is sufficient for  $q$
- ▶  $q$  if  $p$
- ▶  $q$  when  $p$
- ▶ a necessary condition for  $p$  is  $q$
- ▶  $p$  implies  $q$
- ▶  $p$  only if  $q$
- ▶ a sufficient condition for  $q$  is  $p$
- ▶  $q$  whenever  $p$
- ▶  $q$  is necessary for  $p$
- ▶  $q$  follows from  $p$

# Conditional Statements

“No underage drinking”

If somebody is drinking alcohol, then they are over 18 y/o.

- ▶ A is drinking beer
- ▶ B is drinking coda
- ▶ C is 55 y/o
- ▶ D is 15 y/o

| $p$ | $q$ | $p \rightarrow q$ |
|-----|-----|-------------------|
| 0   | 0   | 1                 |
| 0   | 1   | 1                 |
| 1   | 0   | 0                 |
| 1   | 1   | 1                 |

## Remark

We seek to find out whether a certain promise or guarantee is kept. That is,

- ▶ either  $p$  is false,
- ▶ or  $q$  is true.

| $p$ | $q$ | $\neg p$ | $\neg p \vee q$ |
|-----|-----|----------|-----------------|
| 0   | 0   | 1        | 1               |
| 0   | 1   | 1        | 1               |
| 1   | 0   | 0        | 0               |
| 1   | 1   | 0        | 1               |

## Converse/Inverse/Contrapositive/Negation

Given  $p \rightarrow q$ ,

- ▶ Converse:  $q \rightarrow p$
- ▶ Inverse:  $\neg p \rightarrow \neg q$
- ▶ Contrapositive:  $\neg q \rightarrow \neg p$
- ▶ Negation:  $\neg(p \rightarrow q)$

### Remark

A proposition is equivalent to its contrapositive.

| $p$ | $q$ | $p \rightarrow q$ | $\neg q \rightarrow \neg p$ |
|-----|-----|-------------------|-----------------------------|
| 0   | 0   | 1                 | 1                           |
| 0   | 1   | 1                 | 1                           |
| 1   | 0   | 0                 | 0                           |
| 1   | 1   | 1                 | 1                           |

## Biconditional Statements

if and only if (iff)

| $p$ | $q$ | $p \rightarrow q$ | $q \rightarrow p$ | $p \leftrightarrow q$ |
|-----|-----|-------------------|-------------------|-----------------------|
| 0   | 0   | 1                 | 1                 | 1                     |
| 0   | 1   | 1                 | 0                 | 0                     |
| 1   | 0   | 0                 | 1                 | 0                     |
| 1   | 1   | 1                 | 1                 | 1                     |

## Compound proposition

Order of operations (use brackets “( . . . )” when in doubt)

- ▶  $\neg$
- ▶  $\wedge$
- ▶  $\vee$
- ▶  $\rightarrow$
- ▶  $\leftrightarrow$

# Tautology and Contradiction

In the truth table,

- ▶ Tautology: All cases evaluates to 1. (e.g.,  $p \vee \neg p$ )
- ▶ Contradiction: All cases evaluates to 0. (e.g.,  $p \wedge \neg p$ )

## Equivalence

$p$  and  $q$  are called **equivalent** iff  $p \leftrightarrow q$  is a tautology, denoted by  $p \Leftrightarrow q$ .

## Examples

- ▶  $p \vee \neg p \Leftrightarrow 1$
- ▶  $p \wedge \neg p \Leftrightarrow 0$
- ▶  $p \rightarrow q \Leftrightarrow \neg q \rightarrow \neg p \Leftrightarrow \neg p \vee q$

# Tautological Equivalence

## ► Commutativity

$$p \wedge q \Leftrightarrow q \wedge p$$

$$p \vee q \Leftrightarrow q \vee p$$

## ► Associativity

$$(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r)$$

$$(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$$

$$(p \leftrightarrow q) \leftrightarrow r \Leftrightarrow p \leftrightarrow (q \leftrightarrow r)$$

$$(p \oplus q) \oplus r \Leftrightarrow p \oplus (q \oplus r)$$

## ► Distributivity

$$p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$$

$$p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$$

# Tautological Equivalence

## ► Negation

$$\neg\neg p \Leftrightarrow p$$

$$p \rightarrow q \Leftrightarrow \neg q \rightarrow \neg p$$

$$\neg(p \vee q) \Leftrightarrow (\neg p) \wedge (\neg q)$$

$$\neg(p \wedge q) \Leftrightarrow (\neg p) \vee (\neg q)$$

## ► Identity

$$p \vee 0 \Leftrightarrow p \quad (\max\{p, 0\} = p)$$

$$p \wedge 1 \Leftrightarrow p \quad (\min\{p, 1\} = p)$$

## ► Null

$$p \wedge 0 \Leftrightarrow 0 \quad (\min\{p, 0\} = 0)$$

$$p \vee 1 \Leftrightarrow 1 \quad (\max\{p, 1\} = 1)$$

# Tautological Equivalence

- ▶ Idempotent

$$p \wedge p \Leftrightarrow p$$

$$p \vee p \Leftrightarrow p$$

- ▶ Absorption

$$p \wedge (p \vee q) \Leftrightarrow p$$

$$p \vee (p \wedge q) \Leftrightarrow p$$

- ▶ Cases

$$(p \rightarrow q) \wedge (p \rightarrow r) \Leftrightarrow p \rightarrow (q \wedge r)$$

$$(p \rightarrow q) \vee (p \rightarrow r) \Leftrightarrow p \rightarrow (q \vee r)$$

$$(p \rightarrow r) \wedge (q \rightarrow r) \Leftrightarrow (p \vee q) \rightarrow r$$

$$(p \rightarrow r) \vee (q \rightarrow r) \Leftrightarrow (p \wedge q) \rightarrow r$$

- ▶ Added premise

$$(p \wedge q) \rightarrow r \Leftrightarrow p \rightarrow (q \rightarrow r)$$

$$\Leftrightarrow q \rightarrow (p \rightarrow r)$$

## Added Premise

| $p$ | $q$ | $r$ | $p \wedge q$ | $r$ | $(p \wedge q) \rightarrow r$ | $p$ | $q \rightarrow r$ | $p \rightarrow (q \rightarrow r)$ |
|-----|-----|-----|--------------|-----|------------------------------|-----|-------------------|-----------------------------------|
| 0   | 0   | 0   | 0            | 0   | 1                            | 0   | 1                 | 1                                 |
| 0   | 0   | 1   | 0            | 1   | 1                            | 0   | 1                 | 1                                 |
| 0   | 1   | 0   | 0            | 0   | 1                            | 0   | 0                 | 1                                 |
| 0   | 1   | 1   | 0            | 1   | 1                            | 0   | 1                 | 1                                 |
| 1   | 0   | 0   | 0            | 0   | 1                            | 1   | 1                 | 1                                 |
| 1   | 0   | 1   | 0            | 1   | 1                            | 1   | 1                 | 1                                 |
| 1   | 1   | 1   | 1            | 1   | 1                            | 1   | 1                 | 1                                 |
| 1   | 1   | 0   | 1            | 0   | 0                            | 1   | 0                 | 0                                 |

# Tautological Equivalence

## Example

Prove the absorption rule  $p \wedge (p \vee q) \Leftrightarrow p$ .

Proof.

$$\begin{aligned} p \wedge (p \vee q) &\Leftrightarrow (p \vee 0) \wedge (p \vee q) \\ &\Leftrightarrow p \vee (0 \wedge q) \\ &\Leftrightarrow p \vee 0 \\ &\Leftrightarrow p \end{aligned}$$

□

## Remark

Consider the saturation function with parameter  $a, b \in \mathbb{R}$ ,  $a \leq b$ ,

$$\text{SAT}_{a,b}(x) := \begin{cases} a, & x < a \\ x, & a \leq x \leq b \\ b, & x > b \end{cases}$$

Note that  $\text{SAT}_{a,b}(x) = \min\{b, \max\{a, x\}\} = \max\{a, \min\{b, x\}\}$ .

## Tautological Equivalence

### Remark (Cont.)

Then we can write

$$p \wedge (p \vee q) = \min\{p, \max\{p, q\}\} = \text{SAT}_{p,p}(q) = p$$

or

$$p \wedge (p \vee q) = \min\{p, \max\{q, p\}\} = \text{SAT}_{q,p}(p) = p$$

### Remark

Note that

$$\text{SAT}_{a,b}(x) = \text{MEDIAN}(a, b, x)$$

# Tautological Equivalence

## Example

Show that  $(p \rightarrow r) \wedge (q \rightarrow r) \Leftrightarrow (p \vee q) \rightarrow r.$

Proof.

$$\begin{aligned}(p \rightarrow r) \wedge (q \rightarrow r) &\Leftrightarrow (\neg p \vee r) \wedge (\neg q \vee r) \\&\Leftrightarrow (\neg p \wedge \neg q) \vee r \\&\Leftrightarrow (\neg(p \vee q)) \vee r \\&\Leftrightarrow (p \vee q) \rightarrow r\end{aligned}$$

□

## Example

Show that  $(p \rightarrow r) \vee (q \rightarrow r) \Leftrightarrow (p \wedge q) \rightarrow r.$

Proof.

$$\begin{aligned}(p \rightarrow r) \vee (q \rightarrow r) &\Leftrightarrow (\neg p) \vee r \vee (\neg q) \vee r \\&\Leftrightarrow (\neg p) \vee (\neg q) \vee r \\&\Leftrightarrow (\neg(p \wedge q)) \vee r \\&\Leftrightarrow (p \wedge q) \rightarrow r\end{aligned}$$

□

# CNF and DNF

## Theorem

For any proposition  $\varphi$ , there is a proposition  $\varphi_{cnf}$  over the same Boolean variables and in CNF such that  $\varphi \Leftrightarrow \varphi_{cnf}$ .

## Example

- ▶  $\varphi = p \vee q \quad \varphi_{cnf} = (p \vee q)$
- ▶  $\varphi = p \wedge q \quad \varphi_{cnf} = (p) \wedge (q)$
- ▶  $\varphi = p \rightarrow q \quad \varphi_{cnf} = (\neg p \vee q)$
- ▶  $\varphi = p \leftrightarrow q \quad \varphi_{cnf} = (\neg p \vee q) \wedge (\neg q \vee p)$
- ▶  $\varphi = p \oplus q \quad \varphi_{cnf} = (p \vee q) \wedge (\neg q \vee \neg p)$

| $p$ | $q$ | $p \oplus q$ |
|-----|-----|--------------|
| 0   | 0   | 0            |
| 0   | 1   | 1            |
| 1   | 0   | 1            |
| 1   | 1   | 0            |

Since

$$\neg(\varphi_{cnf}) = (\neg p \wedge \neg q) \vee (p \wedge q)$$

Then by DeMorgan's law

$$\varphi_{cnf} = (p \vee q) \wedge (\neg q \vee \neg p)$$

# CNF and DNF

## Theorem

For any proposition  $\varphi$ , there is a proposition  $\varphi_{dnf}$  over the same Boolean variables and in DNF such that  $\varphi \Leftrightarrow \varphi_{dnf}$ .

## Example

- ▶  $\varphi = p \vee q \quad \varphi_{dnf} = (p) \vee (q)$
- ▶  $\varphi = p \wedge q \quad \varphi_{dnf} = (p \wedge q)$
- ▶  $\varphi = p \rightarrow q \quad \varphi_{dnf} = (\neg p) \vee (q)$
- ▶  $\varphi = p \leftrightarrow q \quad \varphi_{dnf} = (p \wedge q) \vee (\neg q \wedge \neg p)$
- ▶  $\varphi = p \oplus q \quad \varphi_{dnf} = (\neg p \wedge q) \vee (p \wedge \neg q)$

| $p$ | $q$ | $p \oplus q$ |
|-----|-----|--------------|
| 0   | 0   | 0            |
| 0   | 1   | 1            |
| 1   | 0   | 1            |
| 1   | 1   | 0            |

## Rules of Inference (Implication)

- ▶ Detachment (Modus ponens):  $(p \rightarrow q) \wedge p \Rightarrow q$
- ▶ Indirect reasoning (Modus tollens):  $(p \rightarrow q) \wedge \neg q \Rightarrow \neg p$
- ▶ Disjunctive addition:  $p \Rightarrow (p \vee q)$
- ▶ Conjunctive simplification:  $(p \wedge q) \Rightarrow p$
- ▶ Disjunctive syllogism:  $(p \vee q) \wedge \neg p \Rightarrow q$
- ▶ Hypothetical syllogism:  $(p \rightarrow q) \wedge (q \rightarrow r) \Rightarrow p \rightarrow r$
- ▶ Resolution:  $(p \vee q) \wedge (\neg p \vee r) \Rightarrow q \vee r$

## Proof by Contrapositive

$$p \rightarrow q \Leftrightarrow (\neg q \rightarrow \neg p)$$

## Proof by Contradiction

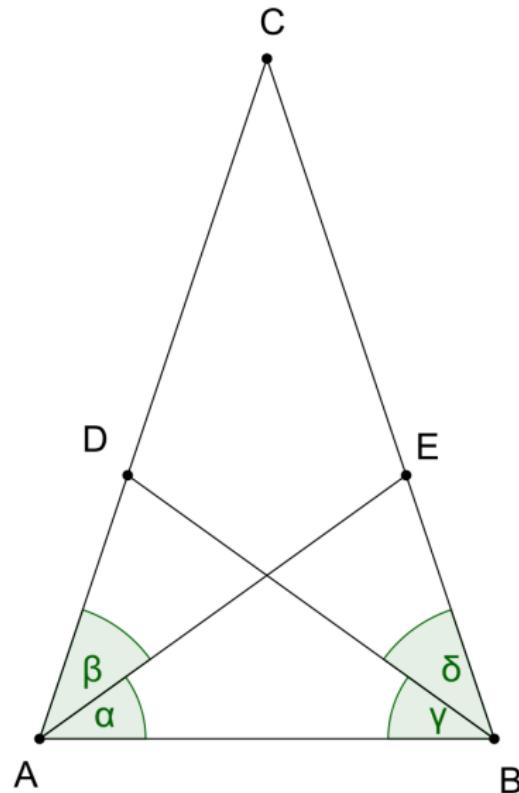
$$p \rightarrow q \Leftrightarrow (p \wedge \neg q) \rightarrow 0$$

# Direct Proofs Might Be Hard

Theorem (Steiner-Lehmus)

*Every triangle with two angle bisectors of equal lengths is isosceles.*

$$|AE| = |BD|, \alpha = \beta, \gamma = \delta \\ \Rightarrow \triangle ABC \text{ is isosceles.}$$



# The Natural Numbers

## Definition

Let  $m, n \in \mathbb{N}$  be natural numbers.

- (i) We say that  $n$  is **greater than or equal to**  $m$ , writing  $n \geq m$ , if there exists some  $k \in \mathbb{N}$  such that  $n = m + k$ . If we can choose  $k \neq 0$ , we say  $n$  is **greater than**  $m$  and write  $n > m$ .
- (ii) We say that  $m$  **divides**  $n$ , writing  $m | n$ , if there exists some  $k \in \mathbb{N}$  such that  $n = m \cdot k$ .
- (iii) If  $2 | n$ , we say that  $n$  is even.
- (iv) If there exists some  $k \in \mathbb{N}$  such that  $n = 2k + 1$ , we say that  $n$  is odd.
- (v) Suppose that  $n > 1$ . If there does not exist any  $k \in \mathbb{N}$  with  $1 < k < n$  such that  $k | n$ , we say that  $n$  is **prime**.

## Remark

It can be proven that every number is either even or odd and not both. We also assume this for the purposes of our examples.

# Infinitude of Primes

## Theorem

*There are infinitely many prime numbers.*

## Proof (NOT due to Euclid).

Assume that there are only finitely many primes, say  $\mathbb{P} = \{p_1, \dots, p_k\}$ . consider the integer  $N = p_1 p_2 \cdots p_k + 1$ , observe that  $p_i \nmid N$  for any  $i = 1, \dots, k$ , so  $N$  must be a prime, but  $N \notin \mathbb{P}$ , contradiction!

□

## Proof (by Euclid).

Consider a finite set of primes  $\{p_1, \dots, p_k\}$ . Let  $N = p_1 p_2 \cdots p_k + 1$ , so

- ▶ either  $N$  is a prime;
- ▶ or  $N$  is not a prime, so  $N$  must admit a prime factor, which is not in  $\{p_1, \dots, p_k\}$ . Call this new prime  $p_{k+1}$ .

So we can always generate a new prime from a finite set of primes.

□

## Remark

Euclid's proof of the infinitude of primes is **NOT** a proof by contradiction.

## Proof by Contradiction

Recall a proof by contradiction admits the form

$$p \rightarrow q \Leftrightarrow (p \wedge \neg q) \rightarrow 0$$

Specifically, for some proposition  $r$ ,

$$\begin{aligned} p \rightarrow q &\Leftrightarrow (p \wedge \neg q) \rightarrow 0 \\ &\Leftrightarrow (p \wedge \neg q) \rightarrow (r \wedge \neg r) \\ &\Leftrightarrow (p \wedge \neg q \rightarrow r) \wedge (p \wedge \neg q \rightarrow \neg r) \end{aligned}$$

What if  $r = q$ ?

$$\begin{aligned} p \rightarrow q &\Leftrightarrow (p \wedge \neg q) \rightarrow (q \wedge \neg q) \\ &\Leftrightarrow (p \wedge \neg q \rightarrow q) \wedge \underbrace{(p \wedge \neg q \rightarrow \neg q)}_{=1} \\ &\Leftrightarrow p \wedge \neg q \rightarrow q \\ &\Leftrightarrow \neg q \rightarrow (p \rightarrow q) \end{aligned}$$

# Statements

## Examples

- ▶ “ $3 > 2$ ” is a **true statement**.
- ▶ “ $x^3 > 10$ ” is not a statement, because we can not decide whether it is true or not.
- ▶ “the cube of any natural number is greater than 10” is a **false statement**.

The last example can be written using a **statement variable**  $n$ :

- ▶ “For any natural number  $n$ ,  $n^3 > 10$ ”

The first part of the statement is a **quantifier** (“for any natural number  $n$ ”), while the second part is called a **statement form** or **predicate** (“ $n^3 > 10$ ”).

# Statements

## Definition

A function  $P : X \rightarrow \{\top, \perp\}$  is called a **predicate** on its domain  $X$ .

## Remark

A statement form or predicate becomes a statement (which can then be either true or false) when the variable takes on a specific value.

## Example

If  $P(x)$  stands for " $x^3 > 10$ ", then

- ▶  $P(10) = \top$ , i.e.,  $10^3 > 10$  is a TRUE statement;
- ▶  $P(1) = \perp$ , i.e.,  $1^3 > 10$  is a FALSE statement.

Recall We indicate that an **element**  $x$  is a member of a **set**  $X$  by writing  $x \in X$ . We may characterize the elements of a set  $X$  by some predicate  $P$ :

$$x \in X \Leftrightarrow P(x).$$

We write  $X = \{x : P(x)\} = \{x \mid P(x)\}$ , which is called the **set-builder notation**.

# Logical Quantifiers

There are two types of quantifiers:

- ▶ the ***universal quantifier***, denoted by the symbol  $\forall$ , read as “for all” and
- ▶ the ***existential quantifier***, denoted by  $\exists$ , read as “there exists.”

## Definition

Let  $M$  be a set and  $A(x)$  be a predicate. Then we define the quantifier  $\forall$  by

$$(\forall x \in M)A(x) \Leftrightarrow A(x) \text{ is true for all } x \in M$$

We define the quantifier  $\exists$  by

$$(\exists x \in M)A(x) \Leftrightarrow A(x) \text{ is true for at least one } x \in M$$

We may also write  $\forall x \in M: A(x)$  or  $\bigvee_{x \in M} A(x)$  instead of  $(\forall x \in M)A(x)$ .

Similarly for  $\exists$ .

# Logical Quantifiers

We may also state the domain before making the statements, as in the following example.

## Examples

Let  $x$  be a real number. Then

- ▶  $\forall x: x > 0 \Rightarrow x^3 > 0$  is a true statement;
- ▶  $\forall x: x > 0 \Leftrightarrow x^2 > 0$  is a false statement;
- ▶  $\exists x: x > 0 \Leftrightarrow x^2 > 0$  is a true statement.

Sometimes mathematicians put a quantifier at the end of a statement form; this is known as a ***hanging quantifier***. Such a hanging quantifier will be interpreted as being located just before the statement form:

$$\exists y: y + x^2 > 0 \quad \forall x$$

is equivalent to  $\exists y \forall x: y + x^2 > 0$ .

## Contraposition and Negation of Quantifiers

We do not actually need the quantifier  $\exists$  since

$$\begin{aligned}\exists x \in M : A(x) &\Leftrightarrow A(x) \text{ is true for at least one } x \in M \\ &\Leftrightarrow A(x) \text{ is not false for all } x \in M \\ &\Leftrightarrow \neg \forall x \in M (\neg A(x))\end{aligned}\tag{1}$$

The equivalence (1) is called **contraposition of quantifiers**. It implies that the negation of  $\exists x \in M : A(x)$  is equivalent to  $\forall x \in M : \neg A(x)$ . For example,

$$\neg (\exists x \in \mathbb{R} : x^2 < 0) \Leftrightarrow \forall x \in \mathbb{R} : x^2 \not< 0.$$

Conversely,

$$\neg (\forall x \in M : A(x)) \Leftrightarrow \exists x \in M : \neg A(x).$$

## Vacuous Truth

If the domain of the universal quantifier  $\forall$  is the empty set  $M = \emptyset$ , then the statement  $\forall x \in M: A(x)$  is defined to be true regardless of the predicate  $A(x)$ . It is then said that  $A(x)$  is **vacuously true**.

### Example

Let  $M$  be the set of real numbers  $x$  such that  $x = x + 1$ . Then the statement

$$\forall_{x \in M} x > x$$

is true.

This convention reflects the philosophy that a universal statement is true unless there is a counterexample to prove it false. While this may seem a strange point of view, it proves useful in practice.

This is similar to saying that “All pink elephants can fly.” is a true statement, because it is impossible to find a pink elephant that can’t fly.

# Nesting Quantifiers

We can also treat predicates with more than one variable as shown in the following example.

## Examples

In the following examples,  $x, y$  are taken from the real numbers.

- ▶  $\forall x \forall y: x^2 + y^2 - 2xy \geq 0$  is equivalent to  $\forall y \forall x: x^2 + y^2 - 2xy \geq 0$ .  
Therefore, one often writes  $\forall x, y: x^2 + y^2 - 2xy \geq 0$ .
- ▶  $\exists x \exists y: x + y > 0$  is equivalent to  $\exists y \exists x: x + y > 0$ , often abbreviated to  $\exists x, y: x + y > 0$ .
- ▶  $\forall x \exists y: x + y > 0$  is a true statement.
- ▶  $\exists x \forall y: x + y > 0$  is a false statement.

As is clear from these examples, the order of the quantifiers is important if they are different.

## Set Theory Examples

- Extensionality Axiom

$$\forall A, B (\forall x (x \in A \Leftrightarrow x \in B) \Rightarrow A = B)$$

- Empty Set Axiom

$$\exists B \forall x (x \notin B)$$

- Pairing Axiom

$$\forall u, v \exists B \forall x (x \in B \Leftrightarrow x = u \vee x = v)$$

- Union Axiom

$$\forall a, b \exists B \forall x (x \in B \Leftrightarrow x \in a \vee x \in b)$$

- Powerset Axiom

$$\forall a \exists b \forall x (x \in B \Leftrightarrow x \subseteq a)$$

where  $x \subseteq a$  is shorthand for  $\forall t (t \in x \Rightarrow t \in a)$ .

# Set-Theoretic Proofs

## Basic Facts

$$x \in A \cup B \Leftrightarrow (x \in A) \vee (x \in B)$$

$$x \notin A \cup B \Leftrightarrow (x \notin A) \wedge (x \notin B)$$

$$x \in A \cap B \Leftrightarrow (x \in A) \wedge (x \in B)$$

$$x \notin A \cap B \Leftrightarrow (x \notin A) \vee (x \notin B)$$

$$x \in A - B \Leftrightarrow (x \in A) \wedge (x \notin B)$$

$$x \notin A - B \Leftrightarrow (x \notin A) \vee (x \in B)$$

$$A \subset B \Leftrightarrow (x \in A) \rightarrow (x \in B)$$

$$A = B \Leftrightarrow (A \subset B) \wedge (B \subset A)$$

## Remark

- ▶ Prove something exists: sufficient to find an example.
- ▶ Prove  $P(x)$  for all  $x \in A$ : Take any  $x \in A$  and continue. (or use induction if  $A = \mathbb{N}$ , more on this later.)

# Duality in Propositional Logic

- ▶  $\vee$  vs  $\wedge$
- ▶ 0 vs 1

| Basic Law  | Property       | Dual Law   |
|--|----------------|--|
| $p \vee q \Leftrightarrow q \vee p$                                  | Commutativity  | $p \wedge q \Leftrightarrow q \wedge p$                            |
| $(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$                | Associativity  | $(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r)$      |
| $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$ | Distributivity | $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ |
| $p \vee 0 \Leftrightarrow p$   | Identity       | $p \wedge 1 \Leftrightarrow p$                                     |
| $p \wedge \neg p \Leftrightarrow 0$                                  | Negation       | $p \vee \neg p \Leftrightarrow 1$                                  |
| $p \vee p \Leftrightarrow p$   | Idempotent     | $p \wedge p \Leftrightarrow p$                                     |
| $p \wedge 0 \Leftrightarrow 0$                                       | Null           | $p \vee 1 \Leftrightarrow 1$                                       |
| $p \wedge (p \vee q) \Leftrightarrow p$                              | Absorption     | $p \vee (p \wedge q) \Leftrightarrow p$                            |
| $\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$                | DeMorgan's     | $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$              |

# Duality in Set Theory

- ▶  $\cup$  vs  $\cap$
- ▶  $\emptyset$  vs  $U$  ( $U$  is the universe)

| Basic Law  | Property       | Dual Law   |
|--|----------------|--|
| $A \cup B = B \cup A$                            | Commutativity  | $A \cap B = B \cap A$                            |
| $(A \cup B) \cup C = A \cup (B \cup C)$          | Associativity  | $(A \cap B) \cap C = A \cap (B \cap C)$          |
| $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ | Distributivity | $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ |
| $A \cup \emptyset = A$                           | Identity       | $A \cap U = A$                                   |
| $A \cap A^c = \emptyset$                         | Negation       | $A \cup A^c = U$                                 |
| $A \cup A = A$                                   | Idempotent     | $A \cap A = A$                                   |
| $A \cap \emptyset = \emptyset$                   | Null           | $A \cup U = U$                                   |
| $A \cap (A \cup B) = A$                          | Absorption     | $A \cup (A \cap B) = A$                          |
| $(A \cup B)^c = A^c \cap B^c$                    | DeMorgan's     | $(A \cap B)^c = A^c \cup B^c$                    |

# Russell's Paradox

## Barber Paradox

The barber is the “one who shaves all those, and those only, who do not shave themselves”.

Question: does the barber shave himself?

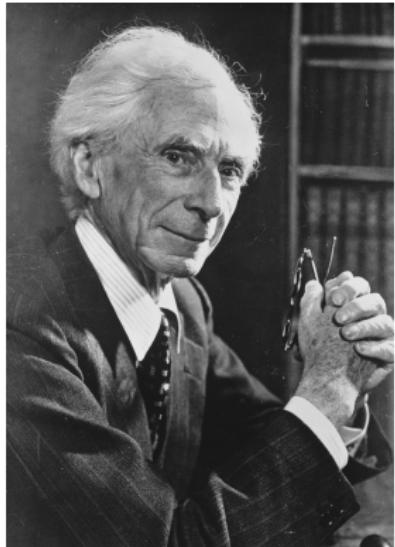
Consider the set of all sets that do not contain themselves:

$$S = \{A \mid A \text{ is a set, and } A \notin A\}$$

If such a set exists, then

- ▶  $S \in S \rightarrow S \notin S$
- ▶  $S \notin S \rightarrow S \in S$

Contradiction!



Bertrand Russell,  
Nov. 1957

# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Mathematical Induction

Typically one wants to show that some statement frame  $P(n)$  is true for all  $n \in \mathbb{N}$  with  $n \geq n_0$  for some  $n_0 \in \mathbb{N}$ . Mathematical induction works by establishing two statements:

- (I) the **base case**:  $P(n_0)$  is true.
- (II) the **inductive case**:  $P(n + 1)$  is true whenever  $P(n)$  is true for  $n \geq n_0$ , i.e.,

$$(\forall n \in \mathbb{N}, n \geq n_0)(P(n) \Rightarrow P(n + 1))$$

In the inductive case,  $P(n)$  is called **inductive hypothesis**, often abbreviated as **IH**.

Note that (II) does not make a statement on the situation when  $P(n)$  is false; it is permitted for  $P(n + 1)$  to be true even if  $P(n)$  is false.

The principle of mathematical induction now claims that  $P(n)$  is true for all  $n \geq n_0$  if (I) and (II) are true.

## Introductory Example

### Example

Consider the statement

$$\sum_{k=0}^n k = \frac{n(n+1)}{2} \quad \text{for all } n \in \mathbb{N}.$$

This is a typical example, in that  $P(n): \sum_{k=0}^n k = \frac{n(n+1)}{2}$  is a predicate which is to be shown to hold for all natural numbers  $n \in \mathbb{N}$ .

We first establish that  $P(0)$  is true:

$$\sum_{k=0}^0 k = 0 \quad \text{and} \quad \frac{0(0+1)}{2} = 0,$$

so  $P(0): 0 = 0$  is true.

## Introductory Example

We next show that  $P(n) \Rightarrow P(n + 1)$  for all  $n \in \mathbb{N} \setminus \{0\}$ . This means we show that  $\sum_{k=0}^{n+1} k = \frac{(n+1)(n+2)}{2}$  if  $\sum_{k=0}^n k = \frac{n(n+1)}{2}$ . Let  $n$  now be any  $n$  for which  $P(n)$  is true. We then write

$$\sum_{k=0}^{n+1} k = \left( \sum_{k=0}^n k \right) + n + 1$$

If  $P(n)$  is true for this specific  $n$ , we can replace the sum on the right by  $\frac{n(n+1)}{2}$ , yielding

$$\sum_{k=0}^{n+1} k = \frac{n(n+1)}{2} + n + 1 = \frac{(n+1)(n+2)}{2}$$

But this is just the statement  $P(n + 1)$ . Therefore, if  $P(n)$  is true, then  $P(n + 1)$  will also be true. We have shown that  $P(n) \Rightarrow P(n + 1)$ . □

# Application of Induction

We can use induction to prove the *efficiency* and *correctness* of a recursive algorithm.

## Properties of Algorithms

- ▶ ***Input.*** An algorithm has input values from a specified set;
- ▶ ***Output.*** For given input, the algorithm produces output values from a specified set;
- ▶ ***Definiteness.*** The steps of the algorithm are defined precisely;
- ▶ ***Correctness.*** For each input, the algorithm produces the correct output values;
- ▶ ***Finiteness.*** For given input, the algorithm produces output after a finite number of steps;
- ▶ ***Effectiveness.*** Each step of the algorithm can be performed exactly;
- ▶ ***Generality.*** The algorithm is generally applicable, not just for certain input values.

# Factorial

---

**Input:**  $n$ , a positive integer

**Output:**  $n!$

```
1 Function fact(n):  
2   if n = 1 then  
3   |   return 1  
4   else  
5   |   return n · fact(n - 1)  
6   end  
7 end
```

---

## Remark

Recursion is inefficient. Use iteration/tail recursion/memoization instead (smart compiler can do this automatically).

## Correctness of fact

Proof.

- ▶ **base case ( $n = 1$ ):** Observe that  $\text{fact}(1)$  returns 1 immediately, and  $1! = 1$ .
- ▶ **inductive case ( $n \geq 1$ ):** Assume that  $\text{fact}(n)$  returns  $n!$ . We want to show that  $\text{fact}(n + 1)$  returns  $(n + 1)!$ . Indeed, by induction hypothesis,

$$\text{fact}(n + 1) = n \cdot \text{fact}(n) = (n + 1) \cdot n! = (n + 1)!$$

Therefore the recursive algorithm `fact` is correct by induction. □

# Insertion Sort

---

**Input:**  $A[1 \dots n] = \langle a_1, \dots, a_n \rangle$ ,  $n$  unsorted elements

**Output:** all the  $a_i, 1 \leq i \leq n$  in increasing order

```
1 Function insertionSort( $A[1 \dots n]$ ,  $n$ ):
2   if  $n = 1$  then
3     return  $A[n]$ 
4   else
5     insertionSort( $A[1 \dots n-1]$ ,  $n-1$ );
6     key  $\leftarrow A[n]$ ;  $i \leftarrow n - 1$ ;
7     while  $i > 0$  and  $A[i] > \text{key}$  do
8       |  $A[i + 1] \leftarrow A[i]$ ;  $i \leftarrow i - 1$ ;
9     end
10     $A[i + 1] \leftarrow \text{key}$ ;
11    return  $A[1 \dots n]$ ;
12  end
13 end
```

---

# Selection Sort

---

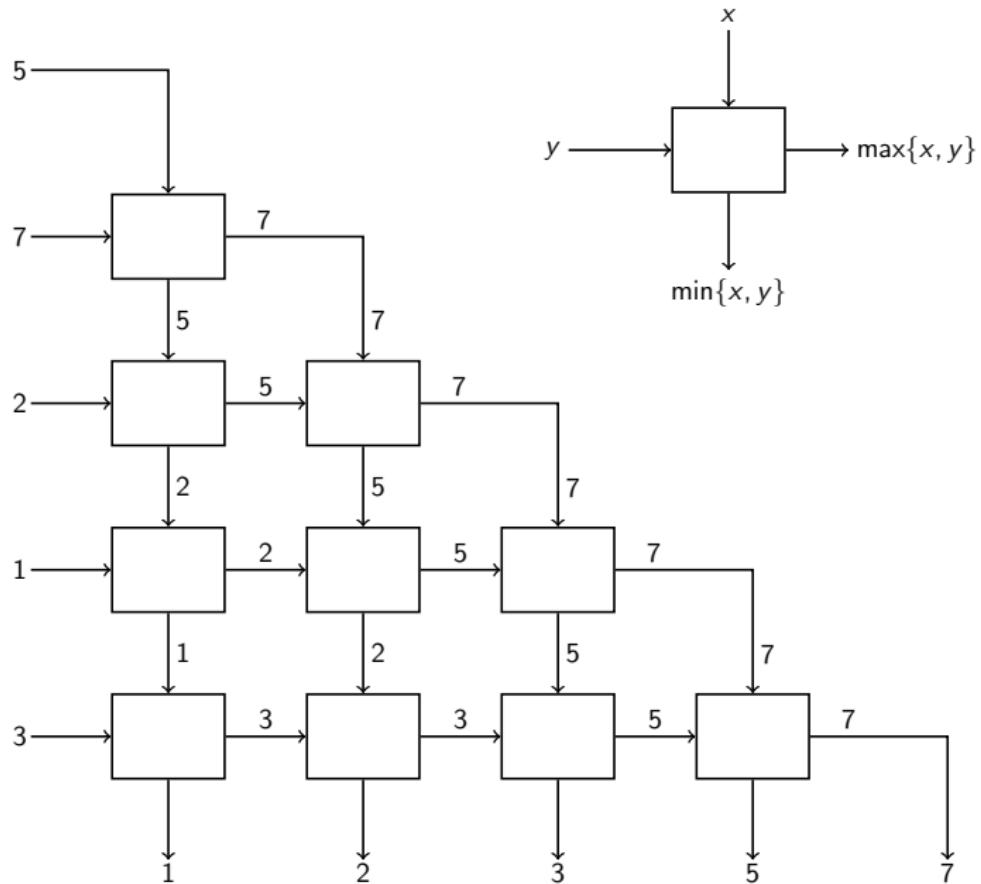
**Input:**  $A[1 \dots n] = \langle a_1, \dots, a_n \rangle$ ,  $n$  unsorted elements

**Output:** all the  $a_i, 1 \leq i \leq n$  in increasing order

```
1 Function selectionSort( $A[1 \dots n]$ ,  $n$ ):
2     if  $n = 1$  then
3         return  $A[n]$ 
4     else
5          $indmax \leftarrow \text{findMaxIndex}(A[1 \dots n], n);$ 
6         swap( $A[n]$ ,  $A[indmax]$ );
7         selectionSort( $A[1 \dots n - 1]$ ,  $n - 1$ )
8     end
9 end
```

---

# Insertion Sort vs Selection Sort



## Correctness of insertionSort and selectionSort

Proof. (Correctness of insertionSort).

- ▶ **base case ( $n = 1$ ):** Trivial since any array of length 1 is sorted.
- ▶ **inductive case ( $n \geq 1$ ):** Assume that  $\text{insertionSort}(A[1 \dots n], n)$  is sorted. We want to show that  $\text{insertionSort}(\langle A[1 \dots n + 1] \rangle, n + 1)$  is also sorted, which is true since  $A[n + 1]$  is inserted into the sorted  $\text{insertionSort}(A[1 \dots n], n)$  to make this happen.

□

Proof. (Correctness of selectionSort).

- ▶ **base case ( $n = 1$ ):** Trivial since any array of length 1 is sorted.
- ▶ **inductive case ( $n \geq 1$ ):** Assume that  $\text{selectionSort}(A[1 \dots n], n)$  is sorted. We want to show that  $\text{selecttSort}(\langle A[1 \dots n + 1] \rangle, n + 1)$  is also sorted. Since  $A[\text{indexmax}] \geq A[i]$  for all  $i < n + 1$ , then after swap, we have  $A[n + 1] \geq A[i]$  for all  $i < n$ . By induction hypothesis,  $A[1 \dots n]$  is already sorted, hence  $A[1 \dots n + 1]$  is sorted.

□

# Fibonacci Numbers

## Example

The Fibonacci Numbers is defined by

$$\begin{aligned}f_0 &= 0, \quad f_1 = 1, \\f_n &= f_{n-1} + f_{n-2}, \quad n \geq 2\end{aligned}$$

Show that

$$f_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right]$$

## Strong (Complete) Induction

The method of induction can be strengthened. We can replace

- (I)  $P(n_0)$  is true.
- (II)  $P(n + 1)$  is true whenever  $P(n)$  is true for  $n \geq n_0$ .

with

- (I)  $P(n_0)$  is true.
- (II')  $P(n + 1)$  is true whenever all the statements  
 $P(n_0), P(n_0 + 1), \dots, P(n)$  are true.

# Fibonacci Numbers

## Proof.

We prove the formula for Fibonacci Numbers by strong induction on  $n$ .

First note that we could let  $\varphi = \frac{1+\sqrt{5}}{2}$ , then we need to show that

$$f_n = \frac{\varphi^n - (-\varphi)^{-n}}{\sqrt{5}}.$$

► **base case ( $n = 0$ ):**

$$\frac{\varphi^0 - (-\varphi)^0}{\sqrt{5}} = 0 = f_0$$

► **base case ( $n = 1$ ):**

$$\frac{\varphi^1 - (-\varphi)^{-1}}{\sqrt{5}} = \dots = 1 = f_1$$

# Fibonacci Numbers

## Proof (Cont.)

- **inductive case ( $n \geq 2$ ):** Assume that the formula holds for all  $k < n$ , then note that  $\varphi^2 = \varphi + 1$ ,

$$\begin{aligned}f_n &= f_{n-1} + f_{n-2} \\&= \frac{\varphi^{n-1} - (-\varphi)^{-(n-1)}}{\sqrt{5}} + \frac{\varphi^{n-2} - (-\varphi)^{-(n-2)}}{\sqrt{5}} \\&= \frac{\varphi^{n-2}(\varphi + 1) - (-\varphi)^{-(n-1)}(1 - \varphi^{-1})}{\sqrt{5}} \\&= \frac{\varphi^n - (-\varphi)^{-n}}{\sqrt{5}}\end{aligned}$$

□

# Prime Factorization

## Theorem (Fundamental Theorem of Arithmetic)

Let  $n \in \mathbb{N} \setminus \{0\}$ , then there exist  $k \geq 0$  prime numbers  $p_1, p_2, \dots, p_k$  such that  $n = \prod_{i=1}^k p_i$ . (also unique up to order, more on this later.)

Proof by strong induction on  $n$ .

- ▶ **base case ( $n = 1$ ):** by convention, 1 is the product of zero prime numbers.
- ▶ **inductive case ( $n \geq 2$ ):** assume the statement is true for all positive integer  $n'$  where  $1 \leq n' \leq n - 1$ .
  - ▶ If  $n$  is prime, which is the product of 1 prime number.
  - ▶ If  $n$  is composite, then by definition there exist positive integers  $a$  and  $b$  such that  $n = a \cdot b$ , with  $2 \leq a, b \leq n - 1$ . By inductive hypotheses, we have  $a = q_1 \cdot q_2 \cdots q_\ell$  and  $b = r_1 \cdot r_2 \cdots r_m$  for prime numbers  $q_1, \dots, q_\ell$  and  $r_1, \dots, r_m$ . Therefore  $n = a \cdot b = q_1 \cdot q_2 \cdots q_\ell \cdot r_1 \cdot r_2 \cdots r_m$  which is the product of  $\ell + m$  prime numbers.



# Prime Factorization

---

```
1 Function primeFactor(n):  
2     if n = 1 then  
3         return ⟨⟩  
4     else  
5         if n is prime then  
6             return ⟨n⟩  
7         else  
8             find factors a, b where  $2 \leq a, b \leq n - 1$  such that  
9                  $n = a \cdot b$ .  
10            ⟨q1, ..., ql⟩ ← primeFactor(a);  
11            ⟨r1, ..., rm⟩ ← primeFactor(b);  
12            return ⟨q1, ..., ql, r1, ..., rm⟩  
13        end  
14    end
```

---

# Quick Sort

---

**Input:**  $A[1 \dots n] = \langle a_1, \dots, a_n \rangle$ ,  $n$  unsorted elements

**Output:** all the  $a_i, 1 \leq i \leq n$  in increasing order

```
1 Function quickSort( $A[1 \dots n]$ ,  $n$ ):
2   if  $n \leq 1$  then
3     return  $A[n]$ 
4   else
5     choose pivot  $\in \{1, \dots, n\}$ ;
6      $L := \langle \rangle$ ;  $R := \langle \rangle$ ;
7     for  $i \in \{1, \dots, n\}$  with  $i \neq \text{pivot}$  do
8       if  $A[i] < A[\text{pivot}]$  then
9          $L \leftarrow L + A[i]$ 
10      else
11         $R \leftarrow R + A[i]$ 
12      end
13       $L \leftarrow \text{quickSort}(L)$ ;
14       $R \leftarrow \text{quickSort}(R)$ ;
15      return  $L + \langle A[\text{pivot}] \rangle + R$ 
16    end
17  end
18 end
```

---

## Quick Sort

Proof. (Correctness of quickSort).

Let  $P(n)$  denote the claim that  $\text{quickSort}(A[1 \dots n])$  returns a sorted array. For simplicity, we assume that the elements in the array are distinct. We will prove  $P(n)$  for all  $n \geq 0$  by strong induction on  $n$ .

- ▶ **base case ( $n = 0, 1$ ):** done b/c any array of length 0 or 1 are already sorted.
- ▶ **inductive case ( $n \geq 2$ ):** assume  $P(0), \dots, P(n-1)$  that for any array  $B[1 \dots k]$  with distinct elements and  $k < n$ ,  $\text{quickSort}(B[1 \dots k])$  returns a sorted array. Let  $A[1 \dots n]$  be an arbitrary array with distinct elements. Let  $pivot \in \{1, \dots, n\}$  be arbitrary. **We need to show that  $x$  appears before  $y$  in  $\text{quicksort}(A[1 \dots n])$  iff  $x < y$ .**

# Quick Sort

Proof. (Correctness of quickSort Cont.)

**Inductive case ( $n \geq 2$ ):**

- ▶ Case 1.  $x = A[\text{pivot}]$ . By  $\text{quickSort}(A[1 \dots n])$ ,  $y \in R$  iff  $x < y$ .
- ▶ Case 2.  $y = A[\text{pivot}]$ . Similar to Case 1.
- ▶ Case 3.  $x, y < A[\text{pivot}]$ . Now  $x, y \in L$ . Since  $A[\text{pivot}] \notin L$ , thus  $L$  is of at most length  $n - 1$ . Hence by IH  $x$  appears before  $y$  in  $\text{quickSort}(L)$  iff  $x < y$ . Furthermore  $x$  appears before  $y$  in  $\text{quickSort}(A[1 \dots n])$  iff  $x < y$ .
- ▶ Case 4.  $x, y > A[\text{pivot}]$ . Similar to Case 3.
- ▶ Case 5.  $x < A[\text{pivot}] < y$ . Trivial.
- ▶ Case 6.  $y < A[\text{pivot}] < x$ . Impossible. □

|    | $L$    | $\text{pivot}$ | $R$    |
|----|--------|----------------|--------|
| 1. |        | $x$            |        |
| 2. |        | $y$            |        |
| 3. | $x, y$ |                |        |
| 4. |        |                | $x, y$ |
| 5. | $x$    |                | $y$    |
| 6. | $y$    |                | $x$    |

$$L = A[1 \dots \text{pivot} - 1]$$

$$R = A[\text{pivot} + 1 \dots n]$$

## Recursive Definitions and the Factorial

Similar to induction, we could make **recursive definitions**. For example, we wish to define a function (to be called the **factorial**)

$$(\cdot)! : \mathbb{N} \rightarrow \mathbb{N}$$

having the properties that

$$0! := 1, \quad n! := n \cdot (n - 1)!, \quad n \in \mathbb{N} \setminus \{0\}.$$

This is an example of a **recursive definition** and we may ask whether such a definition “makes sense”, i.e., whether such a function **exists** and is **unique**.

In the present case, the function is simply

$$n! := \prod_{k=1}^n k, \quad n \in \mathbb{N} \setminus \{0\},$$

This is called a **closed formula** for  $n!$ .

## Recursive Definitions

Recursive definitions often occur naturally in the formulation of a problem, and finding a closed formula can be extremely difficult. In some situations, a closed formula is highly desirable, while at other times, important properties are best expressed through recursive expressions.

For example, there exists a continuous extension of the factorial, given by the **Euler gamma function**, defined for  $t > 0$ ,

$$\Gamma(t) := \int_0^{\infty} z^{t-1} e^{-z} dz, \quad t > 0.$$

It is possible to show that  $\Gamma(1) = 1$  and that

$$\Gamma(t+1) = t\Gamma(t) = t\Gamma((t-1)+1) \quad t > 0.$$

we see that  $\Gamma(n+1) = n!$  for all  $n \in \mathbb{N}$ . Furthermore, we can define functions not just based on their preceding value, but on several such values. For example, The **Fibonacci sequence**.

## Recursive Definitions of Sets

In the same manner, we can define subsets of  $\mathbb{N}$  recursively. For example, consider the set  $S \subset \mathbb{N}$  such that

$$3 \in S \quad \text{and} \quad x, y \in S \Rightarrow x + y \in S. \quad (2)$$

We know that  $3 \in S$ , so  $3 + 3 = 6 \in S$ ,  $3 + 6 = 9 \in S$ ,  $6 + 6 = 12 \in S$  and so on. Our goal is to prove that

$$S = \{n \in \mathbb{N} \mid \exists k \in \mathbb{N} \setminus \{0\}: n = 3k\}.$$

However, this requires a little preparation.

# Recursively Defined Structures

David Liben-Nowell, Connecting Discrete Mathematics and Computer Science, manuscript at [www.cs.carleton.edu/faculty/dln/book/](http://www.cs.carleton.edu/faculty/dln/book/)

## Linked Lists

A ***linked list*** is either;

- ▶  $\langle \rangle$ , known as the ***empty list***; or
- ▶  $\langle x, L \rangle$ , where  $x$  is an arbitrary element and  $L$  is a linked list.

## Binary trees

A ***binary tree*** is either:

- ▶ the empty tree, denoted by `null`; or
- ▶ a root node  $x$ , a ***left subtree***  $T_\ell$ , and a ***right subtree***  $T_r$ , where  $x$  is an arbitrary value and  $T_\ell$  and  $T_r$  are both binary trees.

# Recursively Defined Structures

## Arithmetic Expressions

An **arithmetic expression** is any of the following:

- ▶ any integer  $n$ ;
- ▶  $-E$ , where  $E$  is an arithmetic expression; or
- ▶ where  $E$  and  $F$  are arithmetic expressions and  $\odot \in \{+, -, \cdot, /\}$  is an **operator**.

## Sentences of propositional logic

A **sentence of propositional logic** (also known as a **well-formed formula**, or **wff**) over the propositional variables  $X$  is one of the following

- ▶  $x$ , for some  $x \in X$ ;
- ▶  $\neg P$ , where  $P$  is a wff over  $X$ ; or
- ▶  $P \vee Q$ ,  $P \wedge Q$ , or  $P \rightarrow Q$ , where  $P$  and  $Q$  are wffs over  $X$ .

# Recursively Defined Structures

Natural numbers, recursively defined

A **natural numbers** is either:

- ▶ zero, denoted by 0; or
- ▶ the successor of a natural number  $n$ , denoted by  $\text{succ}(n)$  or  $n^+$  for a natural numbers  $n$ .

Theorem (Recursion on  $\mathbb{N}$ )

Let  $A$  be a set,  $a \in A$ , and  $F : A \rightarrow A$ . Then there **exists** a **unique** function  $h : \mathbb{N} \rightarrow A$  such that

- ▶  $h(0) = a$ ,
- ▶  $h(n^+) = F(h(n))$ ,  $\forall n \in \mathbb{N}$ .

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{n \mapsto n^+} & \mathbb{N} \\ h \downarrow & & \downarrow h \\ A & \xrightarrow[F]{} & A \end{array}$$

# Recursion theorem on $\mathbb{N}$

## Example

For fixed  $k \in \mathbb{N}$ , consider the function

$A_k : \mathbb{N} \rightarrow \mathbb{N}$  given by

- ▶  $A_k(0) := k$
- ▶  $A_k(n^+) := A_k(n)^+$

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{n \mapsto n^+} & \mathbb{N} \\ A_k \downarrow & & \downarrow A_k \\ \mathbb{N} & \xrightarrow{n \mapsto n^+} & \mathbb{N} \end{array}$$

## Example

For fixed  $k \in \mathbb{N}$ , consider the function

$M_k : \mathbb{N} \rightarrow \mathbb{N}$  given by

- ▶  $M_k(0) := 0$
- ▶  $M_k(n^+) := A_k(M_k(n))$

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{n \mapsto n^+} & \mathbb{N} \\ M_k \downarrow & & \downarrow M_k \\ \mathbb{N} & \xrightarrow{A_k} & \mathbb{N} \end{array}$$

# Arithmetic and Order on $\mathbb{N}$

## Definition

- ▶  $n + k := A_k(n)$
- ▶  $n \times k := M_k(n)$
- ▶  $n < m$  if  $n \in m$

## Natural Numbers $\mathbb{N}$

A natural number is either

- ▶  $0 := \{\} = \emptyset$ , or
- ▶  $n + 1 = n^+ := n \cup \{n\}$ ,  $n \in \mathbb{N}$ .

## Usual laws (to be verified)

- ▶ Associative law for addition:  $(a + b) + c = a + (b + c)$
- ▶ Commutative law for addition:  $a + b = b + a$
- ▶ Distributive law:  $a(b + c) = ab + ac$
- ▶ Associative law for multiplication:  $(ab)c = a(bc)$
- ▶ Commutative law for multiplication:  $ab = ba$
- ▶ Order preserved by addition:  $a + c < b + c$
- ▶ Order preserved by multiplication:  $ac < bc$  if  $a < b$  and  $c \neq 0$

## Structural Induction

Structural induction is a useful variant of induction that allows us to prove properties for recursively defined objects.

Structural induction establishes a statement on a recursively defined set in two steps. We call those elements specifically included in the set the basis elements of the set.

1. Establish the statement for the basis elements.
2. Show that if the statement is true for each of the elements used to construct new elements in the recursive step of the definition, the statement holds for these new elements.

The justification for structural induction lies in ordinary induction, applied to the statement

$P(n)$ : The claim is true for all elements of the set generated with  $n$  or fewer applications of the recursive rules for the set.

Structural induction first establishes  $P(0)$  and then  $P(n) \Rightarrow P(n + 1)$ .

# Explicit Representation of Recursively Defined Sets

## Example

Define  $S \subset \mathbb{N}$  to be the set such that

- (i)  $3 \in S$ ,
- (ii)  $x, y \in S \Rightarrow x + y \in S$ .

Let

$$T = \{n \in \mathbb{N} \mid \exists k \in \mathbb{N} \setminus \{0\} \text{ such that } n = 3k\}$$

We want to show that  $S = T$ .

## Proof.

First, we show  $S \subset T$  by structural induction:  $3 = 3 \cdot 1 \in T$ , so the base case is established. Now for  $x, y \in S$  suppose that  $x, y \in T$  so that  $x = 3k$  and  $y = 3k'$  for  $k, k' \in \mathbb{N} \setminus \{0\}$ . Then

$$x + y = 3k + 3k' = 3(k + k')$$

so  $x + y \in T$ . This shows that  $S \subset T$ .

## Explicit Representation of Recursively Defined Sets

### Proof (Cont.)

Next, we show  $T \subset S$  by (ordinary) induction. We claim that

$$\forall k \in \mathbb{N} \setminus \{0\} : 3k \in S.$$

For  $k = 1$ ,  $3k = 3 \cdot 1 = 3 \in S$ , so the base case is established. Now suppose that  $3k \in S$ . Since  $3 \in S$  by definition, we can apply the recursive rule for  $S$  to deduce that

$$3(k + 1) = 3k + 3 \in S.$$

This shows that  $3(k + 1) \in S$  if  $3k \in S$ . By the structural induction principle,  $3k \in S$  for all  $k \in \mathbb{N} \setminus \{0\}$ . This established  $T \subset S$ .

We finally conclude that  $S = T$ . □



# Propositional Logic Using $\neg$ and $\wedge$

## Example

For any logical proposition  $\varphi$  using the connectives  $\{\neg, \wedge, \vee, \rightarrow\}$ , there exists a proposition using only  $\{\neg, \wedge\}$  that is logically equivalent to  $\varphi$ .

## Proof by structural induction.

For a logical proposition  $\varphi$ , let  $A(\varphi)$  denote the property that there exists a  $\{\neg, \wedge\}$ -only proposition logically equivalent to  $\varphi$ . We'll prove by structural induction on  $\varphi$  that  $A(\varphi)$  holds for any well-formed formula  $\varphi$ .

- ▶ **base case:**  $\varphi$  is a variable, say  $\varphi = x$ . No need for connectives from the set  $\{\neg, \wedge\}$ ,  $A(\varphi)$  is vacuously true.
- ▶ **inductive case I:**  $\varphi$  is a negation, say  $\varphi = \neg p$ . Assume the IH  $A(p)$ , we need to show  $A(\neg p)$ . By IH, there exists a  $\{\neg, \wedge\}$ -only proposition  $q \Leftrightarrow p$ . Since  $\neg q \Leftrightarrow \neg p$ , and  $\neg q$  contains only connectives from  $\{\neg, \wedge\}$ , therefore  $A(\neg p)$  follows.

## Propositional Logic Using $\neg$ and $\wedge$

Proof by structural induction, Cont.

- **inductive case II:**  $\varphi$  is a conjunction, disjunction, or implication, say  $\varphi = p_1 \wedge p_2$ ,  $\varphi = p_1 \vee p_2$ ,  $\varphi = p_1 \rightarrow p_2$ . Assume IH  $A(p_1)$  and  $A(p_2)$ , that is, there exist  $\{\neg, \wedge\}$ -only propositions  $q_1$  and  $q_2$  such that  $q_1 \Leftrightarrow p_1$  and  $q_2 \Leftrightarrow p_2$ . We need to show  $A(p_1 \wedge p_2)$ ,  $A(p_1 \vee p_2)$ , and  $A(p_1 \rightarrow p_2)$ . Indeed, note that

$$p_1 \wedge p_2 \Leftrightarrow q_1 \wedge q_2$$

$$p_1 \vee p_2 \Leftrightarrow q_1 \vee q_2 \Leftrightarrow \neg(\neg q_1 \wedge \neg q_2)$$

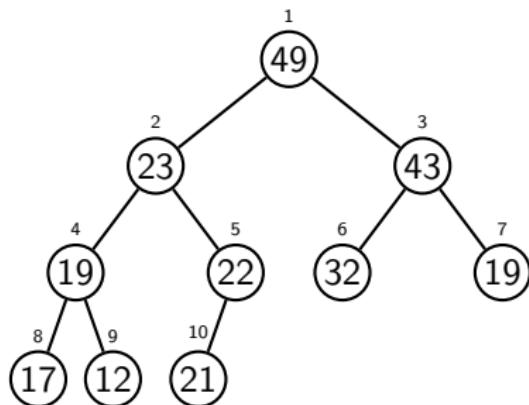
$$p_1 \rightarrow p_2 \Leftrightarrow q_1 \rightarrow q_2 \Leftrightarrow \neg(q_1 \wedge \neg q_2)$$

Since  $q_1$  and  $q_2$  are  $\{\neg, \wedge\}$ -only,  $A(p_1 \wedge p_2)$ ,  $A(p_1 \vee p_2)$ , and  $A(p_1 \rightarrow p_2)$  follow. □

# Heap

A **heap** (not “the heap”) is a data structure that is used to represent a collection of items, each of which has an associated **priority** (e.g.,  $<$ ,  $>$ ,  $\leq$ ,  $\geq$ , or a partial order  $\preceq$  in general). A heap can be represented as

- ▶ a complete (almost) binary tree with no “holes” as you read in left-to-right, top-to-bottom order; or
- ▶ stored more efficiently as an array, in which the elements are stored in that same left-to-right and top-to-bottom order.



|    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
| 49 | 23 | 43 | 19 | 22 | 32 | 19 | 17 | 12 | 21 |

# Heap

## Max Heaps

A heap is either

- ▶ an empty tree, denoted by null; or
  - ▶ a root node  $x$ , a left subtree  $T_\ell$ , and a right subtree  $T_r$ , where all elements of  $T_\ell$  and  $T_r$  are less than or equal to  $x$ , and  $T_\ell$  and  $T_r$  are both heaps.
- 

**Input:**  $A[top \dots bot] = \langle a_{top}, \dots, a_{bot} \rangle = T$

**Output:** all the  $a_i, 1 \leq i \leq n$  as a max heap

1 **Function** `buildHeap(A[top...bot], top):`

2   **if** `2top > bot` **then**

3     **return**

4   **else**

5     `buildHeap(A[top, bot], 2top);`                           // make heap of left child

6     `buildHeap(A[top, bot], 2top + 1);`                   // make heap of right child

7     `heapify(A[top, bot], top);`                           // make whole thing a heap

8   **end**

9 **end**

---

# Do You Want to Build a Max Heap

---

**Input:**  $A[\text{top} \dots \text{bot}] = \langle a_{\text{top}}, T_\ell, T_r \rangle = T$ ,  $a_{\text{top}}$  root,  $T_\ell$ ,  $T_r$  heaps

**Output:** all the  $a_i, 1 \leq i \leq n$  as a max heap

```
1 Function heapify(A[top...bot], top):
2     if 2top > bot then
3         return                                // no child
4     else if 2top = bot then                // only one child on the left
5         if a[top] > a[2top] then swap(a[top], a[bot])
6     else                                // find largest among root and two children
7         if a[top] > a[2top] and a[top] > a[2top + 1] then
8             return                            // already a heap, done
9         else if a[2top] > a[2top + 1] then          // left child largest
10            swap(a[top], a[2top])           // move left child up to the top
11            heapify(A[top...bot], 2top)      // build the rest of the heap
12        else                                // right child largest
13            swap(a[top], a[2top + 1])    // move right child up to the top
14            heapify(A[top...bot], 2top + 1) // build the rest of the heap
15    end
16 end
17 end
```

---

## Correctness of heapify

Proof by structural induction.

Consider the binary tree  $T$  with root  $a[\text{top}]$ ,  $T_\ell$  and  $T_r$  both heaps.

- ▶ **base case:**  $T_\ell = T_r = \text{null}$ : is a heap, b/c children are null.
- ▶ **base case:**  $T_\ell \neq \text{null}$ ,  $T_r = \text{null}$ : either is a heap, or swap  $a[\text{top}]$  with  $a[2\text{top}]$ , done.
- ▶ **inductive case:**  $T_\ell \neq \text{null}$  and  $T_r \neq \text{null}$ : assume the IH that  $\text{heapify}(T_\ell, 2\text{top})$  and  $\text{heapify}(T_r, 2\text{top} + 1)$  returns heaps. Then either
  - ▶  $a[\text{top}]$  is larger than both  $a[2\text{top}]$  and  $a[2\text{top} + 1]$ , we are done (b/c  $T_\ell$  and  $T_r$  are heaps, so are their sub-trees).
  - ▶ or swap  $a[\text{top}]$  with  $a[2\text{top}]$  (left child largest), then  $\text{heapify}(T_\ell, 2\text{top})$  is a heap by IH, and  $T_r$  is a heap by assumption, both have smaller elements.
  - ▶ or swap  $a[\text{top}]$  with  $a[2\text{top} + 1]$  (right child largest), then  $\text{heapify}(T_\ell, 2\text{top} + 1)$  is a heap by IH, and  $T_\ell$  is a heap by assumption, both have smaller elements.



## Correctness of buildHeap

Proof by structural induction.

- ▶ **base case:**  $T_\ell = T_r = \text{null}$ : already a heap, done.
- ▶ **inductive case:** either  $T_\ell \neq \text{null}$  or  $T_r \neq \text{null}$ : assume the IH that  $\text{buildHeap}(T_\ell)$  and  $\text{buildHeap}(T_r)$  return heaps. Apply heapify and we are done. □

### Remark

For iterative algorithms, the proofs of correctness often requires something called “loop invariant”.

## Weak Induction as Special Case of Structural Induction

Note that natural numbers  $\mathbb{N}$  can be recursively defined, i.e., a natural number is either

- ▶ 0, or
- ▶ the successor of a natural number  $n$ , denoted by  $\text{succ}(n)$  or  $n^+$  for a natural numbers  $n$ .

Under this definition, a proof of  $\forall n \in \mathbb{N} : P(n)$  by structural induction and a proof of  $\forall n \in \mathbb{N} : P(n)$  are identical:

- ▶ they have precisely the same **base case**: prove  $P(0)$ ; and
- ▶ they have precisely the same **inductive case**: prove  $P(n) \Rightarrow P(n^+)$ , i.e., prove that  $P(n) \Rightarrow P(n + 1)$ .

# Well-Ordering Principle (WOP)

## Theorem (Well-Ordering Principle)

*Every nonempty collection of natural numbers has a least element.*

### Proof by induction.

We'll prove the contrapositive, namely, that if a collection of natural numbers has no least element, then it must be empty.

Let  $T$  be such a nonempty set of natural numbers. Consider  $S = \mathbb{N} \setminus T$ , since  $T \neq \emptyset$ , then  $S \neq \mathbb{N}$ . We need to show that  $n \in S$  for all  $n \in \mathbb{N}$ .

- ▶ **base case ( $n = 0$ ):** If  $0 \in T$ , then 0 is the least element. So  $0 \notin T$ , i.e.,  $0 \in S$ .
- ▶ **inductive case ( $n \geq 1$ ):** Suppose  $n \in S$ . If any of the numbers less than  $n$  were in  $T$ , then one of them would be least (since there are only finitely many such numbers and every finite set has a least element<sup>1</sup>). So none of the numbers  $0, 1, \dots, n$  is in  $T$ . If  $n+1 \in T$ , then  $n+1$  would be least. So  $n+1 \notin T$ , i.e.,  $n+1 \in S$ . □

---

1. recall any sorting algorithm

# Prime Factorization

Recall

## Theorem (Fundamental Theorem of Arithmetic)

Let  $n \in \mathbb{N} \setminus \{0\}$ , then there exist  $k \geq 0$  prime numbers  $p_1, p_2, \dots, p_k$  such that  $n = \prod_{i=1}^k p_i$ . (also unique up to order, more on this later.)

## Proof by WOP.

Let  $T$  be the set of natural numbers greater than 1 which cannot be written as the product of primes. By WOP,  $T$  has a least element  $n$ . Clearly  $n$  cannot be prime, so  $n$  is composite. Then we can write  $n = ab$ , where neither of  $a$  and  $b$  is 1. So  $a < n$  and  $b < n$ . If both  $a$  and  $b$  had prime factorizations, then so would  $n$ . Therefore at least one of  $a$  and  $b$  does not have a prime factorization (by relabeling, we can assume it is  $a$ ). But  $a < n$  and  $a \in T$ , so  $n$  was not least in  $T$ .

This contradiction establishes that  $T$  has no least element, hence by WOP must be empty. So every natural number greater than 1 can be written as the product of primes. □

## Back to The Beginning

Recall the claim that  $\sum_{k=0}^n k = \frac{n(n+1)}{2}$  for all  $n \in \mathbb{N}$ .

Proof by WOP.

Suppose not. Then there exists some  $p \in \mathbb{N}$  such that  $\sum_{k=0}^p k \neq \frac{p(p+1)}{2}$ . Consider the set of all such numbers:

$$T = \left\{ k \mid 1 + \cdots + k \neq \frac{k(k+1)}{2} \right\}$$

Thus  $p \in T$  by assumption; in particular  $T \neq \emptyset$ .

By WOP,  $T$  has a least element  $N$ .  $N \neq 0$  since  $0 = 0(0+1)/2$ . Therefore  $N > 1$ . But then  $N$  admits a predecessor,  $n = N - 1$ . Since  $n < N$ , then  $n \notin T$  (b/c  $N$  is **least** in  $T$ ). That is,  $\sum_{k=0}^n k = \frac{n(n+1)}{2}$ . Now consider  $\sum_{k=0}^N k = (\sum_{k=0}^n k) + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2} = \frac{N(N+1)}{2}$ , which show that  $N \notin T$ , contradiction! So the initial supposition was incorrect, and thus the claim is true. □

# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
- 4. Relations and Functions**
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Motivation

## Primitive Models

- ▶ equivalence relations ( $=, \cong, \equiv, \sim$ , etc.)
- ▶ partial orders ( $\leq, \subset, \preceq, |$ , etc.)

# Relations

## Definition

A subset  $R \subset A \times B$  is called a (binary) **relation** from  $A$  to  $B$ . If  $A = B$ , we say that  $R$  is a **relation on  $A$** .

## Notation

- ▶  $(x, y) \in R$ ,
- ▶  $xRy$ ,
- ▶ via predefined symbols, e.g.,
  - ▶  $x \preceq y$ , i.e.,  $(x, y) \in \preceq \subset A \times B$ ;
  - ▶  $x \sim y$ , i.e.,  $(x, y) \in \sim \subset A \times B$ .

## Definition

- ▶  $\text{domain}(R) := \{x \mid \exists y(xRy)\}$
- ▶  $\text{range}(R) := \{y \mid \exists x(xRy)\}$

# Relations

## Examples

Suppose  $R \subset A \times B$ .

- ▶  $R = \emptyset$ , the ***empty relation***, with domain  $(\emptyset) = \text{range}(\emptyset) = \emptyset$ .
- ▶ When  $A = B$ , we have the ***identity relation***,

$$\text{id}_A = \{(a, a) \mid a \in A\}$$

The identity relation relates every element to itself. Note that  $\text{domain}(\text{id}_A) = \text{range}(\text{id}_A) = A$ .

- ▶ The relation  $A \times B$  itself. This relation relates every element of  $A$  to every element of  $B$ . Note that  $\text{domain}(A \times B) = A$  and  $\text{range}(A \times B) = B$ .

# Functions

## Definition

A **function** is a relation  $F$  such that

$$\forall x \in \text{dom } F (\exists !y (xFy))$$

## Remark

For a function  $F$  and a point  $x \in \text{dom}(F)$ , the unique  $y$  such that  $xFy$  is called the **value** of  $F$  at  $x$  and is denoted  $F(x)$ . Thus  $(x, F(x)) \in F$ .

## Remark

Given function  $F : A \rightarrow B$ , then  $\forall x, y \in A (x = y \Rightarrow F(x) = F(y))$ .

## Remark

- ▶ Partial function: not (necessarily) everywhere defined.
- ▶ (Total) function: everywhere defined.

## Collatz Conjecture (3n + 1 Problem)

Given  $n \in \mathbb{N} \setminus \{0\}$ , construct the sequence  $c_i(n)$  as follows starting with  $i = 1$ :

$$c_1(n) = n$$
$$c_{i+1}(n) = \begin{cases} c_i(n)/2, & \text{if } c_i(n) \text{ even} \\ 3c_i(n) + 1, & \text{if } c_i(n) \text{ odd} \end{cases}$$

Observe that for  $n = 1$ , we get the infinite periodic sequence

$$1 \rightarrow 4 \rightarrow 2 \rightarrow 1 \rightarrow 4 \rightarrow 2 \rightarrow 1 \rightarrow \dots$$

so we may assume that we stop the first time that the sequence  $c_i(n)$  reaches the value 1 (if at all). Such an index  $i$  is called the **stopping time** of the sequence.

### Conjecture (Collatz)

For any starting integer value  $n \geq 1$ , the sequence  $(c_i(n))$  always reaches 1.

## Collatz Conjecture ( $3n + 1$ Problem)

- ▶  $n = 3$ : stops after 7 steps,

$$3 \rightarrow 10 \rightarrow 5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1.$$

- ▶  $n = 5$ : stops after 5 steps,

$$5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1.$$

- ▶  $n = 6$ : stops after 8 steps,

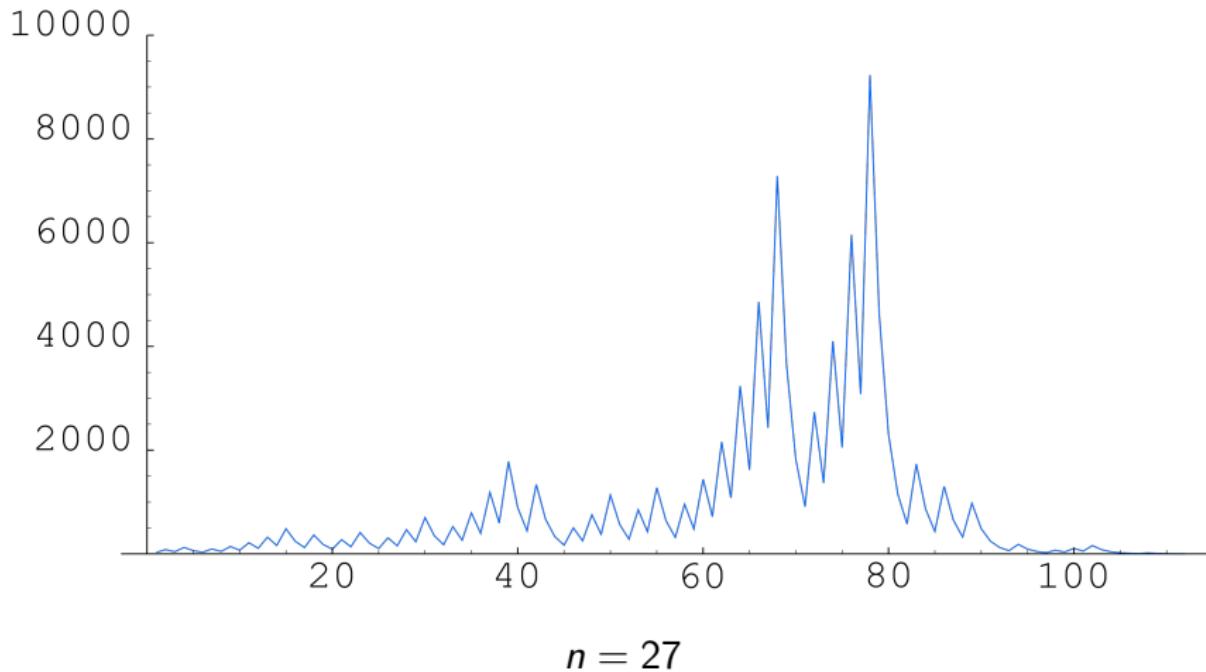
$$6 \rightarrow 3 \rightarrow 10 \rightarrow 5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1.$$

- ▶  $n = 7$ : stops after 16 steps,

$$\begin{aligned} 7 &\rightarrow 22 \rightarrow 11 \rightarrow 34 \rightarrow 17 \rightarrow 52 \rightarrow 26 \\ &\rightarrow 13 \rightarrow 40 \rightarrow 20 \rightarrow 10 \rightarrow 5 \rightarrow 16 \\ &\rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1. \end{aligned}$$

- ▶  $n = 27$ : stops after 111 steps.
- ▶  $n = 97$ : stops after 118 steps.

## Collatz Conjecture ( $3n + 1$ Problem)



$$n = 27$$

cf., Veritasium, The Simplest Math Problem No One Can Solve (Featuring Alex Kontorovich), e.g., <https://www.bilibili.com/video/BV1Ty4y1778o>

## Collatz Conjecture ( $3n + 1$ Problem)

We can define the partial function  $C : \mathbb{N} \setminus \{0\} \rightarrow \mathbb{N}$  as

$$C(n) := \min\{i \in \mathbb{N} \mid c_i(n) = 1\}$$

For example,

|        |   |   |   |   |   |   |    |   |    |    |    |         |
|--------|---|---|---|---|---|---|----|---|----|----|----|---------|
| $n$    | 1 | 2 | 3 | 4 | 5 | 6 | 7  | 8 | 9  | 10 | 11 | $\dots$ |
| $C(n)$ | 0 | 1 | 7 | 2 | 5 | 8 | 16 | 3 | 19 | 6  | 14 | $\dots$ |

The Collatz conjecture is equivalent to asserting that the function  $C$  is a (total) function.

## Functions

For arbitrary sets/relations/functions  $A$ ,  $F$ , and  $G$ ,

- ▶ The **inverse** of  $F$  is the set

$$F^\top = F^{-1} = \{(y, x) \mid xFy\}$$

- ▶ The **composition** of  $F$  and  $G$  is the set

$$F \circ G = \{(x, z) \mid \exists y \in A(xFy \wedge yGz)\}$$

- ▶ The **restriction** of  $F$  to  $A$  is the set

$$F|A = \{(x, y) \in F \mid x \in A\}$$

- ▶ The **image** of  $A$  **under**  $F$  is the set

$$F(A) = \text{ran}(F|A) = \{y \mid (\exists x \in A)x F y\}$$

If  $F$  is a function, then  $F(A) = \{F(x) \mid x \in A\}$ .

# Injection and Surjection

## Definition

Given a function  $F : A \rightarrow B$ , with  $\text{dom } F = A$  and  $\text{ran}(F) \subset B$ , then

- ▶  $F$  is **injective** or **one-to-one** if  $\forall x, y \in A(F(x) = F(y) \Rightarrow x = y)$ ;
- ▶  $F$  is **surjective** or **onto** if  $\text{ran}(F) = B$ .
- ▶  $F$  is **bijection** if it is both injective and surjective.

The above function  $F$  is also called an injection, surjection, or bijection, respectively.

## Theorem

Given a function  $F : A \rightarrow B$ ,  $A \neq \emptyset$ , then

- ▶ There exists a function  $G : B \rightarrow A$  (a “left inverse”) such that  $G \circ F = \text{id}_A \Leftrightarrow F$  is **one-to-one**.
- ▶ There exists a function  $G : B \rightarrow A$  (a “right inverse”) such that  $F \circ G = \text{id}_B \Leftrightarrow F$  is **onto**.

# Injection and Surjection

## Theorem

Let  $f : A \rightarrow B$ ,  $g : B \rightarrow C$

- ▶ If  $g \circ f$  is injective, then  $f$  is injective.
- ▶ If  $g \circ f$  is surjective, then  $g$  is surjective.

## Proof.

- ▶ Given  $x, y \in A$ , then

$$\begin{aligned}f(x) = f(y) &\Rightarrow g(f(x)) = g(f(y)) \Rightarrow (g \circ f)(x) = (g \circ f)(y) \\&\Rightarrow x = y \text{ (b/c } g \circ f \text{ is injective)}\end{aligned}$$

Therefore  $f$  is injective.

- ▶ Since  $g \circ f$  is surjective, then  $\forall z \in C, \exists x \in A$  such that  $g \circ f(x) = g(f(x)) = z$ . Let  $y = f(x) \in B$  (which exists for all  $x \in A$  since  $f$  is a function), then  $\forall z \in C, \exists y \in B$  such that  $z = f(y)$ .  
Therefore  $g$  is surjective. □

# Injection and Surjection

## Theorem

Let  $f : A \rightarrow B$ ,  $g : B \rightarrow C$

- ▶ If  $g \circ f$  is injective, then  $f$  is injective.
- ▶ If  $g \circ f$  is surjective, then  $g$  is surjective.

## Proof (Alternative).

- ▶ Since  $g \circ f$  is injective, then there exists  $h : C \rightarrow A$  such that  $h \circ (g \circ f) = id_A$ , that is, there exists  $h \circ g : C \rightarrow A$  such that  $(h \circ g) \circ f = id_A$ . Therefore  $f$  is left invertible, hence injective.
- ▶ Since  $g \circ f$  is surjective, then there exists  $h : C \rightarrow A$  such that  $(g \circ f) \circ h = id_C$ , that is, there exists  $h \circ g : C \rightarrow A$  such that  $g \circ (f \circ h) = id_C$ . Therefore  $g$  is right invertible, hence surjective.

□

# Relations as Functions

Relation as multivariable boolean functions

## Example

- Given a relation  $R \subset A \times B$ , the associated boolean function is given by

$$\phi_R : A \times B \rightarrow \{\top, \perp\}$$

$$(x, y) \mapsto \begin{cases} \top, & xRy \\ \perp, & \text{otherwise} \end{cases}$$

- Given a boolean function  $\phi : A \times B \rightarrow \{\top, \perp\}$ , the associated relation is given by

$$R_\phi = \{(x, y) \in A \times B \mid \phi(x, y) = \top\}$$

## Relations as Functions

Relation as set functions

### Example

- Given a relation  $R \subset A \times B$ , the associated set function is given by

$$\alpha_R : A \rightarrow \mathcal{P}(B)$$

$$x \mapsto \{y \in B \mid xRy\}$$

- Given a set function  $\alpha : A \rightarrow \mathcal{P}(B)$ , the associated relation is given by

$$R_\alpha = \{(x, y) \in A \times B \mid y \in \alpha(x)\}$$

## Relations as Functions

Relation as set functions

### Example

- Given a relation  $R \subset A \times B$ , the associated set function is given by

$$\beta_R : B \rightarrow \mathcal{P}(A)$$

$$x \mapsto \{y \in A \mid xRy\}$$

- Given a set function  $\beta : B \rightarrow \mathcal{P}(A)$ , the associated relation is given by

$$R_\beta = \{(x, y) \in A \times B \mid y \in \beta(x)\}$$

# Properties of Relations

## Definition

A (binary) relation  $R$  on  $A$ , i.e.,  $R \subset A \times A$ , is

- ▶ **reflexive** if  $aRa \Rightarrow \top$ .
- ▶ **irreflexive** if  $aRa \Rightarrow \perp$ .
- ▶ **total** if  $aRb \vee bRa \Rightarrow \top$ .
- ▶ **transitive** if  $aRb \wedge bRc \Rightarrow aRc$ .
- ▶ **symmetric** if  $aRb \Leftrightarrow bRa$ .
- ▶ **anti-symmetric** if  $aRb \wedge bRa \Rightarrow a = b$ .
- ▶ **asymmetric** if  $aRb \wedge bRa \Rightarrow \perp$ .

# Properties of Relations

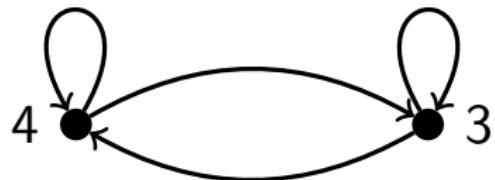
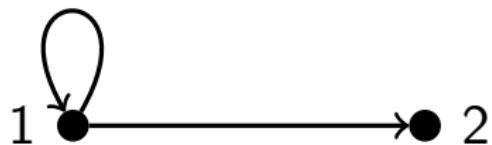
## Example

|                                     | <i>reflexive</i> | <i>transitive</i> | <i>symmetric</i> | <i>antisymmetric</i> |
|-------------------------------------|------------------|-------------------|------------------|----------------------|
| $\leq$ on $\mathbb{R}$              |                  |                   |                  |                      |
| $<$ on $\mathbb{R}$                 |                  |                   |                  |                      |
| $\subset$ on $2^S$                  |                  |                   |                  |                      |
| $\subsetneq$ on $2^S$               |                  |                   |                  |                      |
| $\equiv_n$ on $\mathbb{Z}$          |                  |                   |                  |                      |
| $ $ on $\mathbb{N} \setminus \{0\}$ |                  |                   |                  |                      |
| $ $ on $\mathbb{N}$                 |                  |                   |                  |                      |
| $ $ on $\mathbb{Z}$ ( $0 \mid 0$ )  |                  |                   |                  |                      |
| matrix similarity                   |                  |                   |                  |                      |

## Properties of Relations

### Example

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



# Important Relations

## Definition

A **partial order** on a set  $P$  is a relation that is

- ▶ reflexive
- ▶ antisymmetric
- ▶ transitive

## Example

- ▶ On  $\mathbb{Z}$  (or  $\mathbb{R}$  etc.):  $a \leq b$
- ▶ On  $2^S$  for a given  $S$ :  $A \subset B$
- ▶ On  $\mathbb{N}$ :  $a \mid b$
- ▶ On partitions: refinement

## Irreflexive Relations

Recall that a relation  $R$  on a set  $A$  is **irreflexive** if  $aRa \Rightarrow \perp$  for all  $a \in A$ .

### Remark

- ▶ Irreflexive and non-reflexive are different concepts.
  - ▶ Irreflexive: zero self-loops.
  - ▶ Non-reflexive: missing self-loops.
- ▶ Antisymmetric and non-symmetric are different concepts.
  - ▶ Antisymmetric: no cycle of length 2.
  - ▶ Non-symmetric: exists directed edge of no return.

## Terminologies on Partial Orders

### Non-strict Partial Order (e.g., $\leq$ , $\subseteq$ )

A **reflexive**, **weak**, or **non-strict** partial order is a relation  $\preceq$  over a set  $A$  that is

- ▶ reflexive
- ▶ antisymmetric
- ▶ transitive

### Strict Partial Order (e.g., $<$ , $\subsetneq$ )

An **irreflexive**, **strong**, or **strict** partial order is a relation  $\prec$  over a set  $A$  that is

- ▶ irreflexive
- ▶ asymmetric
- ▶ transitive

### Remark

For every strict partial order there is a unique corresponding non-strict partial order, and vice-versa. The two differ by the identity relation on  $A$ .

### Remark

The term **partial order** typically refers to a non-strict partial order relation.

# Important Relations

## Definition

An **equivalence relation** on a set  $A$  is a relation that is

- ▶ reflexive
- ▶ symmetric
- ▶ transitive

## Example

- ▶ On  $\mathbb{Z}$  (or  $\mathbb{R}$  etc.):  $a = b$
- ▶ On  $\mathbb{Z}$ :  $a \equiv b \pmod{12}$
- ▶ On  $2^S$  for given  $S$ :  $A \equiv B$  iff  $|A| = |B|$
- ▶ On square matrices:  $A \cong B$  iff  $A = PBP^{-1}$

# Equivalence Classes

## Definition

Given an equivalence relation  $R$  on  $A$ , the **equivalence class** containing  $x$  is the set

$$[x]_R := \{t \in A \mid xRt\}$$

## Theorem

Given an equivalence relation  $R$  on  $A$ , then for  $x, y \in A$ ,

$$[x]_R = [y]_R \Leftrightarrow xRy$$

## Proof.

( $\Rightarrow$ ) Let  $[x]_R = [y]_R$ , then  $y \in [y]_R$  (b/c  $yRy$ ) implies  $y \in [x]_R$  (b/c  $[x]_R = [y]_R$ ), hence  $xRy$  (by definition of  $[x]_R$ ).

## Equivalence Classes

Proof.

( $\Leftarrow$ ) Assume  $xRy$ , then take any  $t \in [y]_R$ ,

$$t \in [y]_R \Rightarrow yRt$$

$\Rightarrow xRt$  ( $xRy$  and transitivity)

$$\Rightarrow t \in [x]_R$$

hence  $[y]_R \subset [x]_R$ . The other direction is by symmetry of  $R$ .



# Partition

## Definition

A **partition**  $\Pi$  of a set  $A$  is a set of nonempty subsets of  $A$  that is disjoint and exhaustive, i.e.,

- ▶  $\forall a, b \in \Pi (a \neq b \Rightarrow a \cap b = \emptyset);$
- ▶  $\bigcup \Pi = A.$

An element of a partition is called a **fiber**, a **block**, or an **equivalence class**. An element of such an equivalence class is called a **representative** of the class.

## Examples

- ▶  $A = \emptyset: \Pi = \emptyset.$
- ▶  $A \neq \emptyset: \Pi = \{\{x\} \mid x \in A\},$  or  $\Pi = \{A\}.$
- ▶  $A = \mathbb{N}: \Pi = \{\{\text{even numbers}\}, \{\text{odd numbers}\}\}$

# Partition

## Theorem

Given an equivalence relation  $R$  on  $A$ , then the set  $\{[x]_R \mid x \in A\}$  of all equivalence classes is a partition of  $A$ .

## Proof.

- ▶  $[x]_R \neq \emptyset \forall x. \text{ (b/c } x \in [x]_R)$
- ▶  $\bigcup\{[x]_R \mid x \in A\} = A. \text{ (b/c } x \in [x]_R \subset A)$
- ▶ Suppose  $t \in [x]_R \cap [y]_R$ , then  $tRx$  and  $tRy$ , hence  $xRy$  and  $[x]_R = [y]_R$ .

□

## Quotient set

### Definition

Given an equivalence relation  $R$  on  $A$ , then the **quotient set** is given by

$$A/R := \{[x]_R \mid x \in A\}$$

where  $A/R$  is read “ $A$  modulo  $R$ ”.

### Example

Let  $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ , and define the relation  $\sim$  on  $\mathbb{N}$  by

$$m \sim n \iff m - n \text{ divisible by } 6$$

then  $\mathbb{N}/\sim = \{[0], [1], [2], [3], [4], [5]\}$

## Quotient set

### Example

Let  $F : A \rightarrow B$  and for elements in  $A$  define  $x \sim y \Leftrightarrow F(x) = F(y)$ . The relation  $\sim$  is an equivalence relation on  $A$ . There is a **unique one-to-one** function  $\tilde{F} : A/\sim \rightarrow B$  such that  $F = \tilde{F} \circ q$ , where  $q : A \rightarrow A/\sim$ ,  $x \mapsto [x]$ , is the natural quotient map. The value of  $\tilde{F}$  at a particular equivalence class is the common value at the member of that equivalence class.

$$\begin{array}{ccc} A/\sim & & \\ \uparrow q & \searrow \tilde{F} & \\ A & \xrightarrow{F} & B \end{array}$$

- ▶  $\tilde{F}$  is well-defined. Pick  $[x], [y] \in A/\sim$ ,

$$[x] = [y] \Rightarrow x \sim y \Rightarrow F(x) = F(y) \Rightarrow F([x]) = F([y])$$

- ▶  $\tilde{F}$  one-to-one.  $\tilde{F}([x]) = \tilde{F}([y]) \Leftrightarrow F(x) = F(y) \Leftrightarrow x \sim y \Leftrightarrow [x] = [y]$ .
- ▶  $q$  is surjective. ( $b/c x \in [x] \forall x \in A$ )
- ▶  $\tilde{F}$  is unique.  $\tilde{F} \circ q = \tilde{G} \circ q \Rightarrow \tilde{F} = \tilde{G}$ .

Note that  $\tilde{F} : A/\sim \rightarrow \text{ran}(F)$ ,  $[x] \mapsto F(x)$ , is bijective.

# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Integers

## Definition

Let  $\sim$  be the following equivalence relation on  $\mathbb{N}^2$ ,

$$(a, b) \sim (c, d) \Leftrightarrow a +_{\mathbb{N}} d = b +_{\mathbb{N}} c$$

Explicitly, note that  $\sim \subset \mathbb{N}^2 \times \mathbb{N}^2$ , and

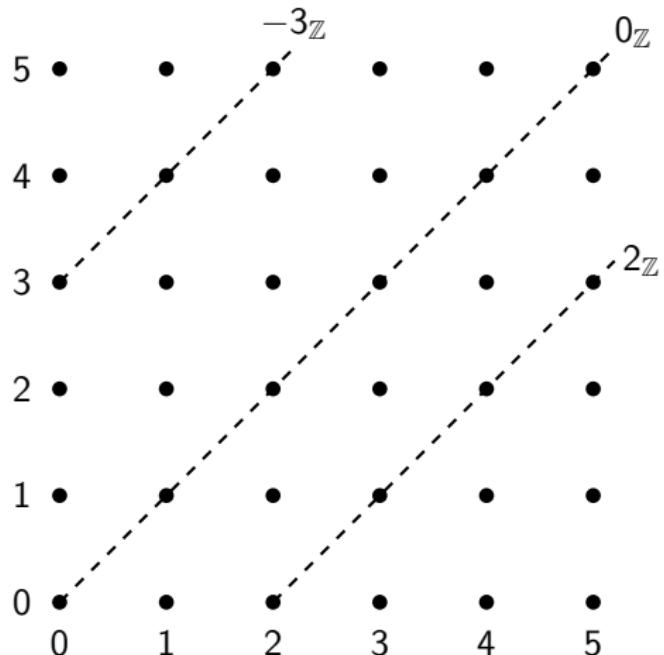
$$\sim = \{((a, b), (c, d)) \in \mathbb{N}^2 \times \mathbb{N}^2 \mid a, b, c, d \in \mathbb{N}, a +_{\mathbb{N}} d = b +_{\mathbb{N}} c\}$$

Define  $\mathbb{Z} = \mathbb{N} \times \mathbb{N}/\sim$ .

## Example

- ▶  $0_{\mathbb{Z}} = \{(0_{\mathbb{N}}, 0_{\mathbb{N}}), (1_{\mathbb{N}}, 1_{\mathbb{N}}), \dots\} = [(0_{\mathbb{N}}, 0_{\mathbb{N}})] = [\{\{\emptyset\}\}] = [\{\{\{\}\}\}]$ .
- ▶  $2_{\mathbb{Z}} = \{(2_{\mathbb{N}}, 0_{\mathbb{N}}), (3_{\mathbb{N}}, 1_{\mathbb{N}}), (4_{\mathbb{N}}, 2_{\mathbb{N}}), \dots\} = [(2_{\mathbb{N}}, 0_{\mathbb{N}})]$ .
- ▶  $-3_{\mathbb{Z}} = \{(0_{\mathbb{N}}, 3_{\mathbb{N}}), (1_{\mathbb{N}}, 4_{\mathbb{N}}), (2_{\mathbb{N}}, 5_{\mathbb{N}}), \dots\} = [(0_{\mathbb{N}}, 3_{\mathbb{N}})]$ .

# Integers



# Integers

## Remark

The relation  $\sim$  is an equivalence relation on  $\mathbb{N} \times \mathbb{N}$ .

## Proof.

Let  $(a, b), (a', b'), (a'', b'') \in \mathbb{N} \times \mathbb{N}$ .

- ▶ Reflexivity. It is clear that  $(a, b) \sim (a, b)$ , since  $a + b = b + a$ .
- ▶ Symmetry. Let  $(a, b) \sim (a', b')$ , then

$$\begin{aligned}(a, b) \sim (a', b') &\Rightarrow a + b' = a' + b \\ &\Rightarrow a' + b = a + b' \Rightarrow (a', b') \sim (a, b)\end{aligned}$$

- ▶ Transitivity. Let  $(a, b) \sim (a', b')$  and  $(a', b') \sim (a'', b'')$ , then we have  $a + b' = a' + b$  and  $a' + b'' = a'' + b'$ . Therefore

$$a + b' + a' + b'' = a' + b + a'' + b'$$

hence  $a + b'' = a'' + b$ , i.e.,  $(a, b) \sim (a'', b'')$ . □

# Integers

## Well-defined operations on equivalence classes

- ▶  $[(a, b)] +_{\mathbb{Z}} [(c, d)] := [(a + c, b + d)].$
- ▶  $[(a, b)] \times_{\mathbb{Z}} [(c, d)] := [(ac + bd, ad + bc)].$
- ▶  $[(a, b)] <_{\mathbb{Z}} [(c, d)]$  if  $a + d <_{\mathbb{N}} b + c.$

### Proof.

We want to show that the operations do not depend on the choice of representatives. choose  $(a, b) \sim (a', b')$  and  $(c, d) \sim (c', d')$ , which indicates that  $a + b' = a' + b$  and  $c + d' = c' + d$ , then

- ▶ We want to show  $a + c + b' + d' = a' + b' + c + d$ . Since
$$(a + c) + (b' + d') = (a + b') + (c + d') \\ = (a' + b) + (c' + d) = (a' + c') + (b + d)$$

Hence  $(a + c, b + d) \sim (a' + c', b' + d')$ , and

$[(a + c, b + d)] = [(a' + c', b' + d')]$ , independent of the choice of representatives.

# Integers

## Proof (Cont.)

► We want to show  $ac + bd + a'd' + b'c' = a'c' + b'd' + ad + bc$ .

Note that  $a + b' = a' + b$  and  $c + d' = c' + d$ , then

$$c(a + b') = c(a' + b)$$

$$d(a' + b) = d(a + b')$$

$$a'(c + d') = a'(c' + d)$$

$$b'(c' + d) = b'(c + d')$$

which simplifies to

$$ac + b'c = a'c + bc$$

$$a'd + bd = ad + b'd$$

$$a'c + a'd' = a'c' + a'd$$

$$b'c' + b'd = b'c + b'd'$$

Add all above together and cancel the unwanted terms.

# Integers

## Proof (Cont.)

- We want to show that  $a + d < b + c$  iff  $a' + d' < b' + c'$ . Recall that  $a + b' = a' + b$  and  $c + d' = c' + d$ , then

$$\begin{aligned} a + d < b + c &\Leftrightarrow a + d + b' + d' < b + c + b' + d' \\ &\Leftrightarrow a' + b + d + d' < b + b' + c' + d \\ &\Leftrightarrow a' + d' < b' + c' \end{aligned}$$

Therefore the ordering  $<_{\mathbb{Z}}$  on  $\mathbb{Z}$  is well-defined.

□

# Rational Numbers

## Definition

Let  $\mathbb{Z}^+ = \{z \in \mathbb{Z} \mid z >_{\mathbb{Z}} 0_{\mathbb{Z}}\}$ , and let  $\sim$  be the following equivalence relation on  $\mathbb{Z} \times \mathbb{Z}^+$ ,

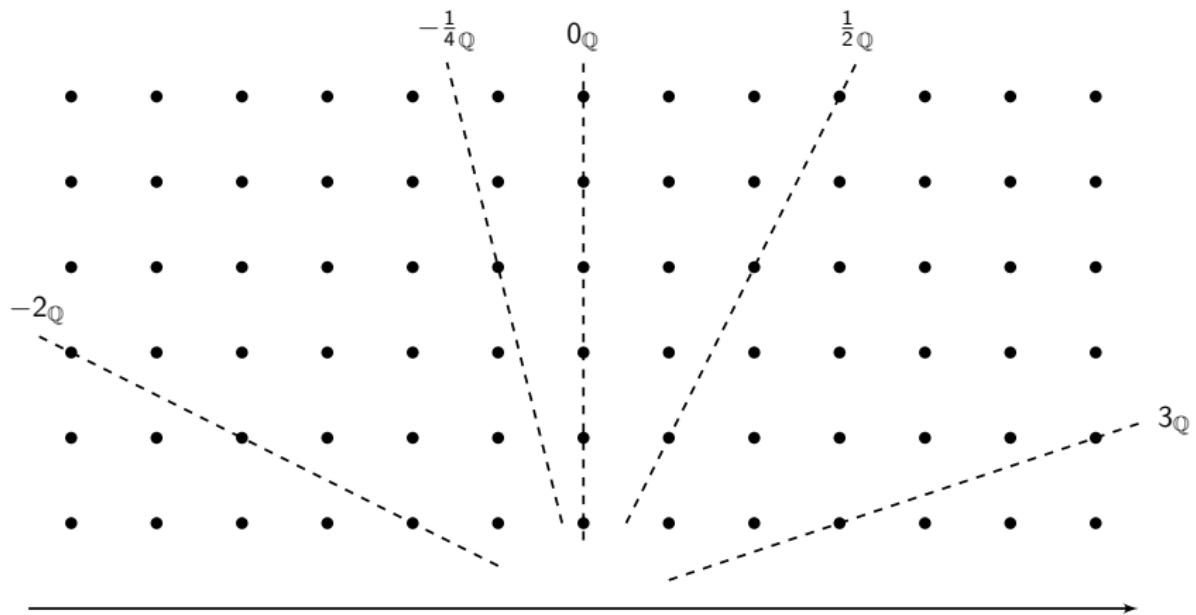
$$(a, b) \sim (c, d) \Leftrightarrow a \times_{\mathbb{Z}} d = b \times_{\mathbb{Z}} c$$

Explicitly, note that  $\sim \subset (\mathbb{Z} \times \mathbb{Z}^+) \times (\mathbb{Z} \times \mathbb{Z}^+)$ , and

$$\sim = \left\{ ((a, b), (c, d)) \in (\mathbb{Z} \times \mathbb{Z}^+)^2 \mid \begin{array}{l} a, c \in \mathbb{Z}, b, d \in \mathbb{Z}^+, \\ a \times_{\mathbb{Z}} d = b \times_{\mathbb{Z}} c \end{array} \right\}$$

Define  $\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}^+ / \sim$ .

# Rational Numbers



# Rational Numbers

## Remark

The relation  $\sim$  is an equivalence relation on  $\mathbb{Z} \times \mathbb{Z}^+$ .

## Proof.

Let  $(a, b), (a', b'), (a'', b'') \in \mathbb{Z} \times \mathbb{Z}^+$ .

- ▶ Reflexivity. Obviously  $(a, b) \sim (a, b)$  since  $ab = ba$ .
- ▶ Symmetry. Let  $(a, b) \sim (a', b')$ , then

$$(a, b) \sim (a', b') \Rightarrow ab' = a'b \Rightarrow a'b = ab' \Rightarrow (a', b') \sim (a, b)$$

- ▶ Transitivity. Let  $(a, b) \sim (a', b')$  and  $(a', b') \sim (a'', b'')$ , then we have  $ab' = a'b$  and  $a'b'' = a''b'$ . Thus  $ab'a'b'' = a'ba''b'$ . Since  $b' \neq 0$ , then  $aa'b'' = a'ba''$ .

- ▶ If  $a' \neq 0$ , then  $ab'' = a''b$ , i.e.,  $(a, b) \sim (a'', b'')$ .
- ▶ If  $a' = 0$ , then  $ab' = 0$ , hence  $a = 0$  ( $b/c b' \neq 0$ ). Similarly  $a'' = 0$ . Therefore  $ab'' = 0 = a''b$ , i.e.,  $(a, b) \sim (a'', b'')$ . □

# Rational Numbers

## Well-defined operations on equivalence classes

- ▶  $[(a, b)] +_{\mathbb{Q}} [(c, d)] := [(a \times_{\mathbb{Z}} d +_{\mathbb{Z}} b \times_{\mathbb{Z}} c, b \times_{\mathbb{Z}} d)].$
- ▶  $[(a, b)] \times_{\mathbb{Q}} [(c, d)] := [(a \times_{\mathbb{Z}} c, b \times_{\mathbb{Z}} d)].$
- ▶  $[(a, b)] <_{\mathbb{Q}} [(c, d)]$  if  $a \times_{\mathbb{Z}} d <_{\mathbb{Z}} b \times_{\mathbb{Z}} c.$

## Proof.

Choose  $(a, b) \sim (a', b')$  and  $(c, d) \sim (c', d')$ , which indicates that  $ab' = a'b$  and  $cd' = c'd$ , with  $b, b', d, d' > 0$ .

- ▶ We want to show that  $(ad + bc, bd) \sim (a'd' + b'c', b'd')$ , that is,  $(ad + bc)b'd' = bd(a'd' + b'c')$ , or  $\textcolor{red}{ab'}dd' + bb'\textcolor{blue}{cd'} = \textcolor{red}{a'b}dd' + bb'\textcolor{blue}{c'd'}$ . Note that this is guaranteed by  $ab' = a'b$  and  $cd' = c'd$ .
- ▶ Since  $ab' = a'b$  and  $cd' = c'd$ , then  $ab'cd' = a'bcd'$ , thus  $(ac, bd) \sim (a'c', b'd')$ .
- ▶  $ad < bc \Leftrightarrow adb'd' < bcb'd' \Leftrightarrow a'bdd' < bb'c'd \Leftrightarrow a'd' < b'c'$  (note that  $b', d' > 0$ ). Therefore the ordering  $<_{\mathbb{Z}}$  on  $\mathbb{Q}$  is well-defined. □

# Real Numbers

## Definition

A **Cauchy sequence** is a function  $s : \mathbb{N} \rightarrow \mathbb{Q}$  such that  $|s_m - s_n|$  is arbitrarily small for all sufficiently large  $m$  and  $n$ ; i.e.,

$$(\forall \varepsilon \in \mathbb{Q}, \varepsilon > 0)(\exists k \in \mathbb{N})(\forall m, n > k)|s_m - s_n| < \varepsilon$$

## Definition

Let  $C$  be the set of all Cauchy sequences. For  $r, s \in C$ , then  $r$  and  $s$  are **equivalent** ( $r \sim s$ ) if  $|s_m - s_n|$  is arbitrarily small for all sufficiently large  $n$ ; i.e.,

$$(\forall \varepsilon \in \mathbb{Q}, \varepsilon > 0)(\exists k \in \mathbb{N})(\forall n > k)|r_n - s_n| < \varepsilon$$

Define  $\mathbb{R}$  to be the quotient set  $C/\sim$ . (This approach to constructing  $\mathbb{R}$  is due to Cantor.)

# Real Numbers

## Definition

A **Dedekind cut** is a set  $x \subset \mathbb{Q}$  such that

- ▶  $x \neq \emptyset$  and  $x \neq \mathbb{Q}$ ;
- ▶  $x$  is closed downward, i.e.,  $\forall p, q \in \mathbb{Q}(p < q \Rightarrow (q \in x \Rightarrow p \in x))$ ;
- ▶  $x$  has no largest element.

Define  $\mathbb{R}$  to be the set of all Dedekind cuts.

## Definition

Let  $x \leq_{\mathbb{R}} y$  if  $x \subset y$ .

## Theorem

*Every subset  $A \subset \mathbb{R}$  that is bounded above admits a least upper bound.*

## Proof.

Take  $x = \bigcup A$ . Claim:  $x$  is a Dedekind cut, and  $\forall y \in A(x \geq y)$ . □



# Equinumerosity

## Definition

A set  $A$  is **equinumerous** to a set  $B$  (written  $A \approx B$ ) if there is a bijection from  $A$  to  $B$ .

## Theorem

For any sets  $A$ ,  $B$ , and  $C$ :

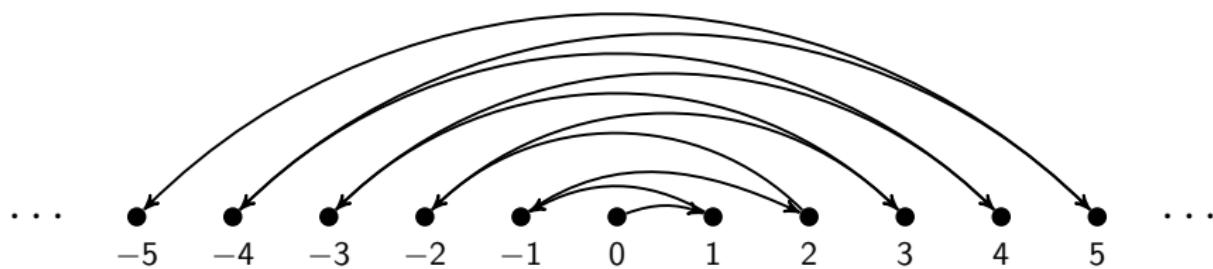
- ▶  $A \approx A$ .
- ▶  $A \approx B \Rightarrow B \approx A$ .
- ▶  $(A \approx B \wedge B \approx C) \Rightarrow A \approx C$ .

## Warning

NOT an equivalence relation since it concerns **all** sets.

# Equinumerosity

- $\mathbb{Z} \approx \mathbb{N}$ .



|        |         |      |      |      |      |      |     |     |     |     |     |     |         |
|--------|---------|------|------|------|------|------|-----|-----|-----|-----|-----|-----|---------|
| $z$    | $\dots$ | $-5$ | $-4$ | $-3$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ | $4$ | $5$ | $\dots$ |
| $f(z)$ | $\dots$ | 10   | 8    | 6    | 4    | 2    | 0   | 1   | 3   | 5   | 7   | 9   | $\dots$ |

$$f : \mathbb{Z} \rightarrow \mathbb{N}$$

$$z \mapsto \begin{cases} 2z + 1, & z \geq 0 \\ -2z, & z < 0 \end{cases}$$

# Equinumerosity

►  $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$ .

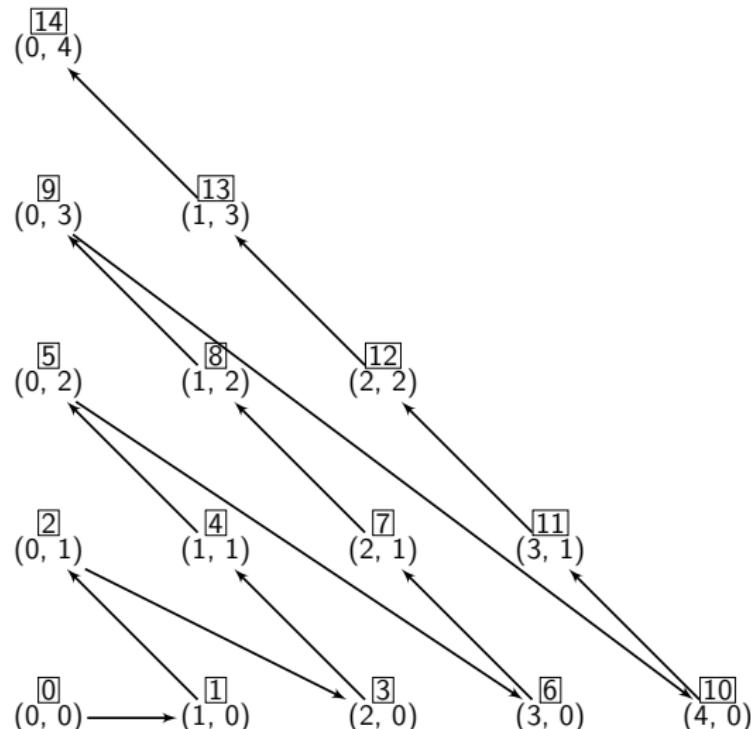
## Cantor's Paring Function

$J : \mathbb{N}^2 \rightarrow \mathbb{N}$ ,

$$J(x, y) = \binom{x + y + 1}{2} + y$$

## Theorem (Fueter-Pólya)

*The only quadratic paring functions are the Cantor polynomials (up to interchanging  $x$  and  $y$ ).*



## Remark

It is unknown whether this the **only polynomial** pairing function.

## Pairing Functions Recursively Defined

Consider Cantor's pairing function

$$J(x, y) = \binom{x+y+1}{2} + y, \quad x, y \in \mathbb{N}$$

We can define it recursively as

$$J(x+1, y) = J(x, y) + x + y + 1$$

$$J(0, y) = \binom{y+1}{2} + y$$

The function  $j : \mathbb{N} \rightarrow \mathbb{N}^2$  is Bijective ( $J^{-1} = j$ ).

A pair  $(x, y) \in \mathbb{N}^2$  is either

- ▶  $(0, 0)$ , or
- ▶  $(y+1, 0)$  as successor of  $(0, y)$ , or
- ▶  $(x-1, y+1)$  as the successor of  $(x, y)$  when  $x \neq 0$ .

## Pairing Functions Recursively Defined

Consider another pairing function

$$P(x, y) = 2^x(2y + 1) - 1, \quad x, y \in \mathbb{N}$$

We can define it recursively as

$$P(x + 1, y) = 2P(x, y) + 1$$

$$P(0, y) = 2y$$

$P$  is Bijective

Recall fundamental theorem of arithmetic.

- ▶ Surjectivity. For all  $z + 1 \in \mathbb{N}$ ,  $z + 1 = 2^x(2y + 1)$  for some  $x, y \in \mathbb{N}$ .
- ▶ Injectivity. Follows from uniqueness of factorization.

## Recursively Defined Functions

- When  $m \geq 1$ , given any two functions  $g : \mathbb{N}^m \rightarrow \mathbb{N}$  and  $h : \mathbb{N} \times \mathbb{N}^m \times \mathbb{N} \rightarrow \mathbb{N}$ , there exists a unique function  $f : \mathbb{N} \times \mathbb{N}^m \rightarrow \mathbb{N}$  defined by

$$\begin{aligned}f(n+1, x) &= h(n, x, f(n, x)) \\f(0, x) &= g(x)\end{aligned}$$

where  $x = (x_1, \dots, x_m) \in \mathbb{N}^m$ .

- When  $m = 0$ , we have

$$\begin{aligned}f(n+1) &= h(n, f(n)) \\f(0) &= f_0\end{aligned}$$

for all  $n \in \mathbb{N}$ , and some fixed  $f_0 \in \mathbb{N}$ .

## Comparing with Differential Equations

- When  $m \geq 1$ , given any two **suitable** functions  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  and  $h : \mathbb{R}^+ \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ , there exists a unique function  $u : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathbb{R}$  such that

$$\begin{aligned}\frac{\partial u}{\partial t} &= h(t, x, u(t, x)) \\ u(0, x) &= g(x)\end{aligned}$$

where  $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ .

- When  $m = 0$ , we have

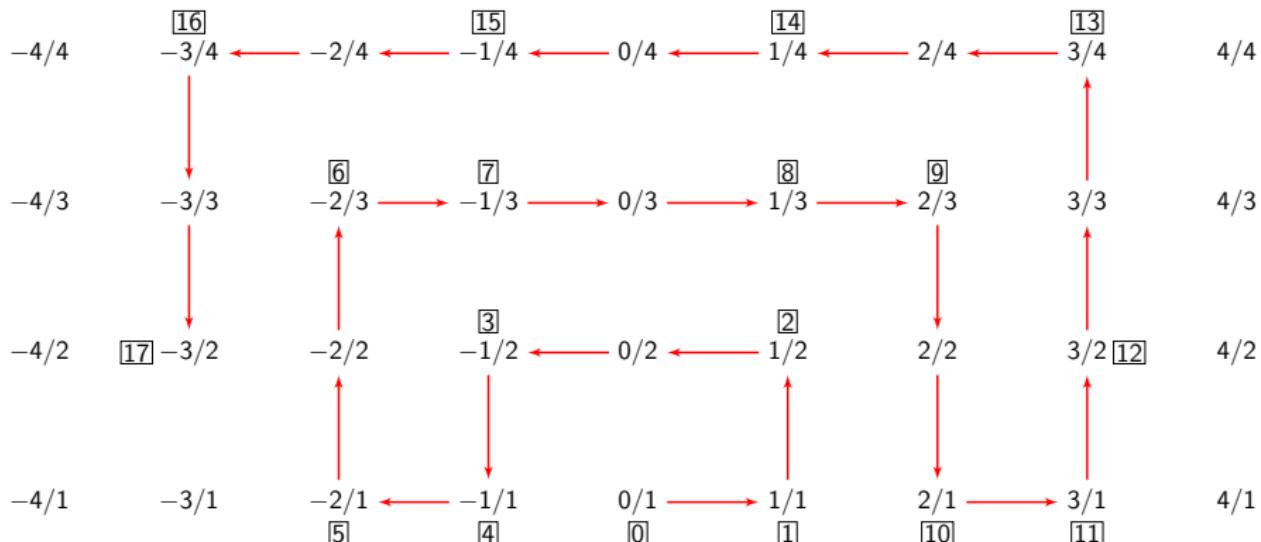
$$\begin{aligned}\frac{du}{dt} &= h(t, u(t)) \\ u(0) &= u_0\end{aligned}$$

for some fixed  $u_0 \in \mathbb{R}$ .

# Equinumerosity

►  $\mathbb{Q} \approx \mathbb{N}$ .

$-4/5$        $-3/5$        $-2/5$        $-1/5$        $0/5$        $1/5$        $2/5$        $3/5$        $4/5$



$$\mathbb{Q} \approx \mathbb{N}$$

Define  $g : \mathbb{N} \rightarrow \mathbb{Q}$ , such that

$$g(0) = [f(0)]$$

$g(n+1) = [f(k)]$  where  $k$  is the first such that

$$\forall i \leq n, g(i) \not\sim f(k)$$

|        |   |   |   |    |    |                |                |    |   |               |               |               |               |               |         |
|--------|---|---|---|----|----|----------------|----------------|----|---|---------------|---------------|---------------|---------------|---------------|---------|
| $n$    | 0 | 1 | 2 | 3  | 4  | 5              | 6              | 7  | 8 | 9             | 10            | 11            | 12            | 13            | $\dots$ |
| $g(n)$ | 0 | 1 | 2 | -2 | -1 | $-\frac{1}{2}$ | $-\frac{3}{2}$ | -3 | 3 | $\frac{3}{2}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{2}{3}$ | $\frac{4}{3}$ | $\dots$ |

# Calkin-Wilf-Newman

- $\mathbb{Q}^+ \approx \mathbb{N}$

## Moshe Newman successor function

The function

$$x \mapsto \frac{1}{[x] + 1 - \{x\}}$$

generates the Calkin-Wilf sequence

$$\frac{1}{1} \rightarrow \frac{1}{2} \rightarrow \frac{2}{1} \rightarrow \frac{1}{3} \rightarrow \frac{3}{2} \rightarrow \frac{2}{3} \rightarrow \frac{3}{1} \rightarrow \frac{1}{4} \rightarrow \frac{4}{3} \rightarrow \dots$$

which contains every **positive** rational number exactly once.

(Aigner & Ziegler, p. 131)

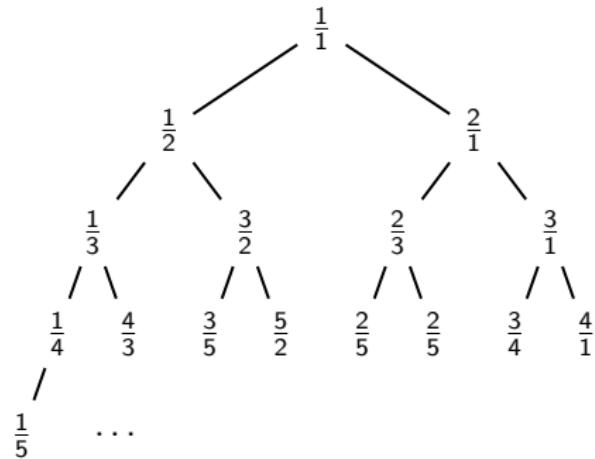
# Calkin-Wilf-Newman

Consider the infinite binary tree,

- ▶  $\frac{1}{1}$  is the root, and
- ▶ every node  $\frac{i}{j}$  has a left child  $\frac{i}{i+j}$  and a right child  $\frac{i+j}{j}$

Properties of the infinite tree:

- ▶ All fractions are reduced.
- ▶ Every reduced fraction  $\frac{r}{s}$  appears in tree.
- ▶ Every reduced fraction  $\frac{r}{s}$  appears exactly once.
- ▶ The denominator of the  $n$ -th fraction in the list equals the numerator of the  $(n + 1)$ -st.



# Cantor's Theorem

## Theorem

- ▶  $\mathbb{R} \not\approx \mathbb{N}$ .
- ▶ For every set  $A$ ,  $A \not\approx \mathcal{P}(A)$ .

## Proof.

Suppose  $\mathbb{R}$  is countable, say

$$x_1 = 0.\textcolor{red}{7}8790984732689\dots$$

$$x_2 = 0.\textcolor{red}{2}3456789098765\dots$$

$$x_3 = 0.98\textcolor{red}{9}65456756889\dots$$

$$x_4 = 0.237\textcolor{red}{8}9237585022\dots$$

$$x_5 = 0.1234\textcolor{red}{5}438765445\dots$$

⋮

Consider  $x_0 = 0.\textcolor{red}{8}4096\dots?$



## Cantor's Theorem

### Proof.

Consider  $f : A \rightarrow \mathcal{P}(A)$ , and  $B = \{x \in A \mid x \notin f(x)\} \in \mathcal{P}(A)$ , e.g.,

$$f : A \longrightarrow \mathcal{P}(A)$$

$$a \mapsto \{c, d\}$$

$$b \mapsto \{e\}$$

$$c \mapsto \{b, c, d, e\}$$

$$d \mapsto \{\}$$

$$e \mapsto A$$

$$f \mapsto \{a, c, e, g, \dots\}$$

$$g \mapsto \{b, k, m, \dots\}$$

⋮

$$B = \{a, b, d, f, g, \dots\}$$

## Cantor's Theorem

### Proof.

Claim:  $f$  is not onto.

Recall  $B = \{x \in A \mid x \notin f(x)\} \in \mathcal{P}(A)$ . If  $f$  is onto, then  $\exists z \in A$  such that  $f(z) = B \in \mathcal{P}(A)$ , yet

- ▶ If  $z \in B$ , then by definition  $z \notin f(z) = B$ .
- ▶ If  $z \notin B$ , then by definition  $z \in f(z) = B$ .

□

## Also by Cantor

### Theorem

*The set  $\mathbb{R}^2$  of all ordered pairs of real numbers has the same size as  $\mathbb{R}$ .*

Proof. (Julius König).

It suffices to prove that the set of all pairs  $(x, y)$ ,  $0 < x, y \leq 1$ , can be mapped bijectively onto  $(0, 1]$ .

Consider the pair  $(x, y)$  and write  $x, y$  in their unique non-terminating decimal expansion as in the following example:

$$\begin{array}{ccccccccc} x = 0. & 3 & & 0 & 1 & & 2 & & 0 & 0 & 7 & & 0 & 8 & \dots \\ y = 0. & 0 & 0 & 9 & & 2 & & 0 & 5 & & 1 & & 0 & 0 & 0 & 8 & \dots \end{array}$$

Now let

$$z = 0.3\ 009\ 01\ 2\ 2\ 05\ 007\ 1\ 08\ 0008\dots$$

□

(Aigner & Ziegler, p. 133)

# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

## Finite Sets

For any  $n \in \mathbb{N}$ , let  $[n] := \{1, \dots, n\}$ , with  $[0] = \emptyset$ .

### Definition

A set  $A$  is **finite** if it is equinumerous to  $[n]$  for some  $n$ . A set is **infinite** iff it is not finite.

### Example

Any natural number is itself a finite set. Recall that for any  $n \in \mathbb{N}$

$$n = \{0, \dots, n - 1\}$$

### Theorem (Pigeonhole Principle)

*No set of the form  $[n]$  is equinumerous to a proper subset of itself, where  $n \in \mathbb{N}$ .*

# Pigeonhole Principle

## Pigeonhole Principle (that we know)

If there are  $n + 1$  pigeons in  $n$  holes, then some hole contains at least 2 pigeons.



# Pigeonhole Principle

## Theorem (Pigeonhole Principle)

No set of the form  $[n]$  is equinumerous to a proper subset of itself, where  $n \in \mathbb{N}$ .

### Proof (Take 1).

Note that any function  $F$  is a surjection onto its image  $\text{im } F$ , we need to show that

$$\nexists f : [n] \rightarrow [n] (f \text{ injective} \wedge \underbrace{f([n])}_{\substack{\text{im } f \\ f \text{ not surjective}}} \subsetneq [n])$$

$$\begin{aligned} &\Leftrightarrow \forall f : [n] \rightarrow [n] (\neg(f \text{ injective} \wedge \neg(f \text{ surjective}))) \\ &\Leftrightarrow \forall f : [n] \rightarrow [n] (\neg f \text{ injective} \vee f \text{ surjective}) \\ &\Leftrightarrow \forall f : [n] \rightarrow [n] (f \text{ injective} \rightarrow f \text{ surjective}) \end{aligned}$$

See (Gallier, p. 133) for the rest of the proof. □

# Pigeonhole Principle

Proof (by induction).

We want to show that for all  $m, n \in \mathbb{N}$ ,

$$m > n \Rightarrow \nexists f : [m] \rightarrow [n] \text{ bijective}$$

It suffices to show that for all  $m, n \in \mathbb{N}$ ,

$$m > n \Rightarrow \nexists f : [m] \rightarrow [n] \text{ injective}$$

or, for all  $m, n \in \mathbb{N}$ ,

$$m > n \Rightarrow \neg(\exists f : [m] \rightarrow [n] \text{ injective})$$

or equivalently, by considering the contrapositive, for all  $m, n \in \mathbb{N}$ ,

$$(\exists f : [m] \rightarrow [n] \text{ injective}) \Rightarrow (m \leq n)$$

# Pigeonhole Principle

## Proof by Induction.

We proceed by induction on  $n$ .

- ▶ **base case. ( $n = 0$ ):** If  $n = 0$ ,  $[0] = \emptyset$ . If  $f : [m] \rightarrow [n]$  is injective, then the only possibility is that  $[m] = \emptyset$ , hence  $m = 0$ .
- ▶ **inductive case. ( $n \geq 1$ ):** Assume the IH that for all  $m \in \mathbb{N}$

$$\exists f : [m] \rightarrow [n - 1] \text{ injective} \Rightarrow m \leq n - 1$$

We want to show that for all  $m \in \mathbb{N}$

$$\exists f : [m] \rightarrow [n] \text{ injective} \Rightarrow m \leq n$$

Suppose that for all  $m \in \mathbb{N}$ , there exists an injective  $f : [m] \rightarrow [n]$ ,

- ▶ If  $f(i) < n$  for all  $i \in [m]$ , then consider  $g : [m] \rightarrow [n - 1], i \mapsto f(i)$ , which is also injective. Hence by IH  $m \leq n - 1 \leq n$ .

# Pigeonhole Principle

## Proof by Induction (Cont.)

- If  $n \in f([m])$ , say,  $f(i_0) = n$  for some  $i_0 \in [m]$ ,  $m \neq 0$ , then  $n \notin f([m] \setminus \{i_0\})$  (since  $f$  is injective). Define

$$g : [m - 1] \rightarrow [n - 1]$$

$$i \mapsto \begin{cases} f(i), & i < i_0 \\ f(i + 1), & i \geq i_0 \end{cases}$$

which is also injective, since for  $i, j \in [m - 1]$ ,

- $i, j < i_0$ .  $g(i) = g(j) \Rightarrow f(i) = f(j) \Rightarrow i = j$ ;
- $i, j \geq i_0$ .  $g(i) = g(j) \Rightarrow f(i + 1) = f(j + 1) \Rightarrow i = j$ ;
- $i < i_0 \leq j$ .  $g(i) = g(j) \Rightarrow f(i) = f(j + 1) \Rightarrow i = j + 1 \Rightarrow i > j$ .  
Impossible!
- $j < i_0 \leq i$ . Also impossible.

Therefore by IH  $m - 1 \leq n - 1$ , hence  $m \leq n$ .



# Pigeonhole Principle

## Proof by Induction.

We proceed by induction on  $m$ .

- ▶ **base case. ( $m = 0$ ):** If  $m = 0$ ,  $[0] = \emptyset$ . Since  $f : \emptyset \rightarrow [n]$  is injective for all  $n \in \mathbb{N}$ , then trivially  $m \leq n$ .
- ▶ **inductive case. ( $m \geq 1$ ):** Assume the IH that for all  $n \in \mathbb{N}$ ,

$$\exists f : [m - 1] \rightarrow [n] \text{ injective} \Rightarrow m - 1 \leq n$$

We want to show that for all  $n \in \mathbb{N}$

$$\exists f : [m] \rightarrow [n] \text{ injective} \Rightarrow m \leq n$$

Suppose that for all  $m \in \mathbb{N}$ , there exists an injective  $f : [m] \rightarrow [n]$ ,

- ▶ If  $f(i) < n$  for all  $i \in [m]$ , define  $g : [m - 1] \rightarrow [n - 1]$  as  $g = f|[m - 1]$ . Then  $g$  is also injective, hence  $m - 1 \leq n - 1$ , and  $m \leq n$ .

# Pigeonhole Principle

## Proof by Induction (Cont.)

- If  $f(i_0) = n$  for some  $i_0 \in [m]$ , thus  $f([m] \setminus \{i_0\}) \subset [n - 1]$  (since for any other  $i \neq i_0$ ,  $f(i) \neq f(i_0) = k$ ). Define

$$g : [m - 1] \rightarrow [n - 1]$$

$$i \mapsto \begin{cases} f(i), & i \neq i_0 \\ f(m), & i = i_0 \end{cases}$$

then  $g$  is also injective, since

- If  $i, j \neq i_0$ , then  $g(i) = g(j) \Rightarrow f(i) = f(j) \Rightarrow i = j$ .
- If  $i \neq i_0$ , then  $g(i) = f(i) \neq f(m) = g(i_0)$  (since  $i < m \Rightarrow f(i) \neq f(m)$ ).

Therefore by IH  $m - 1 \leq n - 1$ , hence  $m \leq n$ . □



# Finite Sets

## Caveat

Given function  $f : A \rightarrow B$ ,

- ▶ If  $A = \emptyset$ , then any function,  $f : \emptyset \rightarrow B$  is (trivially) injective.
- ▶ If  $B = \emptyset$ , then  $f$  is the **empty function** from  $\emptyset$  to itself, and is (trivially) surjective, hence also bijective.

## Corollary

- ▶ *No finite set is equinumerous to a proper subset of itself.*
- ▶  $\mathbb{N}$  is infinite.
- ▶ *Every finite set is equinumerous to a **unique** natural number.*
- ▶ *Any subset of a finite subset is finite.*

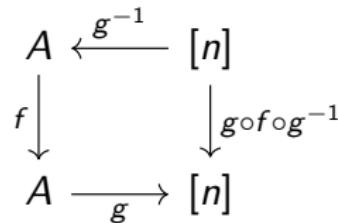
## Finite Sets

### Corollary (Pigeonhole Principle for Finite Sets)

No finite set is equinumerous to a proper subset of itself.

#### Proof.

Since  $A$  is finite, then there exists a bijection  $g : A \rightarrow [n]$  for some  $n \in \mathbb{N}$ . Assume that there exists a bijection  $f$  between  $A$  and some proper subset of  $A$ . Then, consider the function  $g \circ f \circ g^{-1}$ , from  $[n]$  to itself.



Then, note that  $g(a) \in [n] \setminus \text{ran } g \circ f \circ g^{-1}$  for some  $a \in A \setminus \text{ran } f$ , therefore  $g \circ f \circ g^{-1}$  is a bijection from  $[n]$  to some proper subset of itself, which is a contradiction. □

# Finite Sets

## Theorem

Let  $A$  be any finite set. For any function  $f : A \rightarrow A$ , then  $f$  is injective iff  $f$  is surjective.

## Proof.

- ▶ Let  $f$  be an injection. If  $f$  is not surjective, then  $f : A \rightarrow \text{im } f$  is a bijection, hence the finite set  $A$  is equinumerous to a proper subset  $f(A) \subsetneq A$ . Contradiction! Hence  $f$  is surjective.
- ▶ Let  $f$  be a surjection. Then it has a right inverse  $g : A \rightarrow A$  such that

$$f \circ g = id_A$$

Hence  $g$  is injective since it has a left inverse  $f$ . Since  $A$  is finite, then  $g$  is also surjective, hence  $g$  admits a right inverse  $\tilde{f} : A \rightarrow A$  such that  $g \circ \tilde{f} = id_A$ . But  $f = f \circ id_A = f \circ g \circ \tilde{f} = id_A \circ \tilde{f} = \tilde{f}$ , therefore  $f = g^{-1}$  is bijective, hence surjective. □

## Comparing with Linear Transformations

### Example

Let  $L : V \rightarrow W$  be a linear map, with  $\dim V = \dim W < \infty$ , then  $L$  is injective ( $\ker L = \{0\}$ ) iff  $L$  is surjective ( $\text{ran } L = W$ ).

### Proof (Standard).

The result follows from the **rank–nullity theorem**: Given a linear map  $T : V \rightarrow W$ ,

$$\text{rank}(T) + \dim \ker(T) = \dim V$$

□

# Comparing with Linear Transformations

## Example

Let  $L : V \rightarrow W$  be a linear map, with  $\dim V = \dim W < \infty$ , then  $L$  is injective ( $\ker L = \{0\}$ ) iff  $L$  is surjective ( $\text{ran } L = W$ ).

### A Non-standard Proof.

Note that the linear map  $L$  is uniquely determined by its action on the basis vectors. Suppose  $\dim V = \dim W = n$ , then

- If  $L$  is injective and  $\mathcal{B}_V = \{e_1, \dots, e_n\} \subset V$  is a basis for  $V$ , then it is clear that  $|L(\mathcal{B}_V)| = n$ . We need to verify that  $L(\mathcal{B}_V) = \{Le_1, \dots, Le_n\}$  is a basis for  $W$  (such that  $L$  is a surjective linear map). Indeed, since

$$\begin{aligned}\sum_{k=1}^n \alpha_k Le_k = 0 \Rightarrow L\left(\sum_{k=1}^n \alpha_k e_k\right) = 0 && (L \text{ is linear}) \\ \Rightarrow \sum_{k=1}^n \alpha_k e_k = 0 && (L \text{ is injective}) \\ \Rightarrow \alpha_k = 0 \quad \forall k = 1, \dots, n && (\{e_1, \dots, e_n\} \text{ is a basis})\end{aligned}$$

# Comparing with Linear Transformations

## A Non-standard Proof (Cont.)

- If  $L$  is surjective and  $\mathcal{B}_W = \{Le_1, \dots, Le_n\} \subset W$  is a basis for  $W$ , then it is **clear** that  $|L^{-1}(B_W)| = n$ . We need to verify that  $L^{-1}(B_W) = \{e_1, \dots, e_n\}$  is a basis. Indeed,

$$\begin{aligned}\sum_{k=1}^n \alpha_k e_k &= 0 \\ \Rightarrow L\left(\sum_{k=1}^n \alpha_k e_k\right) &= \sum_{k=1}^n \alpha_k L e_k = 0 \\ \Rightarrow \alpha_k &= 0 \quad \forall k = 1, \dots, n \quad (\text{b/c } \{Le_1, \dots, Le_n\} \text{ is a basis})\end{aligned}$$

Therefore  $L^{-1}(0) = \{0\}$ , i.e.,  $\ker L = \{0\}$ . It follows that  $L$  is an injective linear map. □

# Pigeonhole Principle

## Other versions of pigeonhole principle

Let  $r, s \in \mathbb{N} - \{0\}$ , if a set containing at least  $rs + 1$  elements is partitioned into  $r$  subsets, then some subsets contains at least  $s + 1$  elements.

### Example

- ▶ In any group of  $12 \cdot 2 + 1 = 25$  people, at least three were born in the same month.
- ▶ At least two people in London have the same number of hairs on their heads.

## Application of Pigeonhole Principle

### Example

Let  $S \subset \{1, 2, \dots, 200\}$  with  $|S| = 101$ , then  $S$  contains two consecutive integers.

### Proof.

Consider the following sets,

$$S_1 = \{1, 2\}$$

$$S_2 = \{3, 4\}$$

⋮

$$S_{99} = \{197, 198\}$$

$$S_{100} = \{199, 200\}$$

The rest follows by applying the Pigeonhole principle. □



## Application of Pigeonhole Principle

### Example

Let  $S \subset \{1, 2, \dots, 200\}$  with  $|S| = 101$ , then  $S$  contains two integers that one divides the other.

### Proof.

Consider the following sets,

$$S_1 = \{1, 2, 4, 8, \dots, 64, 128\} = \{1, 2, 2^2, 2^3, \dots, 2^7\}$$

$$S_3 = \{3, 6, 12, 24, \dots, 96, 192\} = \{3, 3 \cdot 2, 3 \cdot 2^2, 3 \cdot 2^3, \dots, 3 \cdot 2^6\}$$

$$S_5 = \{5, 10, 20, 40, 80, 160\} = \{5, 5 \cdot 2, 5 \cdot 2^2, 5 \cdot 2^3, 5 \cdot 2^4, 5 \cdot 2^5\}$$

⋮

$$S_{49} = \{49, 98, 196\} = \{49, 49 \cdot 2, 49 \cdot 2^2\}$$

⋮

$$S_{99} = \{99, 198\} = \{99, 99 \cdot 2\}$$

$$S_{101} = \{101\}, \dots, S_{199} = \{199\}$$

The rest follows by applying the Pigeonhole principle. □

# Erdős–Szekeres Theorem

## Theorem (Erdős–Szekeres, 1935)

Let  $A = (a_1, \dots, a_n)$  be a sequence of  $n$  different real numbers. If  $n \geq sr + 1$  then either  $A$  has an increasing subsequence of  $s + 1$  terms or a decreasing subsequence of  $r + 1$  terms (or both).

“The slickest and most systematic” proof (Seidenberg 1959).

### Proof.

Define a function  $f : \mathbb{R} \rightarrow \{1, \dots, n\}^2$ ,  $a_i \mapsto (x_i, y_i)$ , where

- ▶  $x_i$  is the number of terms in the longest increasing subsequence ending at  $a_i$ ,
- ▶  $y_i$  is the number of terms in the longest decreasing subsequence starting at  $a_i$ .

## Erdős–Szekeres Theorem

## Proof. (Cont.)

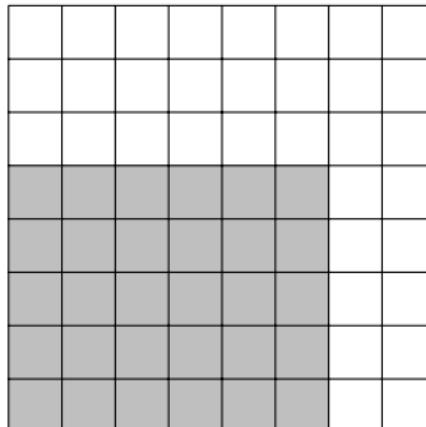
Claim: The mapping  $a_i \mapsto (x_i, y_i)$  is injective, i.e.,

$$\forall i, j \in \{1, \dots, n\}, a_i \neq a_j \Rightarrow (x_i, y_i) \neq (x_j, y_j).$$

Indeed, for a subsequence  $\dots a_i \dots a_j \dots$ , either

- ▶  $a_i < a_j \Rightarrow x_i < x_j$ , or
  - ▶  $a_i > a_j \Rightarrow y_i > y_j$ ,

The rest follows by Pigeonhole principle.



## 300. Longest Increasing Subsequence<sup>2</sup>

Given an integer array `nums`, return the length of the longest strictly increasing subsequence.

A subsequence is a sequence that can be derived from an array by deleting some or no elements without changing the order of the remaining elements.

For example, `[3, 6, 2, 7]` is a subsequence of the array `[0, 3, 1, 6, 2, 2, 7]`.

### Example 1

- ▶ **Input:** `nums = [10, 9, 2, 5, 3, 7, 101, 18]`
- ▶ **Output:** 4
- ▶ **Explanation:** The longest increasing subsequence is `[2, 3, 7, 101]`, therefore the length is 4.

### Example 2

- ▶ **Input:** `nums = [0,1,0,3,2,3]`
- ▶ **Output:** 4

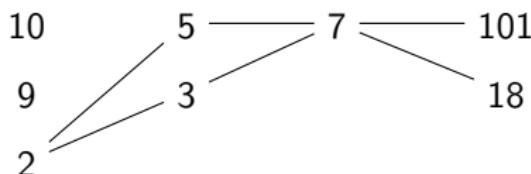
### Example 3

- ▶ **Input:** `nums = [7,7,7,7,7,7,7]`
- ▶ **Output:** 1

## 300. Longest Increasing Subsequence

### Example 1 (Patience Sort)

- ▶ **Input:**  $\text{nums} = [10, 9, 2, 5, 3, 7, 101, 18]$
- ▶ **Output:** 4
- ▶ **Explanation:** A longest increasing subsequence is  $[2, 3, 7, 101]$ , therefore the length is 4.



### Observation

- ▶ Each column is a decreasing subsequence.
- ▶ The length of any increasing subsequence is at most the number of columns (pigeonhole principle).

# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Partial Order

## Definition

An **ordered set** (or **partially ordered set** or **poset**) is an ordered pair  $(P, \leq)$  of a set  $P$  and a binary relation  $\leq$  contained in  $P \times P$ , called the **order** (or the **partial order**) on  $P$  such that  $\leq$  is

- ▶ reflexive:  $a \leq a \Rightarrow \top$
- ▶ antisymmetric:  $a \leq b \wedge b \leq a \Rightarrow a = b$
- ▶ transitive:  $a \leq b \wedge b \leq c \Rightarrow a \leq c$

We write  $x < y$  if  $x \leq y$  and  $x \neq y$ . (Other notation:  $\preceq$  and  $\prec$ )

## Definition

If  $(P, \leq)$  is a poset, and for all  $x, y \in P$ , either  $x \leq y$  or  $y \leq x$ , then it is a **total order** or **linear order**.

## Example of linear/total order

- ▶  $\mathbb{Z}$
- ▶  $\mathbb{Q}$
- ▶  $\mathbb{N}$
- ▶  $\mathbb{R}$

# Partial Order

Pre-order/Quasi-order

- ▶ reflexive:  $a \leq a \Rightarrow \top$
- ▶ transitive:  $a \leq b \wedge b \leq c \Rightarrow a \leq c$

Partial Order

- ▶ antisymmetric:  $a \leq b \wedge b \leq a \Rightarrow a = b$

Total Order

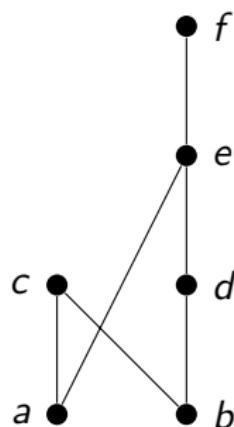
- ▶ total:  $a \leq b \vee b \leq a \Rightarrow \top$

# Hasse Diagram

Hasse/Order Diagram (Idea: keep the most essential component.)

- ▶ Edges are the cover pairs  $(x, y)$  with  $x$  covered by  $y$ ;
- ▶ Edges are drawn such that  $x$  is below  $y$ ;
- ▶ Edges are monotone vertically.

## Example

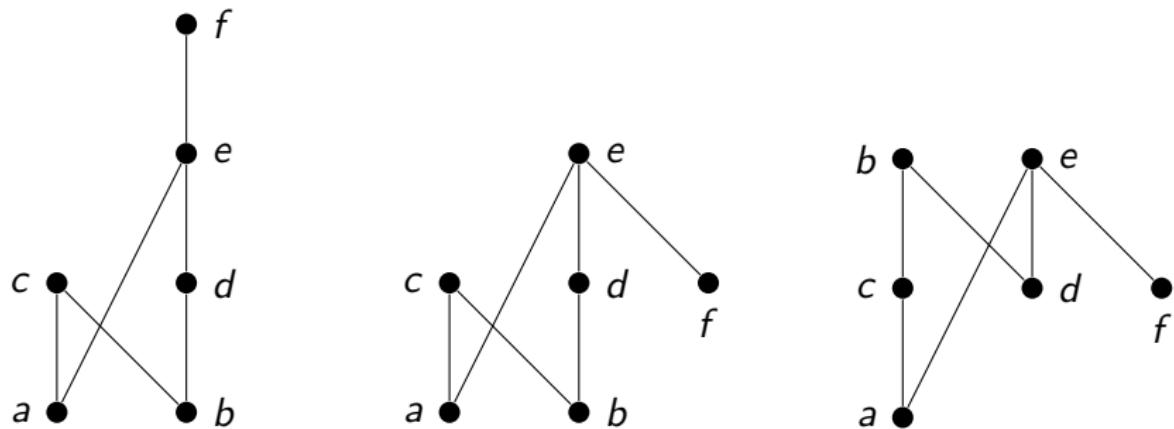


What relation does the Hasse diagram on the left corresponds to?

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

# Hasse Diagrams of Three Different Posets

## Example

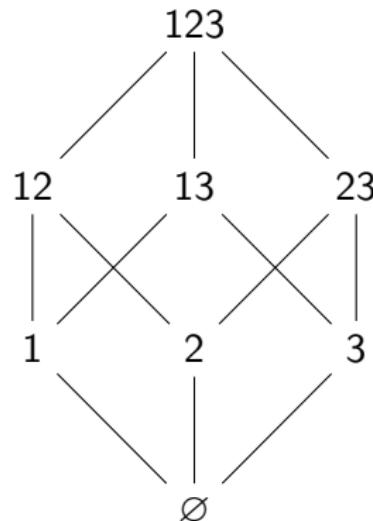
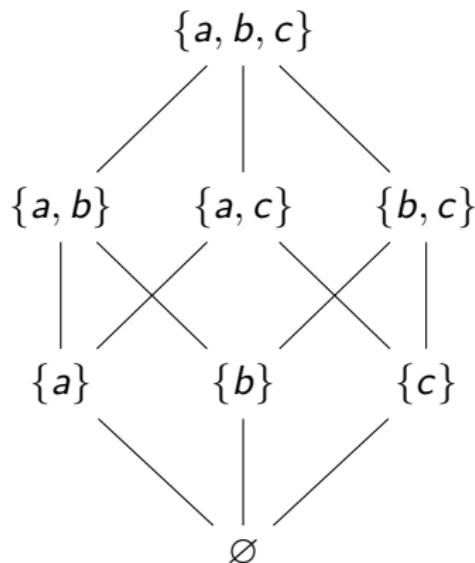


Note that all three are the same as graphs, but not as posets.

# Partial Order

## Examples

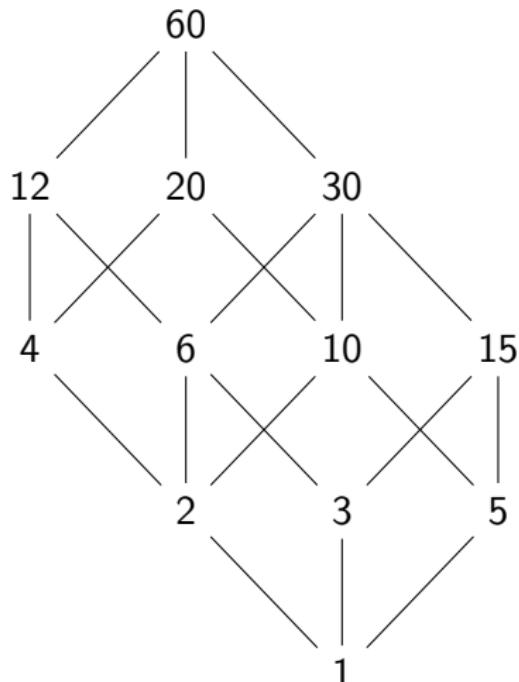
- ▶ Power set/Boolean lattice  $(2^{[n]}, \subseteq)$ .  $[n] = \{1, \dots, n\}$ , subsets of  $[n]$  ordered by inclusion.



# Partial Order

## Examples

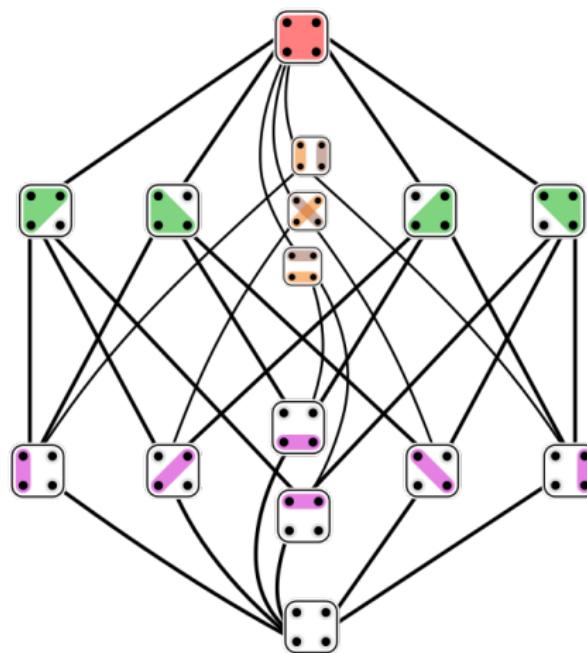
- Divisors of  $n \in \mathbb{N}$ .  $(\mathbb{N}, |)$ . Ordered by divisibility.  $n = 60$ .



# Partial Order

## Examples

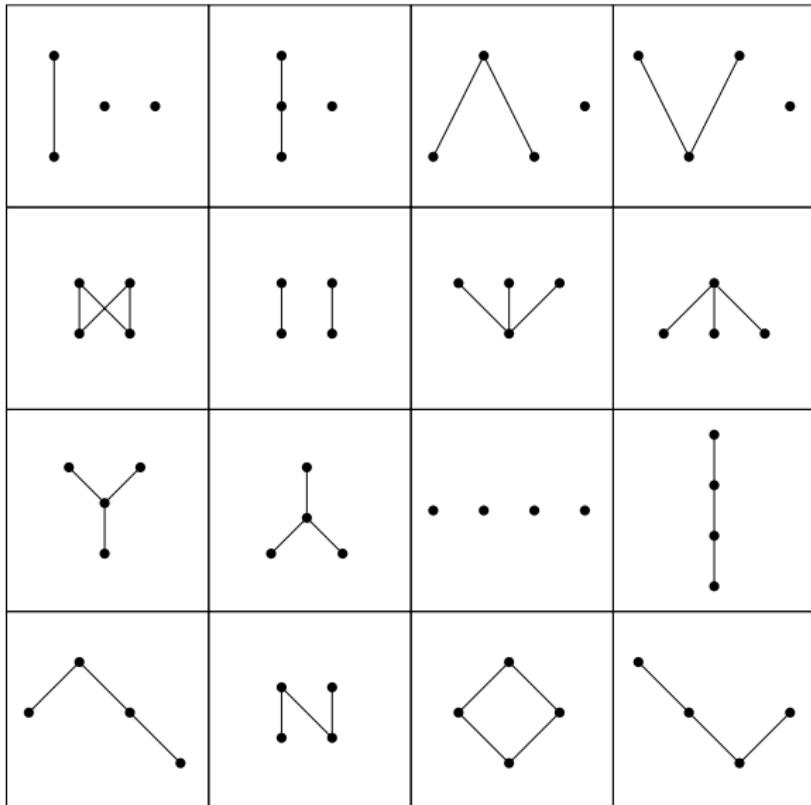
- ▶ Partition of  $[n] = \{1, \dots, n\}$ , ordered by refinement.



# Partial Order

## Examples

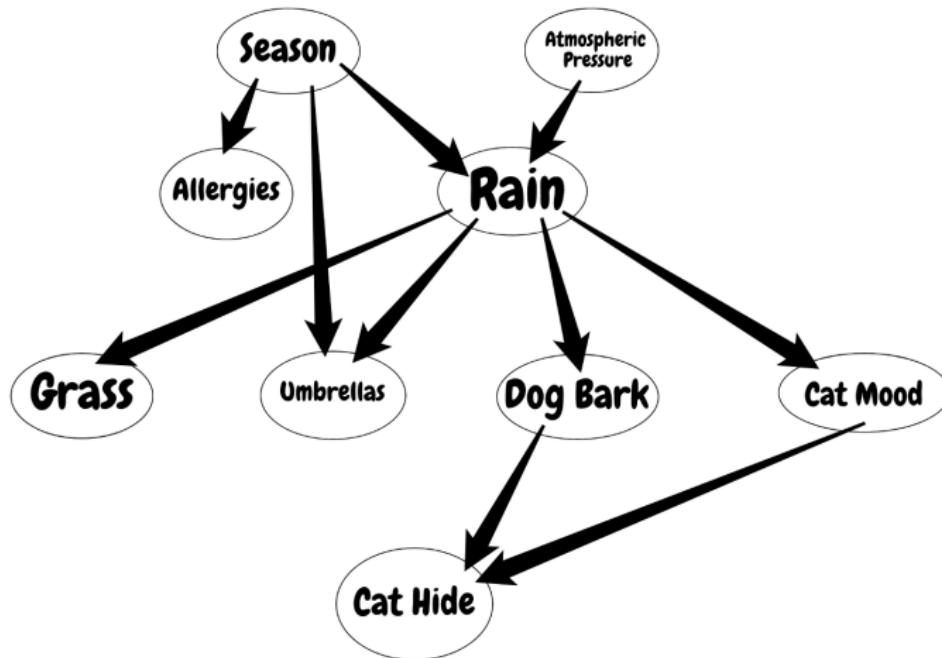
- ▶ All posets on a set with 4 elements (up to relabeling of the points).



# Partial Order

## Examples

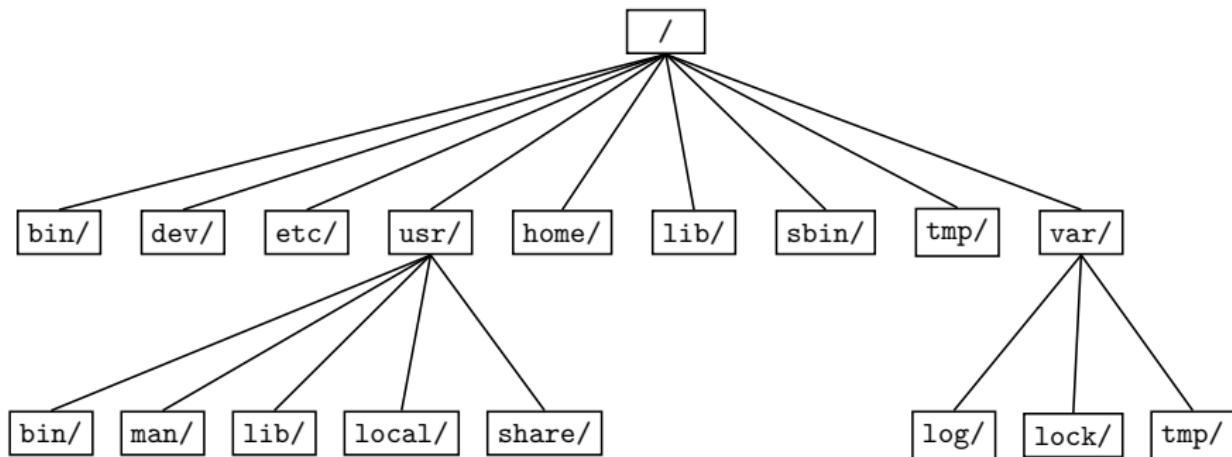
- ▶ Any directed acyclic graph (DAG), e.g., Bayesian network.



# Partial Order

## Examples

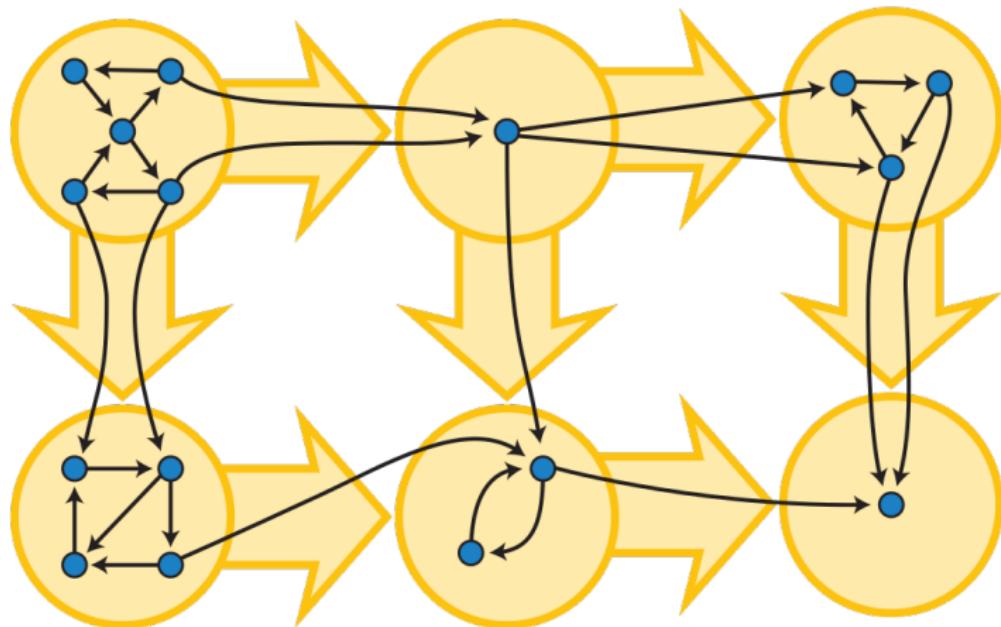
- ▶ Vertices in a rooted tree (e.g., computer directory structure, family tree).



# Partial Order

## Examples

- ▶ Strongly connected components in a directed graph. (cf., preorder)



# Partial Order

## Examples

- ▶ sub-trees/graphs/groups/vector spaces of a trees/graphs/groups/vector spaces.

## Non-example

- ▶  $(\mathbb{Z}, |)$ .  $-1|1$  and  $1|-1$ , but  $1 \neq -1$ .

# Covers in a Poset

## Definition

Let  $P$  be an ordered set. Then  $y \in P$  is called a cover of  $x \in P$  if  $x < y$  and for all  $z \in P$ ,  $x \leq z \leq y$  implies  $z \in \{x, y\}$ . We also say that  $y$  covers  $x$ , or  $x$  is covered by  $y$ . Such  $x$  and  $y$  are called **adjacent**.

## Examples

- ▶ In  $(\mathcal{P}([6]), \subseteq)$ ,  $\{1, 3\}$  is covered by  $\{1, 3, 5\}$ , but not covered by  $\{1, 2, 3, 4\}$ .
- ▶ In  $\mathbb{Z}$ , each  $k \in \mathbb{Z}$  is covered by  $k + 1$ , and covers  $k - 1$ .
- ▶ In  $(\mathbb{N}, |)$ , 15 is covered by 105, 14 is not covered by 84.
- ▶ In  $\mathbb{R}$  and  $\mathbb{Q}$ , no two elements are covers of each other.

## More Definitions

### Definition

Let  $(P, \leq)$  be a poset, and  $a, x, y, z \in P$ .

- ▶ If  $a \in P$  but  $\nexists x \in P$  such that  $x < a$ , then  $a$  is a **minimal element**.
- ▶ If  $a \leq x$  for all  $x \in P$ , then  $a$  is the **minimum element**.
- ▶ If  $z \in P$  but  $\nexists x \in P$  such that  $z < x$ , then  $z$  is a **maximal element**.
- ▶ If  $x \leq z$  for all  $x \in P$ , then  $z$  is the **maximum element**.
- ▶ If either  $x < y$  in  $P$  or  $y < x$  in  $P$ , then  $x$  and  $y$  are **comparable** in  $P$ , otherwise  $x$  and  $y$  are **incomparable**.

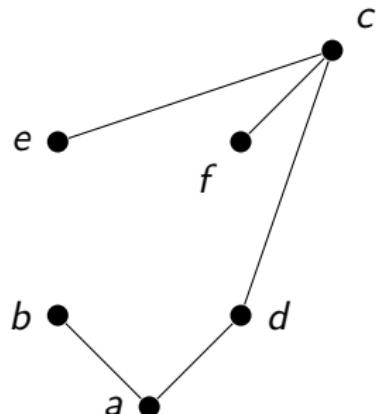
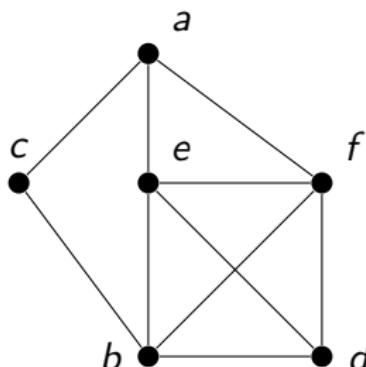
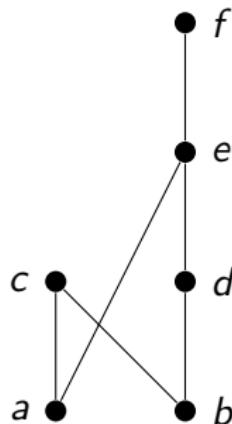
### Definition

Given a poset  $(P, \leq_P)$  and  $Q \subset P$ , then the (binary) relation  $\leq_Q = \leq_P|_{Q \times Q}$  is a partial order on  $Q$ . The induced poset  $(Q, \leq_Q)$  is called **subposet** of  $(P, \leq_P)$ .

# Comparability and Incomparability Graphs

With a poset  $(P, \leq)$ , we associate a **comparability graph**  $G_1 = (P, E_1)$  and an **incomparability graph**  $G_2 = (P, E_2)$ , where  
 $E_1 = \{\{x, y\} \in \binom{P}{2} \mid x, y \text{ comparable}\}$  and  
 $E_2 = \{\{x, y\} \in \binom{P}{2} \mid x, y \text{ incomparable}\}.$

## Example

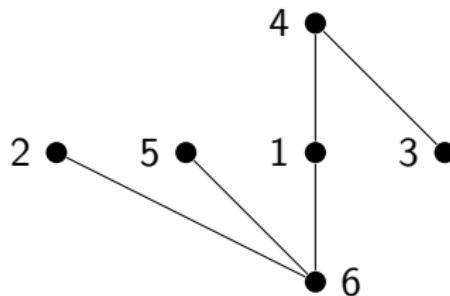


Note that a comparability graph and an incomparability graph are complement graph of each other. The **complement** of graph  $G = (V, E)$  is  $\overline{G} = (V, \binom{V}{2} - E)$ .

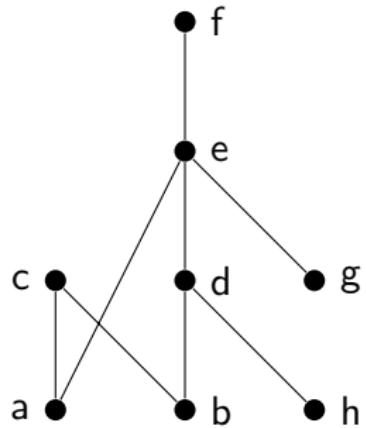
## An Example

Let  $P = \{1, 2, 3, 4, 5, 6\}$ , and  $\leq = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6), (6, 1), (6, 4), (1, 4), (6, 5), (3, 4), (6, 2)\}$ . Then

- ▶ 6 and 3 are minimal elements.
- ▶ 2, 4, and 5 are maximal elements.
- ▶ 4 is comparable to 6.
- ▶ 2 is incomparable to 3.
- ▶ 1 covers 6, and 3 is covered by 4.
- ▶  $4 > 6$  but 4 does not cover 6.



## Another Example



- ▶ c and f are maximal elements.
- ▶ a, b, g, and h are minimal elements.
- ▶ a is comparable to f.
- ▶ c is incomparable to h.
- ▶ e covers a, and h is covered by d.
- ▶  $e > h$  but e does not cover h.

# Chains and Antichains

## Definition

Given  $(P, \leq)$  poset,

- ▶ A **chain** in a poset is a subset  $C \subset P$  such that any two elements are comparable.
- ▶ An **antichain** in a poset is a subset  $A \subset P$  of incomparable elements.

## Definition

A graph  $G = (V, E)$  is called a **clique** or **complete graph** if  $E = \binom{V}{2}$ .

Conversely, the complement graph of  $G = (V, \binom{V}{2})$ , given by  $\overline{G} = (V, \emptyset)$ , is called an **independent graph** or **independent set**.

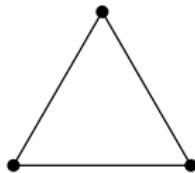
## Remark

- ▶ The comparability graph of a chain is a complete graph.
- ▶ The comparability graph of an antichain is an independent graph.

# Complete Graphs and Independent Graphs

## Complete Graphs $K_n$

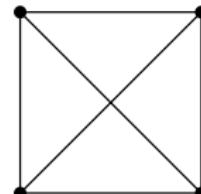
•



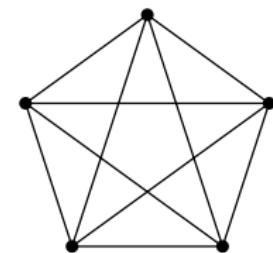
$K_1$

$K_2$

$K_3$



$K_4$



$K_5$

## Independent Graphs $I_n$

|       |       |       |       |       |
|-------|-------|-------|-------|-------|
| •     | •     | •     | •     | •     |
| $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ |

# Chains and Antichains

## Lemma

*Given a chain  $C$  and an antichain  $A$  of a poset,  $|A \cap C| \leq 1$ .*

## Proof.

If  $|A \cap C| \geq 2$ , then we can find two elements that are both comparable and incomparable. Contradiction. □

# Chains and Antichains

## Definition

A chain  $C$  in  $P$  is

- ▶ **maximal** if there exists no chain  $C'$  such that  $C \subsetneq C'$ .
- ▶ **maximum** if for all chain  $C'$ ,  $|C| \not< |C'|$ .

The **height** (not *length*) of a poset  $P$ , denoted by  $h(P)$ , is the maximum size of a chain in  $P$ .

## Definition

An antichain  $A$  in  $P$  is

- ▶ **maximal** if there exists no antichain  $A'$  such that  $A \subsetneq A'$ .
- ▶ **maximum** if for all chain  $A'$ ,  $|A| \not< |A'|$ .

The **width** of a poset  $P$ , denoted by  $w(P)$ , is the maximum size of an antichain in  $P$ .

## Remark

A maximal chain or maximal antichain CANNOT be prolonged by adding a new element.

# Chains and Antichains

## Observation

By pigeonhole principle,

- ▶ If  $P$  can be partitioned into  $t$  antichains, then the height of  $P$  is at most  $t$ .
- ▶ If  $P$  can be partitioned into  $s$  chains, then the width of  $P$  is at most  $s$ .

## Observation

The set of maximal (or minimal) elements is an antichain.

## Theorem (Mirsky's Theorem, 1971)

*A poset of height  $h$  can be partitioned into  $h$  antichains.*

## Proof.

Recursively remove the set of maximal (or minimal) elements.



## Mirsky's Theorem (dual Dilworth)

Proof. (a little more detail).

Denote the set of minimal elements of  $(P, \leq)$  by  $\text{Min}(P)$ . Similarly for  $\text{Max}(P)$ . Thus we have a partition of  $P$  into antichains  $A_1, \dots, A_k$ ,  $k \in \mathbb{N}$ . Since  $|A_i \cap C| \leq 1$  for any chain  $C \subset P$ , then (recall observation)

$$\begin{aligned} k &\geq \max\{|C| : C \text{ is a chain in } P\} \\ &= h(P) \end{aligned}$$

Claim: a chain of length  $k$  can be traced back from  $A_k$ .

Indeed, choose  $x_k \in A_k$ , then  $\exists x_{k-1} \in A_{k-1}$  such that  $x_{k-1} < x_k$ , and so on. Eventually, we have  $x_1 < x_2 < \dots < x_{k-1} < x_k$ . Therefore  $h(P) = k$ .  $\square$

---

**Input:** A partial order  $(P, \leq)$

**Output:** An antichain partition of  $(P, \leq)$

```
1  $i \leftarrow 1$ 
2 while  $P \neq \emptyset$  do
3    $A_i \leftarrow \text{Min}(P)$ 
4    $P \leftarrow P - A_i$ ;
5    $i \leftarrow i + 1$ 
6 end
7 return  $\{A_1, \dots, A_{i-1}\}$ 
```

---

## Dilworth's Theorem

Theorem (Dilworth's Theorem, 1950)

A poset of width  $w$  can be partitioned into  $w$  chains.

Proof.

We use induction on the size of the poset  $P$ .

- ▶ True when  $|P| = 1$ .
- ▶ Assume the theorem is true when  $|P| \leq k$ , then consider a poset  $P$  with  $|P| = k + 1$ , then for each maximal antichain  $A$ , define the downset of  $A$

$$D(A) := \{x \mid x < a \text{ for some } a \in A\}$$

and the upset of  $A$

$$U(A) := \{x \mid x > a \text{ for some } a \in A\}$$

# Dilworth's Theorem

## Proof (Cont.)

**Case I.** Assume there exists a maximum antichain  $A$  with  $D(A) \neq \emptyset$  and  $U(A) \neq \emptyset$ .

Claim:  $\{A, D(A), U(A)\}$  form a partition of  $P$ .

It suffices to show that  $D(A) \cap U(A) = \emptyset$ . Indeed, otherwise let  $x \in D(A) \cap U(A)$ , then  $\exists y \in A$  with  $x < y$ , and  $\exists z \in A$  with  $z < x$ , resulting  $y, z \in A$  comparable, contradiction.

Let  $A = \{a_1, \dots, a_w\}$ , note that  $|A \cup D(A)| \leq k$  and  $|A \cup U(A)| \leq k$ , thus by induction hypothesis, we obtain a chain partition  $\{D_1, \dots, D_w\}$  of  $A \cup D(A)$  with maximal elements  $a_1, \dots, a_w$ .

Similarly we can obtain a chain partition  $\{U_1, \dots, U_w\}$  of  $A \cup U(A)$  with minimal elements  $a_1, \dots, a_w$ .

Glue the chains respectively, we have a chain partition  $\{D_1 \cup U_1, \dots, D_w \cup U_w\}$  of  $P$ .

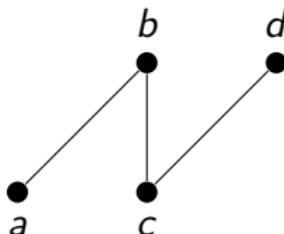
# Dilworth's Theorem

## Proof (Cont.)

**Case II.** Otherwise for every maximum antichain  $A$ , either  $D(A)$  or  $U(A)$  is empty. (Or equivalently, either  $D(A) \cup A = P$  or  $U(A) \cup A = P$  for every maximum antichain  $A$ ). Hence each maximum antichain is either the set of minimal or maximal elements of  $P$ .

Choose  $x \in \text{Min}(P)$  and  $y \in \text{Max}(P)$  with  $x \leq y$  (Possibly  $x = y$ ), then  $\{x, y\}$  is a chain.

Now  $|P - \{x, y\}| \leq k$  and  $P - \{x, y\}$  is of width  $w - 1$  (since each antichain of size  $k$  contains  $x$  or  $y$ ), hence by induction hypothesis,  $P - \{x, y\}$  can be partitioned into  $w - 1$  chains. Add chain  $\{x, y\}$  to obtain the  $w$ -chain partition of  $P$ . □



# An Application of Dilworth's Theorem

Theorem (Erdős–Szekeres, 1935)

Let  $A = (a_1, \dots, a_n)$  be a sequence of  $n$  **different** real numbers. If  $n \geq sr + 1$  then either  $A$  has an increasing subsequence of  $s + 1$  terms or a decreasing subsequence of  $r + 1$  terms (or both).

Proof by Dilworth's Theorem.

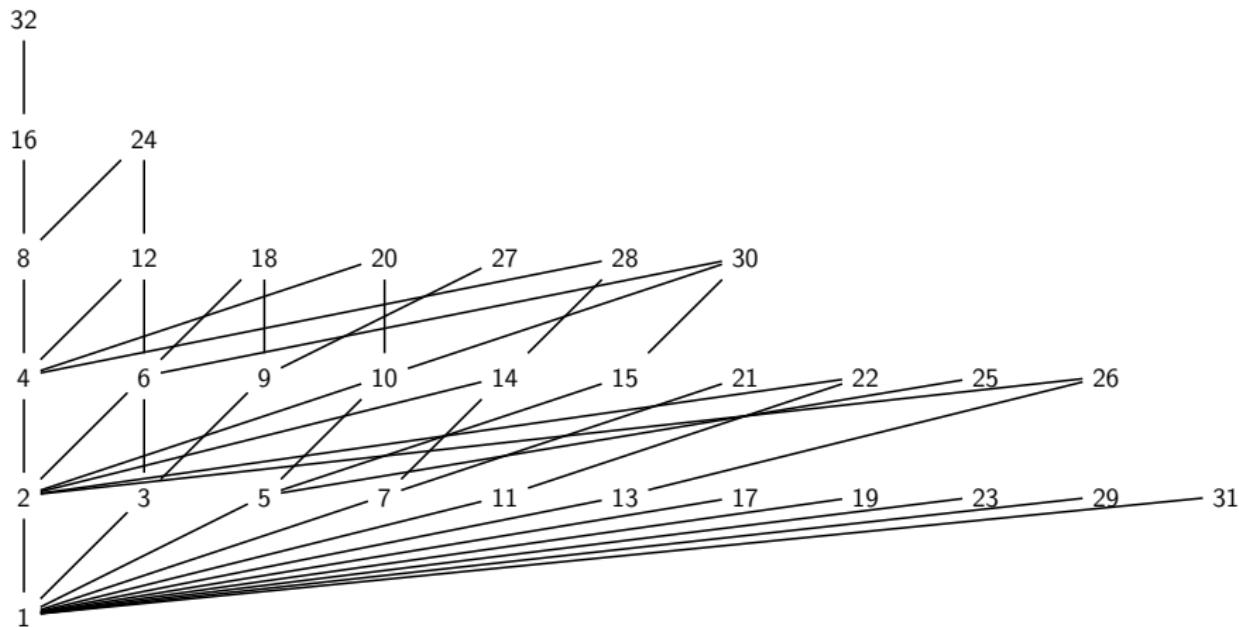
Define the partial order  $\preceq$  on  $A$  by  $a_i \preceq a_j$  iff  $a_i \leq a_j$  and  $i \leq j$ . (Check it!) Then we can observe that an increasing subsequence of  $A$  corresponds to a chain in  $(A, \preceq)$ , and a decreasing subsequence in  $A$  corresponds to an antichain in  $(A, \preceq)$ .

Assume that there is no decreasing subsequence of length  $r + 1$ , then by Dilworth's Theorem, the poset  $(A, \preceq)$  can be **partitioned** into  $k$  chains  $C_1, \dots, C_k$ , with  $k \leq r$ . Therefore  $|C_1| + \dots + |C_k| = n \geq sr + 1$ .

By pigeonhole principle, there exists a chain  $C_j$  with  $|C_j| \geq s + 1$ , which corresponds to an increasing subsequence of length at least  $s + 1$ . □

## Divisibility Revisited

Consider the set  $[32] = \{1, 2, \dots, 31, 32\}$ , ordered by divisibility.



# Several Equivalent Major Theorems in Combinatorics

- ▶ König's Theorem
- ▶ Menger's Theorem (1929)
- ▶ Max-Flow Min-Cut theorem
- ▶ König-Egerváry theorem (1931)
- ▶ Birkhoff-Von Neumann Theorem (1946)
- ▶ ***Hall's Theorem***
- ▶ ***Dilworth's Theorem***

and duality in linear programming.

# Least Upper and Greatest Lower Bounds

## Definition

Let  $P$  be poset and let  $A \subset P$ . Then

- ▶ The element  $u$  is called the **least upper bound (lub)** or **supremum** or **join** of  $A$  if  $u \geq A$  and, for all  $p \in P$  with  $p \geq A$ , we have  $p \geq u$ .
- ▶ The element  $l$  is called the **greatest lower bound (glb)** or **infimum** or **meet** of  $A$  if  $u \leq A$  and, for all  $p \in P$  with  $p \leq A$ , we have  $p \geq l$ .

## Notation

The supremum of a set  $A$  (if exists) is denoted by  $\bigvee A$  and the infimum (if exists) is denoted by  $\bigwedge A$ . For finite sets  $A = \{a_1, \dots, a_n\}$ , we will also use the notation

$$a_1 \vee a_2 \vee \cdots \wedge a_n := \bigvee \{a_1, \dots, a_n\}$$

$$a_1 \wedge a_2 \wedge \cdots \wedge a_n := \bigwedge \{a_1, \dots, a_n\}$$

# Least Upper and Greatest Lower Bounds

## Theorem

Let  $P$  be a poset and let  $A \subset P$  be a subset that has a least upper bound. Then **the** least upper bound of  $A$  is unique.

## Proof.

Let  $A \subset P$  and  $u, u'$  be least upper bounds of  $A$ . Then  $u \leq u'$  and  $u' \leq u$ , hence  $u = u'$ . □

Similar for the greatest lower bound.

# Least Upper and Greatest Lower Bounds

## Examples

- ▶ The poset  $\{x \in \mathbb{Q} \mid x \geq 0, x^2 \leq 2\}$  of  $\mathbb{Q}$  has upper bounds but no least upper bound in  $\mathbb{Q}$ .
- ▶ Every nonempty subset of  $\mathbb{R}$  that has an upper bound has a least upper bound.
- ▶ Let  $P$  be an arbitrary poset. The empty set  $\emptyset$ , as a subset of  $P$ , has a supremum in  $P$  iff  $P$  has a smallest element. The empty set  $\emptyset$  has an infimum iff  $P$  has a largest element.
- ▶ If  $X$  is a set and  $\mathcal{P}(X)$  its power set ordered by inclusion, for each subsets  $A \subset \mathcal{P}(X)$ , we have  $\bigvee A = \bigcup A$  and  $\bigwedge A = \bigcap A$ .

# Lattice

## Definition

Let  $L$  be an ordered set. Then  $L$  is called a *lattice* if any two elements of  $L$  have a supremum and an infimum.  $L$  is called a *complete lattice* iff any subset of  $L$  has a supremum and an infimum.

## Examples

- ▶ Every chain is a lattice, but not every chain is a complete lattice (consider  $\mathbb{N}$ ).
- ▶ If  $X$  is a set, then the power set  $\mathcal{P}(X)$  ordered by set inclusion is a complete lattice.
- ▶ The poset  $\mathbb{N} - \{0\}$  under the divisibility ordering is a lattice.
- ▶ Any nonempty finite lattice is a complete lattice

# Properties of Lattice

## Theorem

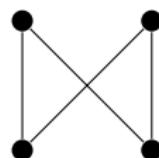
If  $L$  is a lattice, then the following identities hold for all  $a, b, c \in L$ .

- ▶  $a \vee b = b \vee a,$   $a \wedge b = b \wedge a$
- ▶  $(a \vee b) \vee c = a \vee (b \vee c),$   $(a \wedge b) \wedge c = a \wedge (b \wedge c)$
- ▶  $a \vee a = a,$   $a \wedge a = a,$
- ▶  $(a \vee b) \wedge a = a,$   $(a \wedge b) \vee a = a.$

## Remark

Given the Hasse diagram of a lattice,

- ▶ it must contain a least element (denoted by  $\perp$  or  $0$ ), and a greatest element (denoted by  $\top$  or  $1$ ).
- ▶ it must not contain a butterfly (or four-crown).



# Table of Contents

1. Sets (Naive)
2. Logic
3. Induction
4. Relations and Functions
5. Numbers and Equinumerosity
6. Finite Sets and Pigeonhole Principle
7. Partial Order
8. Cardinality

# Cardinals

## Cardinal Number

For any set  $A$ , we will define a set  $\text{card } A$  such that

- ▶ For any sets  $A$  and  $B$ ,

$$\text{card } A = \text{card } B \Leftrightarrow A \approx B$$

- ▶ For a finite set  $A$ ,  $\text{card } A$  is the natural number for which  $A \approx n$ .

## Example

- ▶ Each  $n \in \mathbb{N}$  is a cardinal, e.g.,  $\text{card}\{a, b, c, d\} = 4$  since  $\text{card}\{a, b, c, d\} \approx 4$ .

# Cardinality

## Cardinality

For every  $A$ , there is a unique cardinal  $\kappa$  with  $A \approx \kappa$ . We call that  $\kappa$  the **cardinality** of  $A$ , denoted by  $\text{card } A = \kappa$ .

## Example

- ▶  $\text{card } n = n$ .
- ▶  $\text{card } \mathbb{N} = \aleph_0$  (by Cantor).
- ▶  $\text{card } \mathbb{R} = 2^{\aleph_0}$ .

## Caution

$\{X \mid \text{card } X = \kappa\}$  is NOT a set, except for  $\kappa = 0$ .

# Cardinal Arithmetic

## Definition

Let  $\kappa$  and  $\lambda$  be cardinals, then

- ▶  $\kappa + \lambda = \text{card}(K \cup L)$ , where  $\kappa = \text{card } K$ ,  $\lambda = \text{card } L$ , and  $K \cap L = \emptyset$ .
- ▶  $\kappa \cdot \lambda = \text{card}(K \times L)$ , where  $\kappa = \text{card } K$ ,  $\lambda = \text{card } L$ .
- ▶  $\kappa^\lambda = \text{card}(K^L)$ , where  $\kappa = \text{card } K$ ,  $\lambda = \text{card } L$ , and  $K^L := \{f \mid f : L \rightarrow K\}$ .

## Examples

- ▶  $2 + 2 = 4$ .
- ▶ for  $m, n \in \mathbb{N}$ ,  $m \times n = \text{card}(m \times n)$ , and  $m^n = \text{card}(m^n)$ .
- ▶ For any  $n \in \mathbb{N} - \{0\}$ ,  $n + \aleph_0 = \aleph_0$ , and  $n \cdot \aleph_0 = \aleph_0$ .
- ▶  $\aleph_0 + \aleph_0 = \aleph_0$  and  $\aleph_0 \times \aleph_0 = \aleph_0$ .
- ▶ For any cardinal number  $\kappa$ ,  $\kappa + 0 = \kappa$ ,  $\kappa \cdot 0 = 0$ , and  $\kappa \cdot 1 = \kappa$ .

# Cardinal Arithmetic

## Examples

- ▶  $\kappa^0 = 1$  for any  $\kappa$ , since  $K^\emptyset = \{\emptyset\}$  for any set  $K$ .
  - ▶  $0^0 = 1$ , since  $\emptyset^\emptyset = \{\emptyset\}$ .
- ▶  $0^\kappa = 0$  for any nonzero  $\kappa$ , since  $\emptyset^K = \emptyset$  for nonempty set  $K$ .
- ▶  $2^{\text{card } A} = \text{card}(2^A) = \text{card } \mathcal{P}(A)$ , since  $2^A \approx \mathcal{P}(A)$ .
  - ▶  $\text{card } \mathcal{P}\mathbb{N} = 2^{\aleph_0}$ .
- ▶  $\kappa \neq 2^\kappa$  for any  $\kappa$ , and  $\aleph_0 \neq 2^{\aleph_0}$ . (Cantor's Theorem)
- ▶ For any cardinal  $\kappa$ ,  $\kappa + \kappa = 2 \cdot \kappa$ .

# Cardinal Arithmetic

## Theorem

For any cardinal numbers  $\kappa$ ,  $\lambda$ , and  $\mu$ ,

- ▶  $\kappa + \lambda = \lambda + \kappa$ , and  $\kappa \cdot \lambda = \lambda \cdot \kappa$ .
- ▶  $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$ , and  $\kappa \cdot (\lambda \cdot \mu) = (\kappa \cdot \lambda) \cdot \mu$ .
- ▶  $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$ .
- ▶  $\kappa^{\lambda+\mu} = \kappa^\mu \cdot \kappa^\mu$ .
- ▶  $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$ .
- ▶  $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$ .

## Theorem

Let  $m, n$  be finite cardinals. Then

- ▶  $m + n = m +_{\mathbb{N}} n$ .
- ▶  $m \cdot n = m \cdot_{\mathbb{N}} n$ .
- ▶  $m^n = m^n$ .

# Cardinal Arithmetic

## Theorem

If  $A$  and  $B$  are finite, then  $A \cup B$ ,  $A \times B$ , and  $A^B$  are also finite.

## Proof.

Let  $m = \text{card } A$  and  $n = \text{card } B$ , then

- ▶  $\text{card}(A \times B) = \text{card } A \cdot \text{card } B = m \cdot n = m \cdot_{\mathbb{N}} n \in \mathbb{N}.$
- ▶  $\text{card } A^B = (\text{card } A)^{\text{card } B} = m^n \in \mathbb{N}.$
- ▶ Note  $A \cup B = A \cup (B - A)$  with  $A \cap (B - A) = \emptyset$ .

Also note that  $B - A$  is finite since  $B - A \subset B$  and  $B$  is finite.

Let  $k = \text{card}(B - A)$ , then  $\text{card}(A \cup B) = m + k = m +_{\mathbb{N}} k \in \mathbb{N}$ . □

# Ordering Cardinals

## Definition

A set  $A$  is **dominated** by a set  $B$  (written  $A \preceq B$ ) if there is an injection from  $A$  to  $B$ .

## Examples

- ▶  $A \preceq A$ .
- ▶  $A \preceq B$  if  $A \subset B$ . (Consider the inclusion map  $\iota : A \hookrightarrow B$ .)
- ▶  $A \preceq B$  iff  $A$  is equinumerous to some subset of  $B$ . (Consider a bijection between  $A$  and  $f(A) \subset B$ .)
- ▶  $\mathbb{N} \preceq \mathbb{Z} \preceq \mathbb{Q} \preceq \mathbb{R} \preceq \mathbb{C}$ .
- ▶  $\mathbb{R} \approx (0, 1) \preceq [0, 1] \preceq \mathcal{P}(\mathbb{N}) \approx 2^{\mathbb{N}} \preceq \mathbb{R}$ .

# Ordering Cardinals

## Definition

We write  $\text{card } A \leq \text{card } B$  if  $A \preceq B$ .

**Claim:** This ordering is well-defined.

We need to verify that the definition is independent of the chosen representatives.

Suppose for sets  $A'$  and  $B'$  with  $\text{card } A = \text{card } A'$  and  $\text{card } B = \text{card } B'$ , then  $A \approx A'$  and  $B \approx B'$ . Now if  $A \preceq B$ , then there exist

- ▶  $\alpha : A' \rightarrow A$  bijective;
- ▶  $\beta : A \rightarrow B$  injective;
- ▶  $\gamma : B \rightarrow B'$  bijective.

Thus the overall composition  $\gamma \circ \beta \circ \alpha : A' \rightarrow B'$  is injective, hence  $A' \preceq B'$ .

# Ordering Cardinals

## Definition

We write  $\text{card } A < \text{card } B$  if  $A \preceq B$  and  $A \not\approx B$ .

## Examples

- ▶ If  $A \subset B$ , then  $\text{card } A \leq \text{card } B$ .
- ▶ For all cardinal  $\kappa$ ,  $0 \leq \kappa$ .
- ▶ For all finite cardinal  $n$ ,  $n < \aleph_0$ .
- ▶ If  $m$  and  $n$  are finite cardinals, then  $m \subset n \Rightarrow m \leq n$ .
- ▶ For all cardinal  $\kappa$ ,  $\kappa < 2^\kappa$ . (There is no largest cardinal number.)

# Countable Sets

## Definition

A set  $A$  is **countable** if  $A \preceq \mathbb{N}$ , i.e.,  $\text{card } A \leq \aleph_0$ . Otherwise, it is called **uncountable**.

## Examples

- ▶  $\mathbb{N}$ ,  $\mathbb{Z}$ , and  $\mathbb{Q}$  are countable;  $\mathbb{R}$  is uncountable.
- ▶ A subset of a countable set is countable.
- ▶ The Cartesian product of two countable sets is countable.
- ▶ A countable union of countable sets is countable.
- ▶ If  $X$  is countable and  $f : X \rightarrow Y$  is onto, then  $Y$  is countable.
- ▶ For all infinite set  $A$ ,  $\mathcal{P}(A)$  is uncountable.

## Cantor-Schröder-Bernstein Theorem

Q: Does the ordering on the cardinals induce a “partial ordering”?

For sets  $A$ ,  $B$ , and  $C$ ,

- ▶ reflexivity:  $\text{card } A \leq \text{card } A$ , i.e.,  $A \preceq A$ .
- ▶ transitivity:  $(\text{card } A \leq \text{card } B) \wedge (\text{card } B \leq \text{card } C) \Rightarrow \text{card } A \leq \text{card } C$ ,  
i.e.,  
$$(A \preceq B) \wedge (B \preceq C) \Rightarrow A \preceq C.$$
- ▶ antisymmetry:  
$$(\text{card } A \leq \text{card } B) \wedge (\text{card } B \leq \text{card } A) \Rightarrow ? \text{ card } A = \text{card } B$$
, i.e.,  
$$(A \preceq B) \wedge (B \preceq A) \Rightarrow ? A \approx B.$$

A: Yes.

### Theorem (Cantor-Schröder-Bernstein)

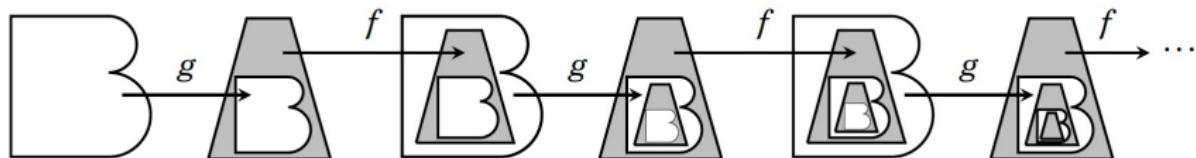
$(\text{card } A \leq \text{card } B) \wedge (\text{card } B \leq \text{card } A) \Rightarrow \text{card } A = \text{card } B$ , i.e.,  
 $(A \preceq B) \wedge (B \preceq A) \Rightarrow A \approx B.$

## Cantor-Schröder-Bernstein Theorem

Let  $f : A \rightarrow B$  and  $g : B \rightarrow A$  be injective.

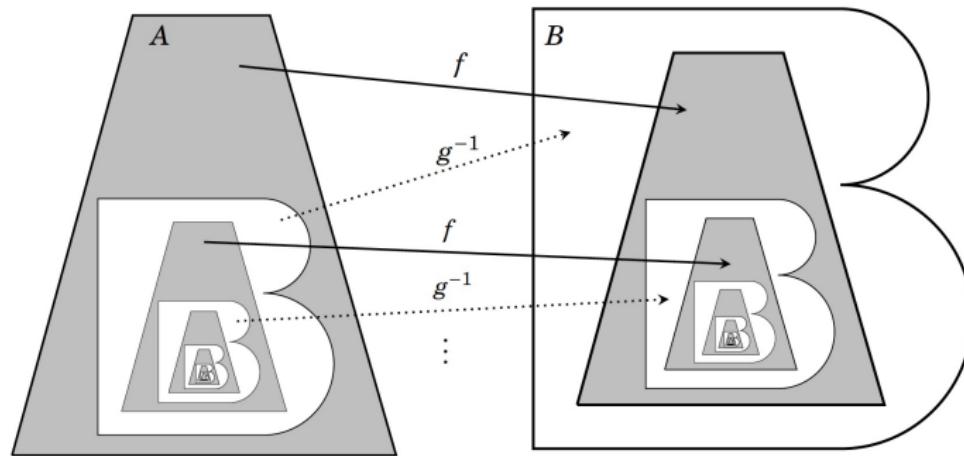


Alternating  $f$  and  $g$ , we get a chain of injections.



# Cantor-Schröder-Bernstein Theorem

Iterate to get a bijection  $h : A \rightarrow B$ .



$$h(x) := \begin{cases} f(x), & x \in \bigcup_{k \in \mathbb{N}} (g \circ f)^k(A - g(B)) \\ g^{-1}(x), & \text{otherwise} \end{cases}$$

## Part II

Basic Number Theory  
and Basic Group Theory

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Divisibility

## Definition

Let  $n, d \in \mathbb{Z}$  with  $d \neq 0$ , we say that  $d$  divides  $n$ , denoted by  $d | n$ , if  $n = dk$ , for some  $k \in \mathbb{Z}$ , i.e.,

$$d | n \Leftrightarrow (\exists k \in \mathbb{Z})(n = dk)$$

By convention,  $0 | n$  only if  $n = 0$ .

The following expressions are equivalent

- ▶  $d$  divides  $n$ .
- ▶  $n$  is divisible by  $d$ .
- ▶  $n$  is a multiple of  $d$ .
- ▶  $d$  is a divisor of  $n$ .
- ▶  $d$  is a factor of  $n$ .

# Divisibility

## Non-divisibility

If  $d$  does not divide  $n$ , we write  $d \nmid n$ . Note that

$$d \nmid n \Leftrightarrow \frac{n}{d} \notin \mathbb{Z}$$

## Examples

- ▶  $n \mid 0$  for all  $n \in \mathbb{Z}$ . ( $0 \equiv \top$ )
- ▶  $1 \mid n$  for all  $n \in \mathbb{Z}$ . ( $1 \equiv \perp$ )
- ▶ If  $d \in \mathbb{Z}$ , then  $d \mid 1 \Rightarrow d = \pm 1$ .
- ▶ If  $d \in \mathbb{N}$  and  $d \mid 2021$ , then  $d = ?$ .

# Prime Numbers

## Definition

A natural number  $p \in \mathbb{N}$  is a prime number (or simply, a prime) if  $p \geq 2$  and if  $p$  is divisible only by itself and 1.

## Remark

A natural number  $p \in \mathbb{N}$  is a prime number if it has exactly two distinct factors. The set of all primes is sometimes denoted by  $\mathbb{P}$ .

## Remark

1 is **NOT** a prime.

For convenience, e.g.,

- ▶ Unique factorization property.
- ▶ Largest power of  $p$  dividing  $n$ .
- ▶ Riemann Zeta function  $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in \mathbb{P}} \frac{1}{1 - p^{-s}}$ .

# Famous Prime Numbers

## Mersenne Primes

Mersenne Prime is a prime of the form  $2^n - 1$ .

- ▶  $2^2 - 1 = 3 \in \mathbb{P}$
- ▶  $2^3 - 1 = 7 \in \mathbb{P}$
- ▶  $2^5 - 1 = 31 \in \mathbb{P}$
- ▶  $2^7 - 1 = 127 \in \mathbb{P}$
- ▶ Necessary condition:  $2^n - 1 \in \mathbb{P} \Rightarrow n \in \mathbb{P}$ .
  - ▶  $2^{11} - 1 = 2047 = 23 \times 89$ .
- ▶ Not all primes are Mersenne.
  - ▶  $5 \in \mathbb{P}$  but is not Mersenne.

# Famous Prime Numbers

## Fermat Primes

$$F_n = 2^{2^n} + 1.$$

- ▶  $F_0 = 2^{2^0} + 1 = 3 \in \mathbb{P}.$
- ▶  $F_1 = 2^{2^1} + 1 = 5 \in \mathbb{P}.$
- ▶  $F_2 = 2^{2^2} + 1 = 17 \in \mathbb{P}.$
- ▶  $F_3 = 2^{2^3} + 1 = 257 \in \mathbb{P}.$
- ▶  $F_4 = 2^{2^4} + 1 = 65537 \in \mathbb{P}.$
- ▶  $F_5 = 2^{2^5} + 1 = 4274967297 = 641 \times 6700417. \text{ (Euler, 1732)}$

## Famous Conjectures

### Goldbach Conjecture (18th century), “1+1”

Can **every** even number greater than 4 be written as the sum of 2 primes?

- ▶  $4 = 2 + 2$
- ▶  $6 = 3 + 3$
- ▶  $8 = 3 + 5$
- ▶  $10 = 5 + 5$
- ▶  $20 = 7 + 13$
- ▶  $200 = 7 + 193$
- ▶  $2040 = 1019 + 1021$

### Jing-run Chen, 1966, “1+2”

All sufficiently large even numbers are the sum of a prime and the product of at most two primes

$$2n = p_1 + p_2 p_3$$

# Famous Conjectures

## Twin Prime Conjecture

Twin primes are a pair of primes which differ by 2:

- ▶ (3, 5); (5, 7); (11, 13); (17, 19); (29, 31); (41, 43); (59, 61); (71, 73); (107, 109); (2027, 2029); (1,000,037, 1,000,039);

Are there infinitely many such pairs?

Yitang Zhang: Bounded gaps between primes, 2014

It is proved that

$$\liminf_{n \rightarrow \infty} (p_{n+1} - p_n) < 7 \times 10^7,$$

where  $p_n$  is the  $n$ -th prime.

# Infinitude of Primes

## Theorem

*There are infinitely many primes.*

## Proof of Euclid.

For any **finite** set  $\{p_1, \dots, p_r\} \subset \mathbb{P}$ , consider the number  $n = p_1 p_2 \cdots p_r + 1$ . Note that  $p_i \nmid n$  for all  $i = 1, \dots, r$ , then

- ▶ either  $n$  is a prime,
- ▶ or  $n$  has a divisor  $p \notin \{p_1, \dots, p_r\}$ .

Either way a new prime is generated from the finite set, hence  $\{p_1, \dots, p_r\}$  cannot be the whole collection of **all** primes. □

## Example

- ▶  $\{2, 3, 7\} \subset \mathbb{P}$ ,  $2 \cdot 3 \cdot 7 + 1 = 43 \in \mathbb{P}$ ;
- ▶  $2 \cdot 3 \cdot 7 \cdot 43 + 1 = 1807 = 13 \times 139$ .

## Euclid's Proof of the Infinity of the Number of Primes

Note that the proof does **not** state that  $n = p_1 \cdot p_2 \cdots p_r + 1$  must be a prime. However, it is interesting to note that it often seems to be the case:

- ▶  $2 + 1 = 3,$
- ▶  $2 \cdot 3 + 1 = 7,$
- ▶  $2 \cdot 3 \cdot 5 + 1 = 31,$
- ▶  $2 \cdot 3 \cdot 5 \cdot 7 + 1 = 211,$
- ▶  $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 + 1 = 2311,$
- ▶  $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 59 \cdot 509,$  etc.

It is not known whether there are infinitely many  $r$  for which  $n$  is prime.

## Infinitely Many Twin Primes?

- ▶ Euclid number:  $E_n = p_1 \cdots p_n + 1$
- ▶ Euclid number of the second kind (also called Kummer number):  
 $E_n = p_1 \cdots p_n - 1.$

# Variant of Euclid's Theorem

## Lemma

Given  $p \in \mathbb{P}$ ,  $k \in \mathbb{N}$ , if  $p \mid k^2 + 1$ , then  $p = 2$  or  $p$  is of the form  $4m + 1$ .

e.g.,

- ▶  $2^2 + 1 = 5$
- ▶  $3^2 + 1 = 2 \times 5$
- ▶  $4^2 + 1 = 17$
- ▶  $5^2 + 1 = 2 \times 13$

We'll prove this later.

## Example

There are infinitely many primes of the form  $4m + 1$ ,  $m \in \mathbb{N}$ . Given  $\{p_1 = 2, p_2, \dots, p_m\} \subset \mathbb{P}$ , take  $n = (p_1 \cdots p_m)^2 + 1$ , then

- ▶ We have a new prime  $p_{m+1} \mid n$ ,
- ▶ Since  $p_{m+1} \in \mathbb{P}$ , and  $p_{m+1} \mid n$ , thus  $p_{m+1} = 2$  or  $p_{m+1}$  is of the form  $4k + 1$ .

# Dirichlet's Theorem

## Theorem

*There are infinitely many primes of the form  $an + b$ , for  $n \in \mathbb{N}$ , and  $a, b$  coprime.*

cf., Stein, Fourier Analysis.

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Greatest Common Divisor

## Definition

Let  $a, b \in \mathbb{Z}$ , not both zero. The **greatest common divisor** of  $a$  and  $b$ , denoted by  $\gcd(a, b)$  or simply  $(a, b)$ , is the positive integer  $d$  satisfying:

- $d$  is a **common divisor** of  $a$  and  $b$ , i.e.,

$$d \mid a \quad \text{and} \quad d \mid b$$

- If  $c$  also divides  $a$  and  $b$ , then  $c \leq d$  (or  $c \mid d$ ). In other words,

$$\forall c \in \mathbb{N}, \text{ if } c \mid a \text{ and } c \mid b, \text{ then } c \leq d.$$

## Example

- $\gcd(72, 63) = 9$
- $\gcd(10^{12}, 6^{18}) = \gcd(2^{12} \cdot 5^{12}, 2^{18} \cdot 3^{18}) = 2^{12}$
- $\gcd(5, 0) = 5$
- $\gcd(0, 0) = 0$

# Calculate $\text{gcd}(m, n)$ , Algorithm 1

Algorithm 1 (assuming  $m \leq n$ )

---

**Input:**  $m, n \in \mathbb{N} \setminus \{0\}$ ,  $m \leq n$

**Output:** Greatest common divisor of  $m$  and  $n$

```
1 Function gcd(m, n):
2     d ← m;
3     while d ∤ n and d ∤ m do
4         |   d ← d - 1
5     end
6     return d
7 end
```

---

Advantage

- ▶ Simple
- ▶ Terminates in finite steps (try  $d = 1$ )
- ▶ Yields the correct answer (which exists)

Disadvantage

- ▶ Slow

Calculate  $\gcd(m, n)$ ,  $m, n \in \mathbb{N} \setminus \{0\}$

### Algorithm 2 (Factorization)

Factor  $m$  and  $n$  as

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$$
$$n = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$$

with  $p_1, \dots, p_k \in \mathbb{P}$ , and  $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k \in \mathbb{N}$ . Then

$$\gcd(m, n) = p_1^{\min\{\alpha_1, \beta_1\}} p_2^{\min\{\alpha_2, \beta_2\}} \cdots p_k^{\min\{\alpha_k, \beta_k\}}$$

Disadvantage

- ▶ Factorization is hard (until the foreseeable future).

Calculate  $\text{gcd}(m, n)$ ,  $m, n \in \mathbb{N} \setminus \{0\}$

Algorithm 3 (Euclidean algorithm, assuming  $m \leq n$ )

---

**Input:**  $m, n \in \mathbb{N} \setminus \{0\}$ ,  $m \leq n$

**Output:** Greatest common divisor of  $m$  and  $n$

```
1 Function gcd(m, n):
2     if n mod m = 0 then
3         return m
4     else
5         return gcd(n mod m, m)
6     end
7 end
```

---

**FACTS:** For  $m, n \in \mathbb{N} \setminus \{0\}$

- If  $m | n$ , then

$$\text{gcd}(n, m) = m.$$

- If  $n = qm + r$  with  $q \geq 0$  and  $0 \leq r < m$ , then

$$\text{gcd}(n, m) = \text{gcd}(m, r).$$

## Proof of Facts

### FACT 1

For  $m, n \in \mathbb{N} \setminus \{0\}$ , if  $m \mid n$ , then  $\gcd(n, m) = m$ .

Proof.

- ▶ Since  $\gcd(n, m) \mid m$ , then  $\gcd(n, m) \leq m$ .
- ▶ Since  $m$  is a common divisor, then  $m \leq \gcd(n, m)$ .

Hence  $\gcd(n, m) = m$

□

## Proof of Facts

### FACT 2

For  $m, n \in \mathbb{N} \setminus \{0\}$ , if  $n = qm + r$  with  $q \geq 0$  and  $0 \leq r < m$ , then  $\gcd(n, m) = \gcd(m, r)$

Proof.

- $\gcd(n, m) \leq \gcd(m, r)$ . Let  $c$  be any common divisor of  $n$  and  $m$ , i.e.,  $c | m$  and  $c | n$ , hence there exist  $k, \ell \in \mathbb{Z}$  such that  $n = ck$  and  $m = c\ell$ . Then

$$\begin{aligned}n &= ck = c\ell q + r \\ \Rightarrow r &= ck - c\ell q = c(k - \ell q) \\ \Rightarrow c &\mid r\end{aligned}$$

Take  $c = \gcd(n, m)$ , hence  $\gcd(n, m)$  divides both  $m$  and  $r$ , so  $\gcd(n, m) \leq \gcd(m, r)$ .

## Proof of Facts

### Proof (Cont.)

- ▶  $\gcd(m, r) \leq \gcd(n, m)$ . Similarly let  $c$  be any common divisor of  $m$  and  $r$ , i.e.,  $c \mid m$  and  $c \mid r$ , hence there exist  $x, y \in \mathbb{Z}$  such that  $m = cx$  and  $r = cy$ . Then

$$n = cxq + cy = (xq + y)c$$

hence  $c \mid n$ . Take  $c = \gcd(m, r)$ , hence  $c = \gcd(m, r)$  divides both  $n$  and  $m$ , so  $\gcd(m, r) \leq \gcd(n, m)$ .

Combine the two results.



# Division Algorithm

## Theorem ((Long) Division Algorithm)

Given  $m, n \in \mathbb{N} \setminus \{0\}$ , there exist unique integers  $q$  and  $r$  with  $q \geq 0$  and  $0 \leq r < m$  so that  $n = qm + r$ .

### Proof.

Existence by induction on  $n$ . Let

$$S = \{n \in \mathbb{N} \mid (\forall m > 0)(\exists q, r \text{ with } q \geq 0 \text{ and } 0 \leq r < m)(n = qm + r)\}$$

- ▶  $1 \in S$ . ( $1 = 1 \cdot 1 + 0$  for  $m = 1$ , and  $1 = 0m + 1$  for  $m > 1$ )
- ▶ Let  $k \in S$ . Then for  $m > 0$ , there exist  $q, r$  such that  $k = qm + r$ .  
Now
  - ▶  $k + 1 = qm + (r + 1)$ , if  $r + 1 < m$ ;
  - ▶  $k + 1 = (q + 1)m + 0$ , if  $r + 1 = m$ .

Thus  $k + 1 \in S$ .

# Division Algorithm

## Proof (Cont.)

Uniqueness.

Suppose  $n = q_1m + r_1 = q_2m + r_2$ , then  $r_1 - r_2 = (q_2 - q_1)m$ , thus if  $q_1 \neq q_2$ , then  $m \mid (r_1 - r_2)$ .

But  $|r_1 - r_2| < m$ , hence  $r_1 - r_2 = 0$ .

But then  $q_1 = q_2$ , contradiction.



## Remark

Note that  $(q_1 - q_2)m + (r_1 - r_2) = 0$  implies  $q_1 - q_2 = 0$  and  $r_1 - r_2 = 0$ , which is basically applying the long division algorithm to 0.

# Euclidean Algorithm

## Euclidean Algorithm

Given positive integers  $n$  and  $m$ , we can repeat the division algorithm to obtain a series of equations

$$n = mq_1 + r_1, \quad 0 < r_1 < m$$

$$m = r_1q_2 + r_2, \quad 0 < r_2 < r_1$$

$$r_1 = r_2q_3 + r_3, \quad 0 < r_3 < r_2$$

$$\vdots$$

$$r_{j-2} = r_{j-1}q_j + r_j, \quad 0 < r_j < r_{j-1}$$

$$r_{j-1} = r_jq_{j+1}$$

Then  $\gcd(n, m) = r_j$ .

Remark: By induction and the two facts, the Euclidean algorithm terminates within finite number of steps and produce the correct answer.

# Euclidean Algorithm

## Example

$n = 42823$  and  $m = 6409$

$$\begin{aligned} 42823 &= 6409 \times 6 + 4369 & (42823, 6409) \\ 6409 &= 4369 \times 1 + 2040 & = (6409, 4369) \\ 4369 &= 2040 \times 2 + 289 & = (4369, 2040) \\ 2040 &= 289 \times 7 + 17 & = (2040, 289) \\ 289 &= 17 \times 17 + 0 & = (289, 17) = 17 \end{aligned}$$

## Remark

The Euclidean algorithm provides a solution to the Diophantine equation

$$mx + ny = \gcd(m, n)$$

by back-tracking.

# Euclidean Algorithm

## Example (Cont.)

Consider the Diophantine equation

$$42823x + 6409y = 17 = \gcd(42823, 6409).$$

### Euclidean Algorithm

$$42823 = 6409 \times 6 + 4369$$

$$6409 = 4369 \times 1 + 2040$$

$$4369 = 2040 \times 2 + 289$$

$$2040 = 289 \times 7 + 17$$

$$289 = 17 \times 17 + 0$$

### Back-Tracking

$$17 = 2040 - 289 \times 7$$

$$= 2040 - (4369 - 2040 \times 2) \times 7$$

$$= 2040 \times 15 - 4369 \times 7$$

$$= (6409 - 4369) \times 15 - 4369 \times 7$$

$$= 6409 \times 15 - 4369 \times 22$$

$$= 6409 \times 15 - (42823 - 6409 \times 6) \times 22$$

$$= 6409 \times (15 + 6 \times 22) - 42823 \times 22$$

$$= 6409 \times 147 - 42823 \times 22$$

Let's take  $x = -22$  and  $y = 147$ .

# Euclidean Algorithm

## Example (Cont.)

$$\begin{aligned}\frac{42823}{6409} &= 6 + \frac{6369}{6409} = 6 + \frac{1}{1 + \frac{2040}{4369}} = 6 + \frac{1}{1 + \frac{1}{2 + \frac{289}{2040}}} \\&= 6 + \frac{1}{1 + \frac{1}{2 + \frac{1}{7 + \frac{17}{289}}}} = 6 + \frac{1}{1 + \frac{1}{2 + \frac{1}{7 + \frac{1}{17}}}}\end{aligned}$$

# Euclidean Algorithm

## Example (Cont.)

$$6 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{7 + \cfrac{1}{17}}}} = 6 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{7}}} = 6 + \cfrac{1}{1 + \cfrac{7}{15}} = 6 + \cfrac{15}{22} = \cfrac{147}{22}$$

Now

$$\frac{42823}{6409} = \frac{2519}{377} \leqslant \frac{147}{22}?$$

Of course

$$377 \times 147 - 2519 \times 22 = 1$$

i.e.,

$$6409 \times 147 - 42823 \times 22 = 17$$

Calculate  $\text{gcd}(m, n)$ ,  $m, n \in \mathbb{N} \setminus \{0\}$

#### Algorithm 4 (Binary Euclidean/GCD Algorithm)

---

**Input:**  $m, n \in \mathbb{N} \setminus \{0\}$

**Output:** Greatest common divisor of  $m$  and  $n$

1 **Function**  $\text{gcd}(m, n)$ :

```
2   if  $n = m$  then return  $m$ ;  
3   else if  $2 | m$  and  $2 | n$  then return  $2\text{gcd}(m/2, n/2)$ ;  
4   else if  $2 | m$  then return  $\text{gcd}(m/2, n)$ ;  
5   else if  $2 | n$  then return  $\text{gcd}(m, n/2)$ ;  
6   else if  $m > n$  then return  $\text{gcd}(m - n, n)$ ;  
7   else return  $\text{gcd}(m, n - m)$ ;
```

8 **end**

---

#### FACTS:

- ▶ If  $2 | m$  and  $2 | n$ , then  $\text{gcd}(m, n) = 2\text{gcd}(m/2, n/2)$ .
- ▶ If  $2 | m$  and  $2 \nmid n$ , then  $\text{gcd}(m, n) = \text{gcd}(m/2, n)$

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Groups

## Definition

A group is a pair  $(G, \cdot)$ , where  $G$  is a set, and  $\cdot : G \times G \rightarrow G$ ,  $(g, h) \mapsto g \cdot h = gh$ , is a law of composition (aka group law) that has the following properties:

- ▶ The law of composition is associative:  $(ab)c = a(bc)$  for all  $a, b, c \in G$ .
- ▶  $G$  contains an identity element  $1$ , such that  $1a = a1 = a$  for all  $a \in G$ .
- ▶ Every element  $a \in G$  has an inverse, an element  $b$  such that  $ab = ba = 1$ .

An **abelian** group is a group whose law of composition is commutative.

## Example

- ▶  $(\mathbb{Z}, +)$
- ▶  $(\mathbb{R} \setminus \{0\}, \cdot)$
- ▶ The set of  $n \times n$  invertible matrices  $GL_n(\mathbb{R})$  or  $GL_n(\mathbb{C})$ .

# Elementary Properties of Groups

## Theorem

Given a group  $G$ ,  $a, b, c \in G$ , then

- ▶ there exists a unique identity element.
- ▶  $ba = ca \Rightarrow b = c$  and  $ab = ac \Rightarrow b = c$ .
- ▶ For all  $a \in G$ , there exists a unique element  $b \in G$  such that  $ab = ba = 1$ .
- ▶  $(ab)^{-1} = b^{-1}a^{-1}$ .

# Subgroup

## Definition

A subset  $H$  of a group  $G$  is a subgroup if it has the following properties:

- ▶ Closure: If  $a, b \in H$ , then  $ab \in H$ .
- ▶ Identity:  $1 \in H$ .
- ▶ Inverses: If  $a \in H$ , then  $a^{-1} \in H$ .

## Subgroups of the Additive Group $(\mathbb{Z}, +)$

A subset  $S$  of  $(\mathbb{Z}, +)$  is a subgroup if

- ▶ Closure: If  $a, b \in S$ , then  $a + b \in S$ .
- ▶ Identity:  $0 \in S$ .
- ▶ Inverses: If  $a \in S$ , then  $-a \in S$ .

For  $a \in \mathbb{Z}$ , a subgroup of  $(\mathbb{Z}, +)$  is given by

$$a\mathbb{Z} = \{n \in \mathbb{Z} \mid n = ka \text{ for some } k \in \mathbb{Z}\}$$

## Subgroup of $(\mathbb{Z}, +)$

### Subgroups of the Additive Group $(\mathbb{Z}, +)$

A subset  $S$  of  $(\mathbb{Z}, +)$  is a subgroup if

- ▶ Closure:  $a, b \in S \Rightarrow a + b \in S$ .
- ▶ Identity:  $0 \in S$ .
- ▶ Inverses:  $a \in S \Rightarrow -a \in S$ .

For  $a \in \mathbb{Z}$ , a subgroup of  $(\mathbb{Z}, +)$  is given by integers divisible by  $a$  as,

$$a\mathbb{Z} = \{n \in \mathbb{Z} \mid n = ka \text{ for some } k \in \mathbb{Z}\}$$

### Remark

We write  $H \leq G$  if  $H$  is subgroup of  $G$ .

### Example

- ▶  $a = 0$  yields the trivial group  $(\{0\}, +)$ .
- ▶  $a = 1$  yields the whole of  $(\mathbb{Z}, +)$ .

# Subgroup of $(\mathbb{Z}, +)$

## Theorem

Let  $S$  be a *subgroup* of the additive group  $(\mathbb{Z}, +)$ , then

- ▶ either  $S$  is the trivial subgroup  $(\{0\}, +)$ ,
- ▶ or it has the form  $a\mathbb{Z}$ , where  $a$  is the *smallest positive integer* in  $S$ .

## Proof.

Let  $S$  be a subgroup of  $(\mathbb{Z}, +)$ , then  $0 \in S$ . If  $S = \{0\}$ , then we are done. Otherwise,  $\exists n \in \mathbb{Z} \cap S - \{0\}$ , then  $\pm n \in S$  by subgroup property of  $S$ , hence either  $n$  or  $-n$  is a positive integer.

Next we show  $S = a\mathbb{Z}$ , where  $a$  is the smallest positive integer of  $S$ .

- ▶  $a\mathbb{Z} \subset S$ . Let  $z \in a\mathbb{Z}$ , then  $z = ka$  for some  $k \in \mathbb{Z}$ . Suppose  $z > 0$ , since  $a \in S$ , then  $ka \in S$  for  $k \in \mathbb{N}$  by induction and closure. Also  $-ka \in S$  by the inverse property. Similar goes for  $z < 0$ . If  $z = 0 \in a\mathbb{Z}$ , then also  $z = 0 \in S$ .

## Subgroup of $(\mathbb{Z}, +)$

### Proof (Cont.)

- $a\mathbb{Z} \supset S$ . Take  $n \in S$ , then  $n = qa + r$  for some  $q \in \mathbb{Z}$  and  $0 \leq r < a$ . Now since  $qa \in \mathbb{Z}a \subset S$ , and  $n \in S$ , then  $r = n - qa \in S$ . But  $a$  is the smallest positive integer in  $S$ , hence  $r = 0$ . Therefore  $n = qa$  for some  $q \in \mathbb{Z}$ , thus  $n \in a\mathbb{Z}$ .

Therefore  $a\mathbb{Z} = S$ .



### Definition

Given  $a, b \in \mathbb{Z}$ , then the subgroup  $S$  generated by  $a$  and  $b$ , denoted by

$$S = \mathbb{Z}a + \mathbb{Z}b = \{n \in \mathbb{Z} \mid n = ra + sb \text{ for some integers } r, s\}$$

It is also the **smallest subgroup** that contains both  $a$  and  $b$ .

### Remark

Since  $S \subset \mathbb{Z}$  is a subgroup, then  $S = d\mathbb{Z}$  for some  $d \in \mathbb{Z}$ .

# Subgroup of $(\mathbb{Z}, +)$

## Theorem

Let  $a, b \in \mathbb{Z}$ , not both zero, and let  $d$  be the positive integer that generates the subgroup  $S = \mathbb{Z}a + \mathbb{Z}b$ , i.e.,  $\mathbb{Z}d = \mathbb{Z}a + \mathbb{Z}b$ . Then

1.  $d \mid a$  and  $d \mid b$ .
2. For  $e \in \mathbb{Z}$ , if  $e \mid a$  and  $e \mid b$ , then  $e \mid d$ .
3. There are integers  $r$  and  $s$  such that  $d = ra + sb$ .

Note that  $d = \gcd(a, b)$ .

## Proof.

1.  $a \in \mathbb{Z}d$  and  $b \in \mathbb{Z}d$ .
3.  $d \in \mathbb{Z}a + \mathbb{Z}b$ .
2. Let  $d = ra + sb$ , then  $e \mid a$  and  $e \mid b$  implies  $e \mid (ra + sb)$ , therefore  $e \mid d$ . □

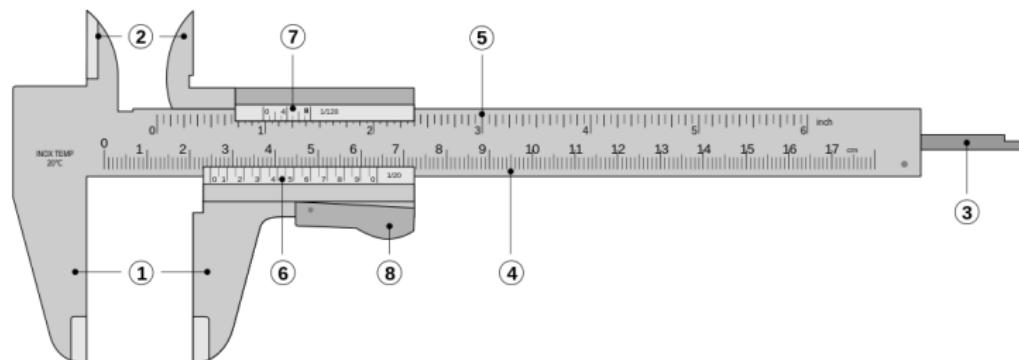
# Subgroup of $(\mathbb{Z}, +)$

## Corollary (Bézout Identity)

Given  $a, b \in \mathbb{Z}$  such that  $\gcd(a, b) = 1$ , i.e.,  $a$  and  $b$  relatively prime or coprime **iff** there exist  $r, s \in \mathbb{Z}$  such that  $ra + sb = 1$ .

## Remark

The proof is just by letting  $d = 1$ . In this case  $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$ .



Vernier Caliper

## Subgroup of $(\mathbb{Z}, +)$

### Corollary

Let  $p$  be prime, and  $a, b \in \mathbb{Z}$ . If  $p \mid ab$ , then  $p \mid a$  or  $p \mid b$ .

### Proof.

Suppose  $p \nmid a$ , then  $\gcd(a, p) = 1$ . Therefore  $\exists r, s \in \mathbb{Z}$  such that  $ra + sp = 1$ . Hence  $rab + spb = b$ . Note that  $p \mid rab$  and  $p \mid spb$ , thus  $p \mid b$ . □

### Remark

By induction, given  $p \in \mathbb{P}$ , and  $a_1, \dots, a_n \in \mathbb{Z}$ , if  $p \mid a_1 a_2 \cdots a_n$ , then  $p \mid a_i$  for some factor  $a_i$  of the product.

### Corollary

If  $c \mid ab$  and  $\gcd(b, c) = 1$ , then  $c \mid a$ .

# Fundamental Theorem of Arithmetic

## Theorem

*Every positive integer can be written uniquely (up to order) as a product of primes (with possibly only one factor).*

## Remark

Convention: 1 is the product of empty set of primes

## Proof.

- ▶ Existence: If  $n > 1$ , then either  $n$  is prime, or can be factored into, say  $n = p \cdot (n/p)$  for some prime  $p$ , continue by induction.
- ▶ Uniqueness: Suppose  $n = p_1 \cdots p_r = q_1 \cdots q_s$ , with  $p_i, q_i$  primes. Then  $p_1 \mid (q_1 \cdots q_s)$ , thus  $p_1 = q_i$  for some  $i$ . Cancel  $p_1$  and  $q_i$  and continue by induction. □

# Fundamental Theorem of Arithmetic

## Other versions of Fundamental Theorem of Arithmetic

- ▶ Integers. (allow negative primes and  $-1$ ).
- ▶ Polynomials over a field. (Factor into irreducible polynomials)

## Examples of Non-uniqueness

- ▶ Positive integers of the form  $4n + 1$ . Consider 1, 5, 9, 13, 17, 21,  $25 (= 5^2)$ , 29, 33, 37, 41, 45 ( $= 5 \cdot 9$ ), 49, ...

$$21 \cdot 21 = 9 \cdot 49.$$

- ▶ Consider numbers of the form  $m + n\sqrt{-5}$ ,  $m, n \in \mathbb{Z}$ , then

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

# Fundamental Theorem of Arithmetic

## Riemann Zeta Function

Euler discovered that (equivalent to fundamental theorem of arithmetic)

$$\begin{aligned}\zeta(s) &= \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n^s} \\ &= \frac{1}{1 - 2^{-s}} \cdot \frac{1}{1 - 3^{-s}} \cdot \frac{1}{1 - 5^{-s}} \cdot \frac{1}{1 - 7^{-s}} \cdots \\ &= \prod_{p \in \mathbb{P}} \frac{1}{1 - p^{-s}}\end{aligned}$$

Let  $s = 1$ , then by divergence of the harmonic series, there are infinitely many primes.

## Theorem (Dirichlet)

If  $u, v \in \mathbb{Z}$  are chosen at random, the probability that  $\gcd(u, v) = 1$  is  $\zeta(2)^{-1} = 6/\pi^2 \approx 0.60793$ .

# Fundamental Theorem of Arithmetic

## Example

To illustrate  $\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in \mathbb{P}} \frac{1}{1 - p^{-s}}$ , consider

$$\begin{aligned} & \frac{1}{1 - 2^{-s}} \cdot \frac{1}{1 - 3^{-s}} \cdot \frac{1}{1 - 5^{-s}} \cdot \frac{1}{1 - 7^{-s}} \cdots \\ &= (1 + 2^{-s} + (2^{-s})^2 + (2^{-s})^3 + \cdots) \\ & \quad (1 + 3^{-s} + (3^{-s})^2 + (3^{-s})^3 + \cdots) \\ & \quad (1 + 5^{-s} + (5^{-s})^2 + (5^{-s})^3 + \cdots) \\ & \quad (\cdots) \end{aligned}$$

Note that, for example,

$$(2^{-s})^3 \cdot (3^{-s}) \cdot (5^{-s})^2 = \frac{1}{(2^3 \cdot 3 \cdot 5^2)^s} = \frac{1}{600^s}$$

## Fundamental Theorem of Arithmetic

Also by Euler,  $p \in \mathbb{P}$ ,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{p} \approx \log \log p \text{ "}\approx 3\text{"}$$

| $n$       | $\log \log n$ |
|-----------|---------------|
| $10^3$    | 1.9           |
| $10^9$    | 2.6           |
| $10^9$    | 3.0           |
| $10^{12}$ | 3.3           |
| $10^{15}$ | 3.5           |

# Least Common Multiple

## Theorem

Let  $a, b \in \mathbb{Z} \setminus \{0\}$ , and let  $m = \text{lcm}(a, b)$  be their **least common multiple** — the positive integer that generates the subgroup  $S = a\mathbb{Z} \cap b\mathbb{Z}$ , i.e.,  $m\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$ . Then

- ▶  $a \mid m$  and  $b \mid m$ .
- ▶  $a, b \mid n$  for some  $n \in \mathbb{Z}$ , then  $m \mid n$ .

## Proof.

Note that  $a\mathbb{Z} \cap b\mathbb{Z}$  is a nontrivial subgroup of  $(\mathbb{Z}, +)$ . □

## Remark

Again by induction, if  $n$  is any common multiple of  $a_1, \dots, a_n \in \mathbb{Z}$ , then  $\text{lcm}(a_1, \dots, a_n) \mid n$ .

# Greatest Common Divisor and Least Common Multiple

## Corollary

Given  $a, b \in \mathbb{N} \setminus \{0\}$ , let  $d = \gcd(a, b)$  and  $m = \text{lcm}(a, b)$ , then  $ab = dm$ .

## Proof.

- ▶ Since  $b/d \in \mathbb{Z}$ , then  $a \mid (ab/d)$ . Similarly  $b \mid (ab/d)$ . Therefore by definition of lcm,  $m \mid (ab/d)$ , hence  $dm \mid ab$ .
- ▶ Let  $d = ra + sb$ , then  $dm = ram + sbm$ . Since  $ab \mid ram$  and  $ab \mid sbm$ , then  $ab \mid dm$ .

Therefore  $ab = dm$ .



# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
- 4. Cyclic Groups and Symmetric Groups**
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Cyclic Groups

## Definition

A group is *cyclic* if it can be *generated by* a single element.

## Example

In multiplication notation, The cyclic subgroup  $H \leq G$  generated by  $x \in G$  is the set of all elements that are powers of  $x$ ,

$$\begin{aligned} H &:= \{\dots, x^{-2}, x^{-1}, 1, x, x^2, \dots\} \\ &= \{x^m \mid m \in \mathbb{Z}\} \end{aligned}$$

This is the *smallest subgroup* of  $G$  containing  $x$ , (often) denoted by  $\langle x \rangle$ . If there exists a smallest  $m \in \mathbb{N}$  such that  $x^m = 1$ , we say  $m$  is the *order* of  $x$ , denoted by  $m = |x|$ . Similarly, the *order* of a group  $G$ , denoted  $|G|$ , is given by the number of elements of  $G$ .

## Remark

The powers  $x^n$  may represent distinct elements, or not. For example, given  $-1 \in \mathbb{R}^\times$ , then  $\{(-1)^m \mid m \in \mathbb{Z}\} = \{\pm 1\}$ .

# Cyclic Groups

## Theorem

Let  $\langle x \rangle$  be the cyclic subgroup of a group  $G$  generated by an element  $x$ , and let  $S := \{k \in \mathbb{Z} \mid x^k = 1\}$ , then

1. The set  $S$  is a subgroup of the additive group  $(\mathbb{Z}, +)$ .
2. For  $r, s \in \mathbb{Z}$ ,  $x^r = x^s$  iff  $x^{r-s} = 1$ , i.e.,  $r - s \in S$ .
3. Suppose  $S \neq \{0\}$ , then  $S = n\mathbb{Z}$  for some  $n \in \mathbb{N} \setminus \{0\}$ . The powers  $1, x, x^2, \dots, x^{n-1}$  are distinct elements of the subgroup  $\langle x \rangle$ , and  $|\langle x \rangle| = n$ , i.e., the order of  $\langle x \rangle$  is  $n$ .

## Proof.

1. We check the properties of  $S$

- ▶ Let  $k, \ell \in S$ , then  $x^k = x^\ell = 1$ , hence  $x^{k+\ell} = x^k x^\ell = 1$ , therefore  $k + \ell \in S$ .
- ▶  $x^0 = 1$ , hence  $0 \in S$ .
- ▶ If  $k \in S$ , i.e.,  $x^k = 1$ , then  $x^{-k} = (x^k)^{-1} = 1$ , hence  $-k \in S$ .

# Cyclic Groups

## Proof (Cont.)

2. By straightforward calculation (cancellation law).
3. If  $S \neq \{0\}$ , then since  $S$  is a subgroup of  $(\mathbb{Z}, +)$ , then  $S = n\mathbb{Z}$  for some smallest positive integer  $n \in S$ . For any  $k \in \mathbb{Z}$ ,  $k = qn + r$  for some  $q \in \mathbb{Z}$  and  $0 \leq r < n$ . Thus  $x^k = x^{qn+r} = x^{qn}x^r = x^r$ . Note that  $1, x, x^2, \dots, x^{n-1}$  are distinct since  $n$  is the smallest power such that  $x^n = 1$ .

□

## Remark

- ▶ If  $|x| = \infty$ , then  $x^r = x^s$  iff  $r = s$  (since  $r - s \in \{0\}$ ).
- ▶ If  $|x| < \infty$ , say,  $|x| = n \in \mathbb{N}$ , then  $x^r = x^s$  iff  $n \mid r - s$ , i.e.,  $r \equiv s \pmod{n}$  (since  $r - s \in n\mathbb{Z}$ ).
- ▶  $|x| = |\langle x \rangle|$ .
- ▶ If  $|x| = n$  and  $x^k = 1$ , then  $n \mid k$ .

# Cyclic Groups

## Examples

- ▶  $(\mathbb{Z}, +)$
- ▶  $\mathbb{Z}/8\mathbb{Z} = \langle 1 \rangle = \langle 3 \rangle = \langle 5 \rangle = \langle 7 \rangle$ .
- ▶  $\langle r \mid r^n = 1 \rangle$ , where  $r$  represents counterclockwise rotation of  $2\pi/n$ .
- ▶  $\{e^{2\pi ik/n} \mid k \in \mathbb{Z}\} = \langle e^{2\pi i/n} \rangle$ ,  $n \in \mathbb{N} \setminus \{0\}$ .
- ▶  $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\equiv$ , where  $a \equiv b$  if  $n \mid a - b$ , i.e.,  $a - b \in n\mathbb{Z}$ , for given  $n \in \mathbb{N} \setminus \{0\}$ .

## Nonexamples

- ▶ The Klein four group  $V = \left\{ \begin{bmatrix} \pm 1 & \\ & \pm 1 \end{bmatrix} \right\}$ .
- ▶ The quaternion group  $H = \{\pm \mathbf{1}, \pm \mathbf{i}, \pm \mathbf{j}, \pm \mathbf{k}\}$ , where

$$\mathbf{1} = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, \quad \mathbf{i} = \begin{bmatrix} i & \\ & -i \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} & i \\ -i & \end{bmatrix}$$

# Cyclic Group

## Theorem

Let  $n, k \in \mathbb{N} \setminus \{0\}$ . Given group  $G$  and  $x \in G$  with  $|x| = n \in \mathbb{N} \setminus \{0\}$ , then  $\langle x^k \rangle = \langle x^{\gcd(n,k)} \rangle$  and  $|x^k| = n/\gcd(n, k)$ .

## Proof.

Note that since  $|x| = n$ , we have

$$\begin{aligned}\langle x^k \rangle &= \{(x^k)^t \mid t \in \mathbb{Z}\} = \{x^{kt+ns} \mid t, s \in \mathbb{Z}\} \\ &= \{x^d \mid d \in k\mathbb{Z} + n\mathbb{Z}\} = \{x^d \mid d \in \gcd(n, k)\mathbb{Z}\} \\ &= \{(x^{\gcd(n,k)})^r \mid r \in \mathbb{Z}\} = \langle x^{\gcd(n,k)} \rangle\end{aligned}$$

If  $(x^k)^t = 1$  for some  $t \in \mathbb{N}$ , then  $n \mid kt$ , also note that  $k \mid kt$ , hence  $\text{lcm}(n, k) \mid kt$ , i.e.,  $nk/\gcd(n, k) \mid kt$ , therefore  $n/\gcd(n, k) \mid t$ . Besides,

$$(x^k)^{n/\gcd(n,k)} = x^{kn/\gcd(n,k)} = x^{\text{lcm}(n,k)} = 1$$



# Cyclic Groups

## Remark

- ▶ Let  $|\langle x \rangle| < \infty$ , then  $y \in \langle x \rangle \Rightarrow |y| \text{ divides } |\langle x \rangle|$ .
- ▶ Let  $|x| = n \in \mathbb{N} \setminus \{0\}$ , then

$$\langle x^i \rangle = \langle x^j \rangle \Leftrightarrow |x^i| = |x^j| \Leftrightarrow \gcd(n, i) = \gcd(n, j)$$

In particular,

$$\langle x \rangle = \langle x^j \rangle \Leftrightarrow |x| = |x^j| \Leftrightarrow \gcd(n, j) = 1$$

For example,

$$\langle k \rangle = \mathbb{Z}/n\mathbb{Z} \Leftrightarrow \gcd(n, k) = 1$$

# Cyclic Groups

## Theorem (Fundamental Theorem of Cyclic Groups)

- ▶ Every subgroup of a cyclic group is cyclic.
- ▶ If  $|\langle x \rangle| = n \in \mathbb{N} \setminus \{0\}$ , then the order of any subgroup of  $\langle x \rangle$  divides  $n$ .
- ▶ For each  $k \mid n$  with  $k > 0$ , the group  $\langle x \rangle$  has exactly one subgroup of order  $k$ , i.e.,  $\langle x^{n/k} \rangle$ .

### Proof.

- ▶ Suppose  $G = \langle x \rangle$  is cyclic, i.e.,  $G = \{x^t \mid t \in \mathbb{Z}\}$ . If  $H \leq G$ , then  $H = \{x^t \mid t \in S \leq \mathbb{Z}\}$ ,<sup>3</sup> where  $S = m\mathbb{Z}$ ,  $m \in \mathbb{N}$ . Hence  $H = \{x^t \mid t \in m\mathbb{Z}, m \in \mathbb{N}\} = \{(x^m)^t \mid t \in \mathbb{Z}\} = \langle x^m \rangle$ , which is cyclic.
- ▶ Consider  $H \leq \langle x \rangle$ , then  $H = \langle x^m \rangle$  for some  $m \in \mathbb{N} \setminus \{0\}$ . Now  $|\langle x^m \rangle| = |x^m| = n/\gcd(n, m)$ , which divides  $n$ .
- ▶ If  $|\langle x^m \rangle| = k = n/\gcd(n, m)$ , then  $\langle x^m \rangle = \langle x^{\gcd(n, m)} \rangle = \langle x^{n/k} \rangle$ . □

---

3. The map  $f : S \rightarrow G$ ,  $t \mapsto x^t$  is a group homomorphism.

# Applications of Cyclic Groups

## Euler's Totient Function

The **Euler's Totient Function**, or the **Euler phi function**, denoted  $\varphi(n)$  or  $\phi(n)$  counts the number of positive integers less than  $n$  and relatively prime to  $n$ , i.e.

$$\varphi(n) = |\{k \in \mathbb{N} \mid \gcd(k, n) = 1, 1 \leq k \leq n\}|$$

In particular, given  $p \in \mathbb{P}$ ,

- ▶  $\varphi(p) = p - 1$ .
- ▶  $\varphi(p^k) = p^k - p^{k-1}$  for  $k \in \mathbb{N} \setminus \{0\}$ . Since the numbers

$$1 \cdot p, 2 \cdot p, 3 \cdot p, \dots, p^{k-1} \cdot p$$

are NOT relatively prime to  $p$ .

|              |   |   |   |   |   |   |   |   |   |    |    |    |
|--------------|---|---|---|---|---|---|---|---|---|----|----|----|
| $n$          | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $\varphi(n)$ | 1 | 1 | 2 | 2 | 4 | 2 | 6 | 4 | 6 | 4  | 10 | 4  |

# Applications of Cyclic Groups

## Lemma

Given a cyclic group  $C$  with order  $|C| = n$ , if  $d > 0$  and  $d \mid n$ , then the number of elements of order  $d$  in  $C$  is given by  $\varphi(d)$ .

## Proof.

Since the group has **exactly one** subgroup of order  $d$ , which is also cyclic.

Denote this subgroup by  $C_d = \langle x \rangle$  for some  $x \in C$  with  $x^d = 1$ . Now, since  $\langle x^k \rangle = \langle x \rangle$  iff  $|x^k| = |x| = d$  iff  $\gcd(d, k) = 1$ , hence the number of elements of order  $d$  is given by  $\varphi(d)$ , which is independent of  $n$ . □

## Divisor Sum (Gauss)

Given  $n \in \mathbb{N} \setminus \{0\}$ , then

$$\sum_{d|n} \varphi(d) = n$$

where the sum is over all positive divisor  $d$  of  $n$ .

# Applications of Cyclic Groups

## Proof 1 (Counting generators).

Consider the cyclic group of order  $n$ , denoted by  $C_n$ . Since  $C_n$  can be partitioned into disjoint sets each containing generators of order  $d$  with  $d \mid n$ , each block of size  $\varphi(d)$ , therefore the equality follows. □

## Proof 2.

Consider the set of  $n$  fractions  $\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$ , and put each fraction in lowest terms of the form  $\frac{c}{d}$  where  $d$  is a positive divisor of  $n$ , and  $\gcd(c, d) = 1$ . For each denominator  $d$  there are  $\varphi(d)$  relatively prime numerators. The total number of fractions is given by  $\sum_{d|n} \varphi(d)$ .

For example, consider  $n = 20$ , then we have

$$\frac{1}{20}, \frac{2}{20}, \frac{3}{20}, \frac{4}{20}, \frac{5}{20}, \frac{6}{20}, \frac{7}{20}, \frac{8}{20}, \frac{9}{20}, \frac{10}{20}, \frac{11}{20}, \frac{12}{20}, \frac{13}{20}, \frac{14}{20}, \frac{15}{20}, \frac{16}{20}, \frac{17}{20}, \frac{18}{20}, \frac{19}{20}, \frac{20}{20}$$

which can be put into lowest terms as

$$\frac{1}{20}, \frac{1}{10}, \frac{3}{20}, \frac{1}{5}, \frac{1}{4}, \frac{3}{10}, \frac{7}{20}, \frac{2}{5}, \frac{9}{20}, \frac{1}{2}, \frac{11}{20}, \frac{3}{5}, \frac{13}{20}, \frac{7}{10}, \frac{3}{4}, \frac{4}{5}, \frac{17}{20}, \frac{9}{10}, \frac{19}{20}, \frac{1}{1}$$
 □

# Euler's Totient Function

A function  $f : \mathbb{N} \setminus \{0\} \rightarrow \mathbb{N} \setminus \{0\}$  is **multiplicative** if  $f(1) = 1$  and  $f(m_1 m_2) = f(m_1)f(m_2)$  for  $\gcd(m_1, m_2) = 1$ .

## Theorem

*The Euler's Totient Function  $\varphi$  is multiplicative.*

This is a consequence of the following more general fact.

## Theorem

*If  $f$  is any function such that the sum*

$$g(m) = \sum_{d|m} f(d)$$

*is multiplicative, then  $f$  is itself multiplicative.* (The converse is also true.  
cf., Graham, Knuth, & Patashnik, Concrete Mathematics, 2ed)

## Proof.

Induction on  $m$ .

**base case ( $m = 1$ ):** True because  $f(1) = g(1) = 1$ .

## Euler's Totient Function

### Proof (Cont.)

**inductive case ( $m > 1$ ):** assume the inductive hypothesis that  $f(m_1 m_2) = f(m_1)f(m_2)$  if  $\gcd(m_1, m_2) = 1$  and  $m_1 m_2 < m$ . Now if  $m = m_1 m_2$  and  $\gcd(m_1, m_2) = 1$ , then

$$g(m_1 m_2) = \sum_{d|m_1 m_2} f(d) = \sum_{d_1|m_1} \sum_{d_2|m_2} f(d_1 d_2)$$

where  $\gcd(d_1, d_2) = 1$  since all divisors of  $m_1$  are relatively prime to divisors of  $m_2$ . By induction hypothesis,  $f(d_1 d_2) = f(d_1)f(d_2)$  except possibly when  $d_1 = m_1$  and  $d_2 = m_2$ . Thus

$$g(m_1 m_2) = \sum_{d_1|m_1} f(d_1) \sum_{d_2|m_2} f(d_2) - f(m_1)f(m_2) + f(m_1 m_2)$$

But we also have  $g(m_1 m_2) = g(m_1)g(m_2)$ , hence  $f(m_1 m_2) = f(m_1)f(m_2)$ .

□

# Symmetric Group

## Symmetric Group $S_n$

Given  $n \in \mathbb{N} \setminus \{0\}$ , we have the following **symmetric group of degree  $n$** ,

$$\begin{aligned} S_n &= \{\text{All permutations on } n \text{ letters/numbers}\} \\ &= \text{Sym}\{1, 2, 3, \dots, n\} \\ &= \{f : [n] \rightarrow [n] \mid f \text{ bijective}\} \end{aligned}$$

Note that it is a finite group of order  $n!$ , i.e.,  $|S_n| = n!$ .

## Examples

- ▶  $S_1 = \{e\}$ .
- ▶  $S_2 = \{e, \tau\}$ , where  $e, \tau : [2] \rightarrow [2]$ , with

$$\begin{aligned} e(1) &= 1, & e(2) &= 2 \\ \tau(1) &= 2, & \tau(2) &= 1 \end{aligned}$$

| $\circ$ | e      | $\tau$ |
|---------|--------|--------|
| e       | e      | $\tau$ |
| $\tau$  | $\tau$ | e      |

Observe that  $\tau \circ \tau = e$ , i.e.,  $\tau = \tau^{-1}$ .

# Symmetric Group

## Examples

- $S_3 = \{e, \tau, \tau', \tau'', \sigma, \sigma'\}$ .

Use cycle notation, such that

$$e = () = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

$$\tau = (12) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\sigma = (123) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$\tau' = (23) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

$$\sigma' = (132) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

$$\tau'' = (13) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

Abbreviate  $\tau \circ \sigma$  as  $\tau\sigma$ ,

$$\tau\sigma(1) = \tau(\sigma(1)) = \tau(2) = 1$$

$$\tau\sigma(2) = \tau(\sigma(2)) = \tau(3) = 3$$

$$\tau\sigma(3) = \tau(\sigma(3)) = \tau(1) = 2$$

Hence  $\tau\sigma = \tau'$ .

## Symmetric Group

Similarly, we have the following multiplication table

| $\circ$   | $e$       | $\tau$    | $\tau'$   | $\tau''$  | $\sigma$  | $\sigma'$ |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| $e$       | $e$       | $\tau$    | $\tau'$   | $\tau''$  | $\sigma$  | $\sigma'$ |
| $\tau$    | $\tau$    | $e$       | $\sigma$  | $\sigma'$ | $\tau'$   | $\tau''$  |
| $\tau'$   | $\tau'$   | $\sigma'$ | $e$       | $\sigma$  | $\tau''$  | $\tau$    |
| $\tau''$  | $\tau''$  | $\sigma$  | $\sigma'$ | $e$       | $\tau$    | $\tau'$   |
| $\sigma'$ | $\sigma'$ | $\tau'$   | $\tau''$  | $\tau$    | $e$       | $\sigma$  |
| $\sigma$  | $\sigma$  | $\tau''$  | $\tau$    | $\tau'$   | $\sigma'$ | $e$       |

## Corollary

The group  $S_n$  is nonabelian for  $n \geq 3$ .

## Proof.

Consider the subgroup  $S_3 \subset S_n$ . □



# Facts About General Permutations

## Cycle Notation

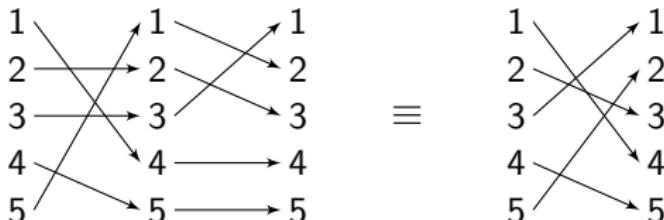
- ▶ Every permutation of a finite set can be written as a cycle or as a product of disjoint cycles.
- ▶ If the pair of cycles  $\alpha = (a_1 a_2 \cdots a_m)$  and  $\beta = (b_1 b_2 \cdots b_n)$  have no entries in common, i.e.,  $\alpha$  and  $\beta$  are **disjoint**, then  $\alpha\beta = \beta\alpha$ . (Such  $\alpha$  is called a **cycle of length m** or an **m-cycle**.)
- ▶ The order of a permutation of a finite set written in disjoint cycle form is the **least common multiple** of the lengths of the cycles.

- $|(132)(45)| = 6$

- $|(123)(456)(78)| = 6$

- $|(1432)(56)| = 4$

- $|(123)(145)| = |(14523)| = 5$



# Facts About General Permutations

## Cycles and Transpositions

A permutation of the form  $(ab)$  where  $a \neq b$  is called a **transposition**.

- ▶ Every permutation in  $S_n$ ,  $n > 1$ , is a product of transpositions.
- ▶ If  $\sigma = \beta_1\beta_2 \cdots \beta_r$ , where  $\beta_i$ 's are transpositions, then  $r$  is even.

## Even and Odd Permutations

A permutation that can be expressed as a product of an even/odd number of transpositions is called an **even/odd** permutation. (Note that this parity is well-defined.) For each permutation  $\sigma$ , define

$$\text{sgn}(\sigma) = \begin{cases} +1, & \text{if } \sigma \text{ is an even permutation.} \\ -1, & \text{if } \sigma \text{ is an odd permutation.} \end{cases}$$

- ▶ The set of even permutations in  $S_n$  forms a subgroup of  $S_n$ , denoted  $A_n$ , is called the **alternating group of degree  $n$** .
- ▶  $|A_n| = n!/2$  for  $n > 1$ .

# The Determinant

## Definition

Given a matrix  $A \in M_n(\mathbb{C})$ , the **determinant** function is given by

$$\det : M_n(\mathbb{C}) \rightarrow \mathbb{C}$$

$$(a_{ij}) \mapsto \det(a_{ij}) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

where  $S_n$  is the set of all permutations of the set  $\{1, \dots, n\} \subset \mathbb{N}$ , and  $\operatorname{sgn}(\sigma)$  the sign of the permutation  $\sigma$ .

# The Determinant

An equivalent definition of the determinant is as follows.

## Definition

The determinant  $\det : M_n(\mathbb{C}) \cong \underbrace{\mathbb{C}^n \times \cdots \times \mathbb{C}^n}_{n \text{ times}} \rightarrow \mathbb{C}$  is the **unique** function satisfying,

- (i) **alternating**, for all  $v \in \mathbb{C}^n$ ,  $\det(v_1, \dots, v_i, \dots, v_i, \dots, v_n) = 0$ , or equivalently **skew-symmetric**, i.e.,

$$\begin{aligned}\det(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n) \\ = -\det(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n)\end{aligned}$$

- (ii) **multilinear**, i.e., for all  $\lambda, \mu \in \mathbb{C}$ ,  $v_i, u \in \mathbb{C}^n$ ,  $i = 1, \dots, n$ ,

$$\begin{aligned}\det(v_1, \dots, v_{i-1}, \lambda v_i + \mu u, v_{i+1}, \dots, v_n) \\ = \lambda \det(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n) \\ + \mu \det(v_1, \dots, v_{i-1}, u, v_{i+1}, \dots, v_n)\end{aligned}$$

- (iii) **unitary**, i.e.,  $\det I_n = 1$ .

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Homomorphism

## Definition

Given groups  $G, G'$ , a homomorphism is a map  $f : G \rightarrow G'$  such that for all  $x, y \in G$ ,

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

## Examples

- ▶ Trivial homomorphism  $f : G \rightarrow G', x \mapsto 1_{G'} \in G'$ .
- ▶ Inclusion map  $\iota : H \hookrightarrow G, x \mapsto x$ , when  $H$  is a subgroup of  $G$ .
- ▶ The determinant function  $\det : GL_n(\mathbb{R}) \rightarrow \mathbb{R}^\times$ .
- ▶ The sign homomorphism  $\text{sgn} : S_n \rightarrow \{\pm 1\}$ .
- ▶ The exponential map  $\exp : (\mathbb{R}, +) \rightarrow \mathbb{R}^\times, x \mapsto e^x$ .
- ▶ The absolute value map  $|\cdot| : \mathbb{C}^\times \rightarrow \mathbb{R}^\times$ .
- ▶  $f : \mathbb{Z} \rightarrow S_2$ , even number  $\mapsto e$ , odd number  $\mapsto \tau$ .

# Homomorphism

## Example

$\text{sgn} = \det \circ \varphi.$

$$\begin{array}{ccc} & \text{sgn} & \\ S_3 & \xrightarrow[\sim]{\varphi} & GL_3(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times = GL_1(\mathbb{R}) \end{array}$$

1     $\mapsto \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto 1$

(123)     $\mapsto \begin{pmatrix} & 1 & \\ 1 & & \\ & & 1 \end{pmatrix} \mapsto 1$

(132)     $\mapsto \begin{pmatrix} & 1 & \\ & 1 & \\ 1 & & \end{pmatrix} \mapsto 1$

(12)     $\mapsto \begin{pmatrix} 1 & 1 & \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto -1$

(23)     $\mapsto \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto -1$

(31)     $\mapsto \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix} \mapsto -1$

# Homomorphism

## Theorem

Let  $f : G \rightarrow G'$  be a group homomorphism, then

- ▶ If  $a_1, \dots, a_k \in G$ , then  $f(a_1 \cdots a_k) = f(a_1) \cdots f(a_k)$ .
- ▶  $f(1_G) = 1_{G'}$ .
- ▶  $f(a^{-1}) = f(a)^{-1}$  for  $a \in G$ .

## Proof.

- ▶ Induction.
- ▶  $f(1_G) \cdot f(1_G) = f(1_G \cdot 1_G) = f(1_G)$ , thus  $f(1_G) = 1_{G'}$  by cancellation.
- ▶  $f(a^{-1})f(a) = f(a^{-1}a) = f(1_G) = 1_{G'}$ . □

# Image and Kernel of Homomorphisms

A group homomorphism determines two important **subgroups**: its image and its kernel.

## Definition

The **image** of a homomorphism  $f : G \rightarrow G'$ , often denoted by  $\text{im } f$ , or  $f(G)$ , is simply the image of as a map of sets:

$$\text{im } f = \{x \in G' \mid x = f(a) \text{ for some } a \in G\}$$

The **kernel** of  $f$ , denoted by  $\ker f$ , is the set of elements of  $G$  that are mapped to the identity in  $G'$ :

$$\ker f = \{a \in G \mid f(a) = 1_{G'}\}.$$

## Examples

- ▶ The determinant function  $\det : GL_n(\mathbb{R}) \rightarrow \mathbb{R}^\times$ .  $\ker \det = SL_n(\mathbb{R})$ .
- ▶ The sign homomorphism  $\text{sgn} : S_n \rightarrow \{\pm 1\} \cong \mathbb{Z}/2\mathbb{Z}$ .  $\ker \text{sgn} = A_n$ .

# Cosets

## Definition

Given a group  $G$ , if  $H \leq G$  is a subgroup and  $a \in G$ , the notation  $aH$  will stand for the set of all products  $ah$  with  $h \in H$ ,

$$aH = \{g \in G \mid g = ah \text{ for some } h \in H\}$$

This set is called a **left coset** of  $H$  in  $G$

## Example

- ▶  $1H = \{1, (12)\} = H.$
- ▶  $(12)H = \{(12), (12)(12)\} = \{(12), 1\} = H.$
- ▶  $(23)H = \{(23), (23)(12)\} = \{(23), (132)\} = (132)H.$
- ▶  $(31)H = \{(31), (31)(12)\} = \{(31), (123)\} = (123)H.$
- ▶  $(123)H = \{(123), (123)(12)\} = \{(123), (31)\} = (31)H.$
- ▶  $(132)H = \{(132), (132)(12)\} = \{(132), (23)\} = (23)H.$

Hence the left coset of  $\langle(12)\rangle$  in  $S_3$  is  $\{H, (23)H, (31)H\}$ .

# Homomorphisms

## Theorem

Let  $f : G \rightarrow G'$  be a group homomorphism, and let  $a, b \in G$ . Let  $K = \ker f$ . TFAE,

- (i)  $f(a) = f(b)$
- (ii)  $a^{-1}b \in K$
- (iii)  $b \in aK$
- (iv)  $aK = bK$

## Proof.

- (i)  $\Rightarrow$  (ii). Suppose  $f(a) = f(b)$ , then

$$f(a^{-1}b) = f(a^{-1})f(b) = f(a)^{-1}f(b) = 1$$

hence  $a^{-1}b \in \ker f = K$ .

- (ii)  $\Rightarrow$  (i). Reverse the argument above.
- (ii)  $\Leftrightarrow$  (iii). By definition of left coset.
- (iii)  $\Leftrightarrow$  (iv). Check the cosets of  $K$  in  $G$  are equivalence classes. (on this later) □

# Homomorphisms

## Corollary

A homomorphism  $f : G \rightarrow G'$  is injective iff  $\ker f = \{1_G\}$ .

## Proof.

- ▶ ( $\Leftarrow$ ). Suppose  $\ker f = \{1_G\}$ , then by previous theorem  $f(a) = f(b) \Rightarrow a^{-1}b \in \ker f \Rightarrow a^{-1}b = 1_G$ , i.e.,  $a = b$ .
- ▶ ( $\Rightarrow$ ). Since  $\ker f \leq G$ , it is always true that  $1_G \in \ker f$ , i.e.,  $\{1_G\} \subset \ker f$ . It is sufficient to show that  $\ker f \subset \{1_G\}$ , i.e., the only element in  $\ker f$  is  $1_G$ . Indeed, Suppose that  $a, b \in \ker f$ , then  $f(a) = f(b) = 1_{G'}$ , hence  $a = b$  by injectivity. Therefore  $\ker f = \{1_G\}$ . □

# Isomorphisms

## Definition

Given groups  $G$  and  $G'$ , an **isomorphism**  $f : G \rightarrow G'$  is a bijective group homomorphism, i.e., a bijection such that  $f(ab) = f(a)f(b)$  for all  $a, b \in G$ .

## Examples

- ▶  $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_{>0}, \times)$ ,  $x \mapsto e^x$ .
- ▶  $f : S_n \rightarrow GL_n(\mathbb{R})$ ,  $\sigma \mapsto$  permutation matrix.
- ▶  $f : G \rightarrow f(G) = \text{im } f$  is an isomorphism if  $f$  is injective.

## Check if $f : G \rightarrow G'$ is an isomorphism

Verify  $\ker f = \{1_G\}$  and  $\text{im } f = G'$ .

# Isomorphisms

## Theorem

If  $f : G \rightarrow G'$  is an isomorphism, its inverse map  $f^{-1} : G' \rightarrow G$  is also an isomorphism.

## Proof.

Since the inverse of a bijection is also a bijection, we only need to verify that  $f^{-1}$  is a homomorphism, that is,

$$f^{-1}(xy) = f^{-1}(x)f^{-1}(y) \text{ for all } x, y \in G'$$

Indeed. Note that  $f$  is bijective, then for  $x, y \in G'$ ,

$$\begin{aligned} f(f^{-1}(xy)) &= (f \circ f^{-1})(xy) = xy = (f \circ f^{-1})(x)(f \circ f^{-1})(y) \\ &= f(f^{-1}(x))f(f^{-1}(y)) = f(f^{-1}(x) \cdot f^{-1}(y)) \end{aligned}$$

Again, since  $f$  is bijective and we are done. □

## Cosets

Given a subgroup  $H$  of  $G$ , then the cosets of  $H$  are equivalence classes.

Denote  $a \equiv b$  if  $b \in aH$ . Indeed,

- ▶ Reflexivity.  $a = a \cdot 1$  and  $1 \in H$ , hence  $a \equiv a$ .
- ▶ Symmetry. Suppose  $a \equiv b$ , then  $b \in aH$  hence  $b = ah$  for some  $h \in H$ . Hence  $a = bh^{-1}$ , but  $h^{-1} \in H$ . Therefore  $a \in bH$ , i.e.,  $b \equiv a$ .
- ▶ Transitivity. Suppose  $a \equiv b$  and  $b \equiv c$ , then  $b = ah$  and  $c = bh'$  for some  $h, h' \in H$ . Therefore  $c = ahh'$ . Note that  $hh' \in H$  (since  $H$  is a subgroup), hence  $c \in aH$ , i.e.,  $c \equiv a$ .

## Corollary

*The left cosets of a subgroup  $H$  of a group  $G$  partition the group.*

## Remark

The subgroup  $H$  is a particular **left** coset since  $H = 1 \cdot H$ .

## Cosets

### Example

Let  $S_3 = \langle \tau, \sigma \mid \tau^2 = \sigma^3 = 1, \tau\sigma\tau = \sigma^{-1} \rangle$ . Then

- ▶  $H := \langle \tau \rangle = \{1, \tau\} = \tau H,$
- ▶  $\sigma H = \{\sigma, \sigma\tau\} = \sigma\tau H,$
- ▶  $\sigma^2 H = \{\sigma^2, \sigma^2\tau\} = \sigma^2\tau H,$

form a partition of  $S_3$ . Similarly

- ▶  $K := \langle \sigma \rangle = \{1, \sigma, \sigma^2\} = \sigma K = \sigma^2 K,$
- ▶  $\tau K = \{\tau, \tau\sigma, \tau\sigma^2\} = \tau\sigma K = \tau\sigma^2 K,$

also form a partition of  $S_3$ .

# Cosets

## Definition

The number of **left cosets** of a subgroup is called the **index** of  $H$  in  $G$ .  
The index is denoted by  $[G : H]$  (which could be infinite if  $|G| = \infty$ ).

## Example

|       |      |
|-------|------|
| 1     | (12) |
| (132) | (23) |
| (123) | (13) |

|       |      |
|-------|------|
| 1     | (12) |
| (132) | (23) |
| (123) | (13) |

$$[S_3 : \langle (12) \rangle] = 3.$$

$$[S_3 : \langle (123) \rangle] = 2.$$

# Cosets

## Lemma

All left cosets  $aH$  of a subgroup  $H$  of a group  $G$  have the **same** order.

## Proof.

The map  $h \mapsto ah$  induces a bijective map

$$\begin{aligned} H &\mapsto aH \\ a^{-1}(aH) &\leftrightarrow aH \end{aligned}$$

□

## Counting Formula

Note that the cosets all have the same order, and since they **partition** the group, then we have the **Counting Formula**

$$|G| = |H| \cdot [G : H]$$

$$(\text{order of } G) = (\text{order of } H) \cdot \left( \begin{array}{c} \text{number of left} \\ \text{cosets of } H \end{array} \right)$$

# Lagrange's Theorem

## Theorem (Lagrange's Theorem)

*Let  $H$  be a subgroup of a finite group  $G$ . The order of  $H$  divides the order of  $G$ .*

### Proof.

By applying the counting formula. □

## Corollary

*The order of an element of a finite group divides the order of the group.*

### Proof.

Let  $g \in G$ , then  $H := \langle g \rangle \leq G$ , and recall

$$H = \langle g \rangle = \{1, g, g^2, \dots, g^{m-2}, g^{m-1}\}$$

where  $|G| = m = \text{order of } g$ . □

# Lagrange's Theorem

## Corollary

Given a group  $G$ , with  $|G| = p$  prime. Let  $g \in G$ ,  $g \neq 1$ , then  $G = \langle g \rangle$  which is cyclic.

## Proof.

Let  $g \in G$  and  $g \neq 1$ , note that the order of  $g$  divides  $|G| = p$ , which is prime, hence the order of  $g$  is  $p$ . Therefore  $|\langle g \rangle| = p$ . Note that  $\langle g \rangle \subset G$ , with  $|\langle g \rangle| = |G| = p$ , hence  $G = \langle g \rangle$ , which is cyclic.  $\square$

## Remark

- ▶  $g^{|G|} = 1_G$  for all  $g \in G$ .
- ▶ The only subgroups of  $G$  with prime order are  $\{1\}$  and  $G$ .
- ▶ This classifies groups of prime order  $p$ . They form **one** isomorphism class, the class of the cyclic groups of order  $p$ .

# Lagrange's Theorem

## Example

Given group  $G$  of order 6, then

- ▶  $G$  contains an element of order 3. Indeed, if  $G$  has an element of order 6, then it is cyclic, so contains an element of order 3. If  $G$  does not have elements of order 3 or 6, then all non-identity elements of  $G$  have order 2. In this case, for all  $x, y \in G$  we have  $xy = (xy)^{-1} = y^{-1}x^{-1} = yx$ , hence  $G$  is abelian. Then for  $x, y \in G$  with  $x \neq y$ ,  $\{1, a, b, ab\}$  form a subgroup of  $G$  of order 4, but this contradicts Lagrange's theorem. Therefore  $G$  must contain an element of order 3.
- ▶  $G$  contains an element of order 2. Indeed, if it did not, then all non-identity elements would have order 3. But elements of order 3 come in pairs (e.g.,  $x$  and  $x^{-1}$ ), but there are are odd number of non-identity elements (i.e., 5), which is a contradiction. hence there must be an element of order 2.

# Lagrange's Theorem

## Corollary

Let  $G, G'$  be finite groups, and  $f : G \rightarrow G'$  a homomorphism. Then

1.  $|G| = |\ker f| \cdot |\text{im } f|$ ,
2.  $|\ker f|$  divides  $|G|$ ,
3.  $\text{im } f$  divides both  $|G|$  and  $|G'|$ .

$$\begin{array}{ccc} & G/\ker f & \\ q \uparrow & \swarrow \tilde{f} & \\ G & \xrightarrow{f} & \text{im } f \end{array}$$

## Proof.

1. Note that  $\ker f$  is a subgroup, then  $\tilde{f} : G/\ker f \rightarrow \text{im } f$  is a set-theoretic bijection between cosets of  $\ker f$  and elements of  $\text{im } f$ . Thus we have the counting formula

$$[G : \ker f] = |\text{im } f|$$

2. Follows from counting formula.
3. Follows from counting formula and Lagrange's theorem (Note that  $\text{im } f \leq G'$ ).



## Lagrange's Theorem

Compare the previous theorem with the following result in linear algebra.

### Remark (Rank-Nullity Theorem)

Given  $T : V \rightarrow W$  a linear map, then

$$\dim V = \dim \ker T + \dim \text{im } T$$

$$\left[ \begin{array}{ccccccccc} & \xleftarrow{\dim V \text{ columns}} & & & & & & & \\ \boxed{(*)} & * & * & * & * & * & * & * & * \\ 0 & \boxed{(*)} & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \boxed{(*)} & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \boxed{(*)} & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & & & & \end{array} \right]$$

# Right Cosets

## Definition

The **right cosets** of a subgroup  $H \leq G$  are the sets

$$Ha := \{ha \mid h \in H\}$$

## Example

Consider  $\langle(12)\rangle \leq S_3$ .

|       |       |
|-------|-------|
| 1     | (12)  |
| (132) | (23)  |
| (13)  | (123) |

Left cosets of  $\langle(12)\rangle$ .

|       |       |
|-------|-------|
| 1     | (12)  |
| (132) | (23)  |
| (13)  | (123) |

Right cosets of  $\langle(12)\rangle$ .

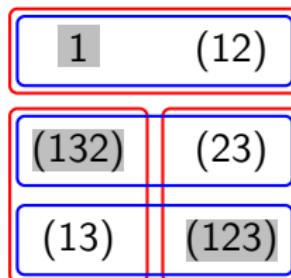
# Group Transversals

## Definition

Given a group  $G$ , and subgroup  $H \leq G$ . A subset  $S \subset G$  is a left/right transversal for  $H$  in  $G$  if every left/right coset of  $H$  contains exactly one element of  $S$ .

## Theorem

*Common transversal always exists for subgroups of finite groups.*

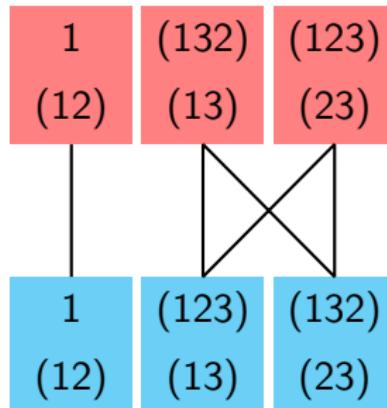


e.g.,  $\{1, (123), (132)\}$  is a common transversal for  $\langle(12)\rangle$  in  $S_3$ .

## Group Transversals

### Proof.

Suppose  $G = \bigcup_{k=1}^n x_k H = \bigcup_{k=1}^n Hy_k$ . We define the partial order on  $P = \{x_1 H, \dots, x_n H\} \cup \{Hy_1, \dots, Hy_n\}$  where  $x < y$  if  $x \in \{x_1 H, \dots, x_n H\}$ ,  $y \in \{Hy_1, \dots, Hy_n\}$ , and  $x \cap y \neq \emptyset$ . We know that the width of the poset is at least  $n$  (e.g.,  $\{x_1 H, \dots, x_n H\}$  is an antichain.) Suppose there exists a subset  $Q \subset P$  containing  $n + 1$  pairwise disjoint sets, then the size of their union exceeds the size of  $G$ , which is impossible. The rest follows from Dilworth theorem. □



# Normal Subgroup

## Definition

Given group  $G$ , and  $a, g \in G$ , the element  $gag^{-1} \in G$  is called the **conjugate of  $a$  by  $g$** .

## Definition

A subgroup  $N$  of  $G$  is a **normal subgroup**, denoted by  $N \trianglelefteq G$ , if for all  $a \in N$  and  $g \in G$ ,  $gag^{-1} \in N$ .

## Theorem

Given groups  $G, G'$ , and  $f : G \rightarrow G'$  a homomorphism, then  $\ker f \trianglelefteq G$ .

## Proof.

Let  $a \in \ker f$  and  $g \in G$ , then

$$f(gag^{-1}) = f(g)f(a)f(g^{-1}) = f(g) \cdot 1_{G'} \cdot f(g)^{-1} = 1_{G'} \quad \square$$

# Normal Subgroup

## Examples

- ▶  $SL_n(\mathbb{R}) \trianglelefteq GL_n(\mathbb{R})$ .
- ▶  $A_n \trianglelefteq S_n$ .
- ▶ Every subgroup of an abelian group is normal.
- ▶ The **center** of a group  $G$ , denoted by  $Z$ , is the set of elements that commute with every element of  $G$ :

$$Z := \{z \in G \mid zx = xz \text{ for all } x \in G\}$$

The center is always a normal subgroup. ( $zx = xz \Leftrightarrow x = zxz^{-1}$ .)

# Normal Subgroup

## Theorem

Let  $H \leq G$ , TFAE

1.  $H \trianglelefteq G$ , i.e.,  $ghg^{-1} \in H$  for all  $h \in H, g \in G$ .
2.  $gHg^{-1} = H$  for all  $g \in G$ .
3.  $gH = Hg$  for all  $g \in G$ .
4. Every left coset of  $H$  is a right coset.
5.  $H = \ker f$  for some homomorphism  $f : G \rightarrow X$ .
6. The quotient group  $G/H$  exists.

## Remark

If  $G$  is abelian, then all subgroups are normal.

## Definition

A group is **simple** if its only normal subgroup are the identity subgroup and the group itself.

# Classification of Finite Simple Groups

## The Periodic Table Of Finite Simple Groups

|               |
|---------------|
| $0, C_1, Z_2$ |
| 1             |
| 1             |

Dynkin Diagrams of Simple Lie Algebras

|                   |               |  |  |   |                                     |                                     |                             |           |
|-------------------|---------------|--|--|---|-------------------------------------|-------------------------------------|-----------------------------|-----------|
| $A_1(4), A_1(5)$  | $A_2(2)$      | $A_1(7)$                               | $B_2(3)$                               | $C_3(3)$                                      | $D_4(2)$                            | $2D_4(2^2)$                         | $G_2(2)'$                   | $C_2$     |
| $A_5$             |               |  |  |   |                                     |                                     |                             | $2$       |
| 60                |               | 168                                    |  |   |                                     |                                     |                             |           |
| $A_1(9), B_2(2)'$ | ${}^2G_2(3)'$ | $A_1(8)$                               | $C_6$                                  |   |                                     |                                     |                             | $C_3$     |
| $A_6$             |               |  |  |   |                                     |                                     |                             | 3         |
| 360               |               | 504                                    |  |   |                                     |                                     |                             |           |
| $A_7$             | $A_1(11)$     | $E_6(2)$                               | $E_7(2)$                               | $E_8(2)$                                      | $F_4(2)$                            | $G_2(3)$                            | ${}^3D_4(2^3)$              | $C_5$     |
| 2320              | 660           | 234 641 575 523<br>605 379 279 400     | 3 301 126<br>605 368 400               | 4245 696<br>211 341 312                       | 76 532 479 683<br>77 483 639 200    | 29 120                              | ${}^2B_2(2^3)$              | 2         |
| $A_8$             | $A_1(13)$     | $E_6(3)$                               | $E_7(3)$                               | $E_8(3)$                                      | $F_4(4)$                            | ${}^3D_4(3^2)$                      | ${}^2E_6(2^2)$              | $A_1(11)$ |
| 20160             | 1092          | 234 641 575 523<br>605 379 279 400     | 3 301 126<br>605 368 400               | 5734 420 792 836<br>471 844 743 400           | 251 596 800                         | 20 560 831 566 912                  | ${}^2B_2(2^3)$              | 2         |
| $A_9$             | $A_1(17)$     | $E_6(4)$                               | $E_7(4)$                               | $E_8(4)$                                      | $F_4(4)$                            | $G_2(5)$                            | ${}^3D_4(4^3)$              | $C_6$     |
| 181440            | 2448          | 234 641 575 523<br>605 379 279 400     | 3 301 126<br>605 368 400               | 5809 300 000                                  | 47 962 350<br>642 790 400           | 238 189 910 264<br>352 349 332 632  | ${}^2B_2(2^7)$              | 5         |
| $A_{10}$          | $A_8(q)$      | $E_6(q)$                               | $E_7(q)$                               | $E_8(q)$                                      | $F_4(q)$                            | $G_2(q)$                            | ${}^3D_4(q^3)$              | $C_7$     |
| $\frac{q!}{2}$    |               | $\frac{q^{15}(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^{24}(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^{30}(q^2-1)(q^2+1)(q^2+5)}{(q-1)^3}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^3D_4(q^3)$              | 7         |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2E_6(q^2)$              | $A_1(13)$ |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2B_2(2^{2k+1})$         | 11        |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2F_4(2^{2k+1})$         | $C_{11}$  |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2G_2(3^{2k+1})$         | 13        |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2B_2(3^7)$              | $Z_p$     |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2F_4(2^5)$              | $C_p$     |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2G_2(3^5)$              | $p$       |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2B_2(2^5)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2F_4(2^5)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2G_2(3^2)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2B_2(2^2)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2D_4(4^2)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(9)$                |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(16)$               |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(25)$               |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_3(9)$                |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_3(11)$               |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(64)$               |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(128)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(256)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(512)$              |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(1024)$             |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(2048)$             |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(4096)$             |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(8192)$             |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(16384)$            |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(32768)$            |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(65536)$            |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(131072)$           |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(262144)$           |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(524288)$           |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(1048576)$          |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(2097152)$          |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(4194304)$          |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(8388608)$          |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(16777216)$         |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(33554432)$         |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(67108864)$         |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(134217728)$        |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(268435456)$        |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(536870912)$        |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(1073741824)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(2147483648)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(4294967296)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(8589934592)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(1717986912)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(3435973824)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(6871947648)$       |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(13743895296)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(27487790592)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(54975581184)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(10995116368)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(21990233336)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(43980466672)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(87960933344)$      |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(175921866688)$     |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(351843733376)$     |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(703687333552)$     |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(1407374671088)$    |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(2814749342176)$    |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(56294986843544)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(112589973686912)$  |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(225179947373824)$  |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(450359894751648)$  |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(9007197891192)$    |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(18014395782384)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(36028791564768)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(72057583129536)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(14411516631920)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(2882303240000)$    |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(5764606400000)$    |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(11529212800000)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(23058425600000)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(46116812800000)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(9223360000000)$    |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(18446720000000)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(36893440000000)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(73786880000000)$   |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(147573760000000)$  |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(295147520000000)$  |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(590295040000000)$  |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(1180590080000000)$ |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(2360180000000000)$ |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(4720360000000000)$ |           |
|                   |               |  |  |   | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | $\frac{q^8(q^2-1)(q^2+1)}{(q-1)^2}$ | ${}^2A_2(9440720000000000)$ |           |
|                   |               |  |  |   |                                     |                                     |                             |           |

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Modular Arithmetic

## Definition

Given  $a, b \in \mathbb{Z}$ ,  $a$  and  $b$  are said to be **congruent modulo  $n$** , i.e.,

$$a \equiv b \pmod{n}$$

if  $n | b - a$ , i.e.,  $b = a + nk$  for some  $k \in \mathbb{Z}$ .

## Remark

This is an equivalence relation. The equivalence classes are called **congruence classes**.

- ▶  $a \equiv a \pmod{n}$  for all  $a \in \mathbb{Z}$ .
- ▶ If  $a \equiv b \pmod{n}$ , then  $b \equiv a \pmod{n}$ .
- ▶ If  $a \equiv b \pmod{n}$  and  $b \equiv c \pmod{n}$ , then  $a \equiv c \pmod{n}$ .

# Modular Arithmetic

## Congruence Classes

Let  $H = n\mathbb{Z} \trianglelefteq \mathbb{Z}$ , then the cosets of  $H$ , i.e., the congruence classes, are given by

$$[a]_n = \bar{a} = a + H = a + n\mathbb{Z} = \{a + kn \mid k \in \mathbb{Z}\}$$

The integers  $0, 1, \dots, n - 1$  are representatives for the  $n$  congruence classes.

## Notation

|            | Multiplicative<br>Notation         | Additive<br>Notation                      |
|------------|------------------------------------|---|
| Operation: | $ab$                               | $a + b$                                   |
| Identity:  | e or 1                             | 0   |
| Inverse:   | $a^{-1}$                           | $-a$                                      |
| Exponents: | $a^n = aa \cdots a$ ( $n$ factors) | $na = a + a + \cdots + a$ ( $n$ summands) |
|            | $a^{-n} = a^{-1} \cdots a^{-1}$    | $(-n)a = -a - a - \cdots - a$             |
|            | $a^m a^n = a^{m+n}$                | $(ma) + (na) = (m+n)a$                    |
|            | $(a^m)^n = a^{mn}$                 | $n(ma) = (mn)a$                           |
| Cosets:    | $aH$                               | $a + H$                                   |

# Modular Arithmetic

In an attempt to prove Fermat's Last Theorem,

## Theorem (Schur, 1916)

Let  $n \in \mathbb{N} \setminus \{0\}$ , then for all sufficiently large primes  $p$ , there are  $x, y, z \in \{1, \dots, p - 1\}$  such that  $x^n + y^n \equiv z^n \pmod{p}$ .

## Less Dramatic Examples

Given  $x, y, z \in \mathbb{Z}$ , then

- ▶  $x^2 + y^2 \equiv 0, 1, 2 \pmod{4}$ .
- ▶  $x^2 + y^2 + z^2 \equiv 0, 1, 2, 3, 4, 5, 6 \pmod{8}$ .
- ▶  $x^3 + y^3 + z^3 \equiv 0, \pm 1, \pm 2, \pm 3 \pmod{9}$ .

# Modular Arithmetic

## Theorem

There are  $n$  congruence classes modulo  $n$ , namely,  $\overline{0}, \overline{1}, \dots, \overline{n-1}$ . The index of the subgroup  $n\mathbb{Z}$  in  $\mathbb{Z}$  is  $[\mathbb{Z} : n\mathbb{Z}] = n$ .

## Proof.

Consider the function  $f : \mathbb{Z} \rightarrow \{0, 1, \dots, n-1\}$ ,  $x \mapsto x \bmod n$ . Note that  $f$  induces a bijection  $\tilde{f} : \mathbb{Z}/n\mathbb{Z} \rightarrow \{0, \dots, n-1\}$ . □

$$\begin{array}{ccc} \mathbb{Z}/n\mathbb{Z} & & \\ \uparrow q & \searrow \tilde{f} & \\ \mathbb{Z} & \xrightarrow{f} & \{0, \dots, n-1\} \end{array}$$

## Remark

- The set of congruence classes modulo  $n$  may be denoted by  $\mathbb{Z}/n\mathbb{Z}$ ,  $\mathbb{Z}/\mathbb{Z}n$ ,  $\mathbb{Z}_n$ , or  $\mathbb{Z}/(n)$ .
- It is the same to say  $\overline{a} = \overline{b}$ ,  $a = b$  in  $\mathbb{Z}/n\mathbb{Z}$ , or  $a \equiv b \pmod{n}$ .

# Modular Arithmetic

We can do “arithmetic” in  $\mathbb{Z}/n\mathbb{Z}$ , e.g.,

$$\bar{a} + \bar{b} = \overline{a + b}$$

$$\bar{a} \cdot \bar{b} = \overline{a \cdot b}$$

which are well-defined.

## Lemma

If  $a \equiv a' \pmod{n}$  and  $b \equiv b' \pmod{n}$ , then  $a + b \equiv a' + b' \pmod{n}$  and  $ab \equiv a'b' \pmod{n}$ .

## Proof.

Assume that  $a \equiv a' \pmod{n}$  and  $b \equiv b' \pmod{n}$ , then  $a' = a + rn$  and  $b' = b + sn$  for some  $r, s \in \mathbb{Z}$ . Then

- ▶  $a' + b' = a + b + (r + s)n$ , hence  $a + b \equiv a' + b' \pmod{n}$ .
- ▶  $a'b' = (a + rs)(b + sn) = ab + (as + rb + rns)n$ , hence  $ab \equiv a'b' \pmod{n}$ .



# Modular Arithmetic

$(\mathbb{Z}/n\mathbb{Z}, +)$  is a group

- ▶ Addition is associative. (inherited from  $\mathbb{Z}$ )

$$(\bar{a} + \bar{b}) + \bar{c} = \overline{a + b + c} = \bar{a} + (\bar{b} + \bar{c})$$

- ▶ Identity:  $\bar{0}$ .
- ▶ Inverses:  $-\bar{a} = \overline{n - a} = \overline{-a}$ .

i.e., the set of cosets of  $n\mathbb{Z} \subset \mathbb{Z}$  form a (quotient) group.

## Inheritance from $\mathbb{Z}$

The associative, commutative, and distributive laws hold for addition and multiplication of congruence classes. e.g.,

$$\begin{aligned}\bar{a}(\bar{b} + \bar{c}) &= \bar{a}(\overline{b + c}) = \overline{a(b + c)} \\ &= \overline{ab + ac} \\ &= \overline{ab} + \overline{ac} = \bar{a}\bar{b} + \bar{a}\bar{c}\end{aligned}$$

# Modular Arithmetic

## Multiplicative Group of Integers Modulo $n$

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{\bar{a} \in \mathbb{Z}/n\mathbb{Z} \mid \exists \bar{c} \in \mathbb{Z}/n\mathbb{Z} \text{ s.t. } \bar{a} \cdot \bar{c} = \bar{1}\}$$

- ▶ Closure: product of inverses are inverse of product.
- ▶ Associativity: inherited from  $\mathbb{Z}$ .
- ▶ Identity:  $\bar{1}$ .
- ▶ Inverses by construction.

## Theorem

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{\bar{a} \in \mathbb{Z}/n\mathbb{Z} \mid \gcd(a, n) = 1\}$$

## Proof.

- ▶ (LHS  $\supset$  RHS). If  $\gcd(a, n) = 1$ , then  $\exists r, s \in \mathbb{Z}$  such that  $ar + ns = 1$ , i.e.,  $ar - 1 \in n\mathbb{Z}$ , or  $\bar{a} \cdot \bar{r} = \bar{1}$ , so  $\bar{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$ .
- ▶ (LHS  $\subset$  RHS). Consider  $\bar{a} \cdot \bar{c} = \bar{1}$ , then  $ac - 1 = nb$  for some  $b \in \mathbb{Z}$ . Hence  $1 = ac + nb \in a\mathbb{Z} + n\mathbb{Z} = \gcd(a, n)\mathbb{Z}$ .



# Modular Arithmetic

## Finding Inverses

For example, we want to solve  $7x \equiv 1 \pmod{31}$ .

### Method I

By Euclidean algorithm, we can find integers  $x = 9$ ,  $y = -2$  such that  $7x + 31y = 1$ , i.e.

$$7 \times 9 + 31 \times (-2) = 1$$

hence  $7 \cdot 9 \equiv 1 \pmod{31}$ , i.e.,  $x \equiv 7^{-1} \equiv 9 \pmod{31}$ .

### Method II (Gauss), for prime modulus

By division algorithm (keep remainder with smallest absolute value),

$$31 = 7 \times 4 + 3 \quad \Rightarrow \quad 7 \times 4 \equiv -3 \pmod{31}$$

$$31 = 3 \times 10 + 1 \quad \Rightarrow \quad 3 \times 10 \equiv -1 \pmod{31}$$

Hence  $7 \cdot 4 \cdot 3 \cdot 10 \equiv 1$ , so  $7^{-1} \equiv 4 \cdot 10 \equiv 9 \pmod{31}$ .

## Fermat's (Little) Theorem

### Theorem (Fermat-I)

Given  $a \in \mathbb{Z}$  and  $p \in \mathbb{P}$ , such that  $(a, p) = 1$ , then

$$a^{p-1} \equiv 1 \pmod{p}$$

### Theorem (Fermat-II)

Given  $a \in \mathbb{Z}$  and  $p \in \mathbb{P}$ , then

$$a^p \equiv a \pmod{p}$$

### Remark

- ▶ (Fermat-I  $\Rightarrow$  Fermat-II). Clear by multiplying  $a$  on both sides.
- ▶ (Fermat-II  $\Rightarrow$  Fermat-I). Clear by multiplying  $a^{-1}$  on both sides.  $a^{-1} \pmod{p}$  exists because  $(a, p) = 1$ .

## Fermat's (Little) Theorem

### Proof of Fermat-II (Euler).

Induction on  $a \in \mathbb{N}$ .

**base case.** ( $a = 0$ ). True.

**inductive case.** ( $a \geq 0$ ). Assume the IH that  $a^p \equiv a \pmod{p}$  for some  $a \in \mathbb{N}$ , we want to show that  $(a+1)^p \equiv a+1 \pmod{p}$  also holds. Note that

$$(a+1)^p = a^p + \binom{p}{1}a^{p-1} + \binom{p}{2}a^{p-2} + \cdots + \binom{p}{p-1}a + 1$$

Now it is sufficient to show that  $p \mid \binom{p}{k}$  for  $p \in \mathbb{P}$ ,  $1 \leq k \leq p-1$ . Indeed, since

$$p! = \binom{p}{k} \cdot (p-k)!k!$$

Now note that  $p \mid p!$  but  $p \nmid [(p-k)!k!]$ , we have  $p \mid \binom{p}{k}$ . □



# Euler's Theorem

## Theorem (Euler)

For  $m \in \mathbb{N} \setminus \{0\}$  and  $a \in \mathbb{Z}$  such that  $\gcd(a, m) = 1$ ,

$$a^{\varphi(m)} \equiv 1 \pmod{m}$$

where  $\varphi(m)$  is the number of invertible integers modulo  $m$ .

## Proof.

Note that  $|(\mathbb{Z}/m\mathbb{Z})^\times| = \varphi(m)$ , which is divisible by the order of  $a \in (\mathbb{Z}/m\mathbb{Z})^\times$  by Lagrange's theorem. (In fact,  $a^{|G|} = 1_G \forall a \in G$ .) □

## Remark

Given  $p \in \mathbb{P}$ ,

- ▶ Fermat's theorem becomes Euler's theorem since  $\varphi(p) = p - 1$ .
- ▶  $\mathbb{Z}/p\mathbb{Z}$  is cyclic due to Lagrange's theorem.
- ▶  $(\mathbb{Z}/p\mathbb{Z})^\times$  is also cyclic, but NOT due to Lagrange's theorem.

# Fermat's Theorem

## Theorem

Given  $p \in \mathbb{P}$ , if  $p \mid n^2 + 1$ , then  $p = 2$  or  $p \equiv 1 \pmod{4}$ .

| $n$       | 1 | 2 | 3    | 4  | 5     | 6  | 7    | 8     |
|-----------|---|---|------|----|-------|----|------|-------|
| $n^2 + 1$ | 2 | 5 | 10   | 17 | 26    | 37 | 50   | 65    |
| $p$       | 2 | 5 | 2, 5 | 17 | 2, 13 | 37 | 2, 5 | 5, 13 |

## Proof.

If  $p$  is odd, then  $p \mid n^2 + 1 \Leftrightarrow n^2 \equiv -1 \pmod{p}$ , hence the order of  $n$  is not 1 or 2. (Note that  $1 \not\equiv -1 \pmod{p}$  since  $p$  is odd.) But since  $n^4 \equiv 1 \pmod{p}$ , we know that the order of  $n$  divides 4, hence the order of  $n$  is exactly 4. Also note that  $\gcd(n, p) = 1$ , hence by Fermat's theorem, we have  $n^{p-1} \equiv 1 \pmod{p}$ , so the order of  $n$  divides  $p - 1$ , that is,  $4 \mid p - 1$ , i.e.,  $p \equiv 1 \pmod{4}$ . □

## Euler's Theorem

### Example

For  $\varphi(8) = 4$ , by Euler's theorem

$$a^4 \equiv 1 \pmod{8}, \quad \text{for all } a \in \mathbb{Z} \text{ s.t. } \gcd(a, 8) = 1$$

Note that  $\gcd(a, 8) = 1$  leads to  $a = 1, 3, 5, 7$ . In fact,

$$a^2 \equiv 1 \pmod{8}$$

### Example

Let  $m = 35 = 5 \times 7$ , then by Fermat's theorem

- ▶  $a^6 \equiv 1 \pmod{7}$
- ▶  $a^4 \equiv 1 \pmod{5}$

Hence  $a^{\text{lcm}(4,6)} = a^{12} \equiv 1 \pmod{5, 7}$ , i.e.,  $a^{12} \equiv 1 \pmod{35}$ .

By Euler's Theorem,  $a^{\varphi(35)} = a^{24} \equiv 1 \pmod{35}$ .

# Fermat Primes

When is  $2^n + 1$  prime? ( $n > 0$ )

- ▶ If  $n > 1$ , odd, then NO. (since  $3 \mid (2^n + 1)$ )
- ▶ If  $n = ab$ ,  $b$  odd, also NO. (since  $(2^a + 1) \mid (2^n + 1)$ )

Therefore  $n = 2^m$ ,  $m \in \mathbb{N}$ .

## Fermat Primes

$$F_n = 2^{2^n} + 1.$$

- ▶  $F_0 = 2^{2^0} + 1 = 3 \in \mathbb{P}$ .
- ▶  $F_1 = 2^{2^1} + 1 = 5 \in \mathbb{P}$ .
- ▶  $F_2 = 2^{2^2} + 1 = 17 \in \mathbb{P}$ .
- ▶  $F_3 = 2^{2^3} + 1 = 257 \in \mathbb{P}$ .
- ▶  $F_4 = 2^{2^4} + 1 = 65537 \in \mathbb{P}$ .
- ▶  $F_5 = 2^{2^5} + 1 = 4274967297 = 641 \times 6700417$ . (Euler, 1732)

### FACT

If  $m$  is odd, then  $(-1)^m + 1 = 0$ , thus  $x^m + 1$  is divisible by  $x + 1$ . By long division, we have

$$x^m + 1 = (x+1)(x^{m-1} - x^{m-2} + \dots + 1)$$

## Testing Fermat Primes

Check  $F_4 = 2^{2^4} + 1 = 65537$  is prime

Suppose  $p \mid 65537$ ,  $p \leq \sqrt{65537}$ , that is,  $p \mid 2^{16} + 1$ , hence

$$2^{16} \equiv -1 \pmod{p}$$

$$2^{32} \equiv 1 \pmod{p}$$

Hence the order of 2 divides 32 but not 16, that is, the order of 2 is 32 modulo  $p$ . Therefore by Fermat's theorem,

$$2^{p-1} \equiv 1 \pmod{p} \Rightarrow p \equiv 1 \pmod{32}$$

Note that  $p \leq \sqrt{65537}$ , possible  $p$ 's are listed as follows

33, 65, 97, 129, 161, 193, 225

Among which we only need to check 97 and 193.

# Primality Testing of General Numbers

## Fermat Primality Test

Given  $n \in \mathbb{N}$ , calculate  $2^n \pmod n$ ,

- ▶ If  $2^n \not\equiv 2 \pmod n$ , then  $n$  is COMPOSITE.
- ▶ If  $2^n \equiv 2 \pmod n$ , then  $n$  is PROBABLY prime.

Such test is called *probabilistic test*.

## Task: Calculate $2^n \pmod n$

$n$  is usually large, e.g.,  $n \sim 10^{100}$

- ▶  $2^n$  is rediculously large.
- ▶ Takes spatial and temporal resources to calculate.
- ▶ Mod  $n$  after each multiplication of 2 is still slow.

# Fast Modular Exponentiation

Calculate  $a^b \bmod m$

1. Write  $b$  in binary, i.e.,

$$b = (b_{k-1} \cdots b_0)_2 = \sum_{j=0}^{k-1} b_j 2^j = b_{k-1} 2^{k-1} + \cdots + b_1 \cdot 2 + b_0,$$

with  $b_0, \dots, b_{k-1} \in \{0, 1\}$ , then

$$a^b = \prod_{j=0}^{k-1} a^{b_j 2^j} = a^{b_{k-1} 2^{k-1}} \times a^{b_{k-1} 2^{k-1}} \times \cdots \times a^{b_1 \cdot 2} \times a^{b_0}$$

2. Calculate  $a^{2^j} \bmod m$  for  $j = 0, \dots, k - 1$ , by noting that  $a^{2^{j+1}} = (a^{2^j})^2$
3. Multiply the terms for which  $b_k = 1$ .

Such **square and multiply** method is also known as **repeated squaring**.

## Fast Modular Exponentiation

Example: Test if 35 is prime.

Note that  $35 = (100011)_2 = 2^5 + 2^1 + 2^0$ , then

$$2^{35} = 2^{32} \times 2^2 \times 2^1$$

Next calculate

- ▶  $2^1 \equiv 2 \pmod{35}$ .
- ▶  $2^2 \equiv 2^2 \equiv 4 \pmod{35}$ .
- ▶  ~~$2^4 \equiv 4^2 \equiv 16 \pmod{35}$~~
- ▶  ~~$2^8 \equiv 16^2 \equiv 256 \equiv 11 \pmod{35}$~~
- ▶  ~~$2^{16} \equiv 11^2 \equiv 121 \equiv 16 \pmod{35}$~~
- ▶  $2^{32} \equiv 16^2 \equiv 11 \pmod{35}$ .

Now  $2^{35} \equiv 2^{32} \times 2^2 \times 2^1 \equiv 11 \times 4 \times 2 \equiv 18 \not\equiv 2 \pmod{35}$ .

Hence 35 is NOT prime.

# Fast Modular Exponentiation and Egyptian/Ethiopian/Russian Multiplication

Example:  $2^{35} \pmod{35}$

| HALVING | SQUARING  |
|---------|---|
| 35      | $2 \pmod{35}$                                   |
| 17      | $4 \pmod{35}$                                   |
| 8       | <del><math>16 \pmod{35}</math></del>            |
| 4       | <del><math>256 \equiv 11 \pmod{35}</math></del> |
| 2       | <del><math>121 \equiv 16 \pmod{35}</math></del> |
| 1       | $256 \equiv 11 \pmod{35}$                       |

$$2^{35} \equiv 2 \cdot 4 \cdot 11 \pmod{35}.$$

Example:  $35 \times 27$

| HALVING | DOUBLING       |
|---------|----------------|
| 35      | 27             |
| 17      | 54             |
| 8       | <del>108</del> |
| 4       | <del>216</del> |
| 2       | <del>432</del> |
| 1       | 864            |

$$35 \times 27 = 27 + 54 + 864 = 945.$$

# Carmichael Numbers

## Fermat Primality Test

Given  $n \in \mathbb{N}$ , calculate  $a^n \pmod{n}$ ,  $a < n$  (in general for many  $a$ )

- ▶ If  $a^n \not\equiv a \pmod{n}$ , then  $n$  is COMPOSITE. Such  $a$  is called a **Fermat witness**.
- ▶ If  $a^n \equiv a \pmod{n}$ , then
  - ▶  $n$  is prime.
  - ▶  $n$  is composite, such  $a$  is called a **Fermat Liar**.

## Definition

A **Carmichael number** is a composite number  $n$  for which

$$a^n \equiv a \pmod{n} \text{ for all } a.$$

## Remark

Carmichael numbers have **NO** Fermat witnesses.

## Carmichael Numbers

The first few Carmichael numbers are 561, 1105, 1729, 2465, 2821, 6601, 8911, ...

### Example

Let  $n = 561 = 3 \times 11 \times 17$ , note that By Fermat's Theorem, for a coprime to 561,

- ▶  $a^{3-1} \equiv 1 \pmod{3}$
- ▶  $a^{11-1} \equiv 1 \pmod{11}$
- ▶  $a^{17-1} \equiv 1 \pmod{17}$

Now note that  $\text{lcm}(3 - 1, 11 - 1, 17 - 1) = 80$  which divides  $560 = 561 - 1$ . Therefore for a coprime to 561,

$$a^{561-1} \equiv 1 \pmod{3, 11, 17}$$

hence

$$a^{561} \equiv a \pmod{561} \text{ for all } a \in \mathbb{Z}$$

## Carmichael Numbers

For 100-digit numbers, less than 1 in 10<sup>30</sup> are Carmichael numbers. For 200-digit numbers, the chances are even less.

### Remark

If we randomly choose a 200-digit number  $n$ , and test  $\approx 100$  different values of  $a$  without getting a Fermat witness, then we can be almost certain that  $n$  is prime.

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# Sunzi's Problem

Sunzi asks:

*There are certain things whose number is unknown. When divided by 3, the remainder is 2; when divided by 5, the remainder is 3; and when divided by 7, the remainder is 2. What will be the number of things?*

今有物不知其数，三三数之剩二，五五数之剩三，  
七七数之剩二，问物几何？

In Language of Congruences

Find  $x$  such that

$$x \equiv 2 \pmod{3},$$

$$x \equiv 3 \pmod{5},$$

$$x \equiv 2 \pmod{7}.$$

# Sunzi's Problem

## Solution Algorithm

三人同行七十希，五树梅花廿一支，  
七子团圆正半月，除百零五便得知。

## Solution in modern mathematical language

$$x \equiv 2 \times 70 + 3 \times 21 + 2 \times 15 = 233 \equiv 23 \pmod{105}$$

### Remark

- ▶  $70 \equiv 1 \pmod{3}$ ,  $70 \equiv 0 \pmod{5}$ ,  $70 \equiv 0 \pmod{7}$ ;
- ▶  $21 \equiv 0 \pmod{3}$ ,  $21 \equiv 1 \pmod{5}$ ,  $21 \equiv 0 \pmod{7}$ ;
- ▶  $15 \equiv 0 \pmod{3}$ ,  $15 \equiv 0 \pmod{5}$ ,  $15 \equiv 1 \pmod{7}$ ;
- ▶  $105 \equiv 0 \pmod{3}$ ,  $105 \equiv 0 \pmod{5}$ ,  $105 \equiv 0 \pmod{7}$ .

# Sunzi's Problem

## General Form

Given  $x \equiv a_i \pmod{m_i}$ ,  $i = 1, \dots, r$ ,  $a_1, \dots, a_r \in \mathbb{Z}$ , and  $m_1, \dots, m_r$  are pairwise relatively prime. The unique solution is given by

$$x = a_1y_1 + a_2y_2 + \cdots + a_r y_r \pmod{m}$$

where  $m = m_1 \cdots m_r$  and  $y_i = \delta_{ij} \pmod{m_j}$ , e.g.,  $y_i = (m/m_i)^{\varphi(m_i)}$ .

## Lagrange interpolation (also cf., Green's function, matrix inverse)

Given a set of  $k+1$  data points  $(x_0, y_0), (x_1, y_1), \dots, (x_k, y_k)$ , with distinct  $x_j$ 's. The **interpolation polynomial in the Lagrange form** is a linear combination  $L = y_0\ell_0 + \cdots + y_k\ell_k$ , with  $\ell_i$  satisfying  $\ell_i(x_j) = \delta_{ij}$ , e.g.,

$$\ell_j(x) := \prod_{\substack{0 \leq m \leq k \\ m \neq j}} \frac{x - x_m}{x_j - x_m} = \frac{(x - x_0)}{(x_j - x_0)} \cdots \frac{(x - x_{j-1})}{(x_j - x_{j-1})} \frac{(x - x_{j+1})}{(x_j - x_{j+1})} \cdots \frac{(x - x_k)}{(x_j - x_k)}$$

# Sunzi's Problem

## Matrix Inverse

Given an  $n \times n$  invertible matrix  $A$ , its inverse  $A^{-1}$  can be found by solving

$$Ax_1 = e_1, \quad Ax_2 = e_2, \quad \dots, \quad Ax_n = e_n$$

where

$$e_k = [0 \ 0 \ \cdots \ 0 \ \underset{k\text{-th}}{\downarrow} \ 1 \ 0 \ \cdots \ 0]^T$$

Now the general solution to  $Ax = b$  can be solved by first recognizing that  $b = \sum_{k=1}^n b_k e_k$ , then

$$x = A^{-1}b = A^{-1}\left(\sum_{k=1}^n b_k e_k\right) = \sum_{k=1}^n b_k A^{-1}e_k = \sum_{k=1}^n b_k x_k$$

## Remark

Recall the procedure of finding matrix inverse by Gauss-Jordan elimination:  $[A|I_n] \rightsquigarrow [I_n|B]$ , then  $B = A^{-1}$ .

## Sunzi's Problem

### Green's Function

Given a differential equation  $Lu = f$  with boundary condition  $Bu = 0$  over certain domain  $D$ , we first solve the following equation

$$Lg(x; \xi) = \delta(x - \xi), \quad Bg(x; \xi) = 0$$

where the solution  $g(x; \xi)$  is known as the **Green's function**. Now the solution to original equation is given by

$$u(x) = \int_D g(x; \xi) f(\xi) d\xi$$

If the differential operator  $L$  is time-invariant, then the solution is given by a convolution

$$u(x) = \int_D g(x - \xi) f(\xi) d\xi = (g * f)(x)$$

where  $g(x) = g(x; 0)$  and  $Lg(x) = \delta(x)$ .

## Sunzi's Problem

Find the smallest  $x \in \mathbb{N}$  (or all  $x \in \mathbb{Z}$ ) such that

$$x \equiv a_1 \pmod{m_1}$$

$$x \equiv a_2 \pmod{m_2}$$

⋮

$$x \equiv a_r \pmod{m_r}$$

### Remark

- ▶ No constraints on the remainders  $a_1, \dots, a_r$ .
- ▶ The moduli  $m_1, \dots, m_r$  are **pairwise** relatively prime. (This is NOT equivalent to  $\gcd(m_1, \dots, m_r) = 1$ .)

# Product Group

## Definition

Given groups  $G$  and  $G'$ , the product group  $(G \times G', \cdot_{\times})$  is the set  $G \times G'$  equipped with the group law

$$\begin{aligned}\cdot_{\times} : (G \times G') \times (G \times G') &\rightarrow G \times G' \\ ((g, g'), (h, h')) &\mapsto (g, g') \cdot_{\times} (h, h') = (gh, g'h')\end{aligned}$$

## Remark

- ▶ The identity element of  $(G \times G', \cdot_{\times})$  is given by  $(1_G, 1_{G'})$ .
- ▶ The inverse of  $(g, g')$  is  $(g^{-1}, g'^{-1})$ .
- ▶ Associativity is inherited from  $G$  and  $G'$ .

## Chinese Remainder Theorem

Let  $m, n \in \mathbb{N} \setminus \{0\}$  and  $\gcd(m, n) = 1$ , then  $C_{mn} \cong C_m \times C_n$ . ( $C_n$  is the cyclic group of order  $n$ .) Note that  $C_4 \not\cong C_2 \times C_2$ .

# Chinese Remainder Theorem

## Theorem

$$\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \text{ if } \gcd(m, n) = 1.$$

## Proof.

Consider the mapping

$$\begin{aligned} f : \mathbb{Z}/mn\mathbb{Z} &\rightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ x &\mapsto (x, x) \end{aligned}$$

which is obviously a homomorphism. We show that it is bijective.

- ▶ Injectivity. We need to show  $f(x) = (0, 0) \Rightarrow x \equiv 0 \pmod{mn}$ .  
Indeed, since if  $m, n | x$ , then  $\text{lcm}(m, n) | x$ , but  $\gcd(m, n) = 1$ , so  $mn | x$ .
- ▶ Surjectivity. By dimension count (basically pigeonhole). □

# Chinese Remainder Theorem (General Form)

## Theorem

$\mathbb{Z}/m_1 \cdots m_r \mathbb{Z} \cong \mathbb{Z}/m_1 \mathbb{Z} \times \mathbb{Z}/m_2 \mathbb{Z} \times \cdots \times \mathbb{Z}/m_r \mathbb{Z}$  if  $\gcd(m_i, m_j) = 1$  for  $i \neq j$ . (Induction on  $r$ .)

## Lemma (Base case for induction)

Given the system  $x \equiv a \pmod{m}$  and  $x \equiv b \pmod{n}$  with  $\gcd(m, n) = 1$ , the solution can be found as follows,

1. Find  $u$  and  $v$  such that  $mu + nv = 1$ .
2. Then  $t = bmu + anv \pmod{mn}$  is a solution.

## Example

Consider  $x \equiv 2 \pmod{3}$ ,  $x \equiv 3 \pmod{5}$ ,  $x \equiv 2 \pmod{7}$ . We can apply the Euclidean algorithm (or by guessing),

- Then consider the first two, we have  $x \equiv 8 \pmod{15}$ .
- Combine with the third one, we have  $x \equiv 23 \pmod{105}$ .

## Chinese Remainder Theorem (General Form)

### Example (Cont.)

We first solve  $x \equiv 2 \pmod{3}$ ,  $x \equiv 3 \pmod{5}$ , that is,

$$x = 2 + 3y = 3 + 5z \Rightarrow 3y - 5z = 1 \Rightarrow (y, z) = (7, 4)$$

thus  $x = 2 + 3 \cdot 7 = 23 \equiv 8 \pmod{15 = 3 \times 5}$ .

Next we solve  $x \equiv 8 \pmod{15}$ ,  $x \equiv 2 \pmod{7}$ , that is,

$$x = 8 + 15t = 2 + 7t \Rightarrow 7t - 15s = 6 \Rightarrow (t, s) = 6 \times (13, 6) = (78, 36)$$

thus  $x = 8 + 15 \cdot 36 \equiv 23 \pmod{105 = 15 \times 7}$ .

## Solution of a System in an Elementary Fashion

### Example

We solve the congruency

$$17x \equiv 9 \pmod{276}.$$

Instead of solving it directly, we note that  $276 = 3 \cdot 4 \cdot 23$ , so the congruency is equivalent to the system

$$\begin{aligned} 17x &\equiv 9 \pmod{3}, & 17x &\equiv 9 \pmod{4}, & 17x &\equiv 9 \pmod{23}. \\ x &\equiv 0 \pmod{3}, & x &\equiv 1 \pmod{4}, & 17x &\equiv 9 \pmod{23}. \end{aligned}$$

The first congruence gives  $x = 3k$ ,  $k \in \mathbb{Z}$ . Plugging into the second one,

$$3k \equiv 1 \pmod{4}$$

The modular inverse of  $a = 3$  is  $a^{-1} = 3$ , so we obtain  $k \equiv 3 \pmod{4}$ .

## Solution of a System in an Elementary Fashion

### Example (Cont.)

We then have

$$x = 3 \cdot (3 + 4j) = 9 + 12j, \quad j \in \mathbb{Z}.$$

Inserting into the last congruence,

$$17 \cdot (9 + 12j) \equiv 9 \pmod{23}$$

or

$$204j \equiv -144 \pmod{23}.$$

Hence,  $j = 2 + 23t$ ,  $t \in \mathbb{Z}$  and hence

$$x = 33 + 276t$$

or simply  $x \equiv 33 \pmod{276}$ .

# Euler's Phi Function

## Theorem

$$\varphi(mn) = \varphi(m)\varphi(n) \text{ if } \gcd(m, n) = 1.$$

## Proof.

Recall the isomorphism  $f : \mathbb{Z}/mn\mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ ,  $x \mapsto (x, x)$ .

Similarly consider  $f^\times : (\mathbb{Z}/mn\mathbb{Z})^\times \rightarrow (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$ ,  $x \mapsto (x, x)$ .

Obviously  $f^\times$  is a homomorphism. We want to show that  $f^\times$  is an isomorphism, hence  $|(\mathbb{Z}/mn\mathbb{Z})^\times| = |(\mathbb{Z}/m\mathbb{Z})^\times| \cdot |(\mathbb{Z}/n\mathbb{Z})^\times|$ , i.e.,  $\varphi(mn) = \varphi(m)\varphi(n)$ .

- ▶ Injectivity. Note that  $f^\times$  is  $f$  restricted to the subset  $(\mathbb{Z}/mn\mathbb{Z})^\times \subset \mathbb{Z}/mn\mathbb{Z}$ , then  $f^\times$  is injective since  $f$  is.
- ▶ Surjectivity. Given  $a \in (\mathbb{Z}/m\mathbb{Z})^\times$  and  $b \in (\mathbb{Z}/n\mathbb{Z})^\times$ , by Chinese remainder theorem, we know that there exists  $c \in \mathbb{Z}/mn\mathbb{Z}$  such that  $f(c) = (a, b)$ . We show that  $c \in (\mathbb{Z}/mn\mathbb{Z})^\times$ . Indeed, if  $c = a \in (\mathbb{Z}/m\mathbb{Z})^\times$ , then  $m \in (\mathbb{Z}/c\mathbb{Z})^\times$ . Similarly  $n \in (\mathbb{Z}/c\mathbb{Z})^\times$ , thus  $mn \in (\mathbb{Z}/c\mathbb{Z})^\times$ , so  $c \in (\mathbb{Z}/mn\mathbb{Z})^\times$ . (Be careful with representatives and abuse of notation.)



# Chinese Remainder Theorem

## Corollary

*By fundamental theorem of arithmetic, if  $n = p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots$ , then*

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1^{k_1}\mathbb{Z} \times \mathbb{Z}/p_2^{k_2}\mathbb{Z} \times \mathbb{Z}/p_3^{k_3}\mathbb{Z} \times \cdots$$

and

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/p_1^{k_1}\mathbb{Z})^\times \times (\mathbb{Z}/p_2^{k_2}\mathbb{Z})^\times \times (\mathbb{Z}/p_3^{k_3}\mathbb{Z})^\times \times \cdots$$

## Theorem (Gauss)

*The group  $(\mathbb{Z}/n\mathbb{Z})^\times$  is cyclic if and only if  $n$  is 1, 2, 4,  $p^k$ , or  $2p^k$ , where  $p$  is an odd prime and  $k > 0$ .*

# Table of Contents

1. Prime Numbers
2. Euclidean Algorithm
3. Additive Group of Integers
4. Cyclic Groups and Symmetric Groups
5. Homomorphism and Cosets
6. Modular Arithmetic
7. Chinese Remainder Theorem
8. Public Key Cryptography

# RSA (Rivest–Shamir–Adleman) Cryptography

## Goal

Transfer information from A (Alice) to B (Bob).

## Trapdoor Function

Want to find a (bijective) trapdoor function  $f : S \rightarrow S$ ,  $S$  a HUGE set, such that

- ▶ Easy to compute.
- ▶ HARD to invert.
- ▶ Unless one has the secret key.

## Example (Discrete Logarithm)

Having the inverse of  $e \bmod \varphi(n)$ , the Euler's totient function of  $n$ , is the trapdoor:  $f(x) = x^e \pmod{n}$ .

If the factorization is known,  $\varphi(n)$  can be computed, hence  $e^{-1} \bmod \varphi(n)$  can be computed. Its hardness follows from RSA assumption.

## RSA Example

1. (Alice) Choose 2 (large) distinct primes, e.g.,  $p = 17$ ,  $q = 19$ .
2. (Alice) Let  $n = pq = 17 \times 19 = 323$ .
3. (Alice) Let  $A = \varphi(n) = (p - 1)(q - 1) = 16 \times 18 = 288$ . (Keep private!)
4. (Alice) Pick<sup>4</sup>  $E < \varphi(n)$  such that  $\gcd(E, \varphi(n)) = 1$ , say,  $E = 95$ . Publish public key  $(n, E) = (323, 95)$ , with (public) encryption function  $e$  (for Bob)

$$y = e(x) = x^E \pmod{n}, \quad \text{e.g., } y = e(x) = x^{95} \pmod{323}$$

5. (Alice) Compute private key,  $D = E^{-1} \pmod{A}$ . Then the decryption function  $d$  is given by

$$d(y) = y^D = x^{ED} \equiv x \pmod{n}, \quad \text{e.g., } d(y) = y^{191} \pmod{323}$$

---

4. usually choose  $E = 65537$

# RSA Correctness

## Theorem

Given distinct primes  $p, q$ , let  $n = pq$  and  $ed \equiv 1 \pmod{(p-1)(q-1)}$ . Then if  $x < n$  with  $\gcd(x, n) = 1$ , then  $x^{ed} \equiv x \pmod{n}$ .

## Proof.

Since  $\gcd(x, n) = 1$ , we have  $x^{(p-1)(q-1)} \equiv 1 \pmod{n}$ .

Therefore  $ed = 1 + k(p-1)(q-1)$  for some  $k \in \mathbb{Z}$ , then

$$\begin{aligned}x^{ed} &= x^{1+k(p-1)(q-1)} \\&= x \cdot x^{k(p-1)(q-1)} \\&= x \cdot (x^{(p-1)(q-1)})^k \\&\equiv x \pmod{n}\end{aligned}$$

□

## RSA Correctness

Theorem (Stronger, cf., Gallier, p. 316)

For any two distinct prime numbers  $p$  and  $q$ , if  $e$  and  $d$  are any two positive integers such that

1.  $1 < e, d < (p - 1)(q - 1)$ ,
2.  $ed \equiv 1 \pmod{(p - 1)(q - 1)}$ ,

then for every  $x \in \mathbb{Z}$  we have

$$x^{ed} \equiv x \pmod{pq}$$

### Remark

The proof does NOT rely on Euler's theorem (no coprimeness condition).

# Part III

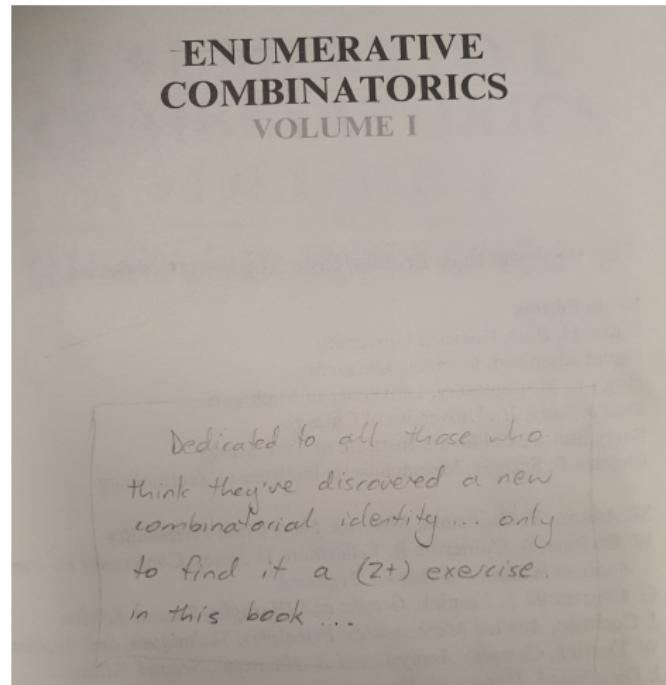
## Counting and Algorithms

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

## Overview

- ▶ The On-Line Encyclopedia of Integer Sequences® (OEIS®)  
<https://oeis.org/>
- ▶ *Enumerative Combinatorics* by Richard P. Stanley



## Twelvefold Way

Distribute  $k$  balls into  $n$  urns. ( $f : B \rightarrow U, |B| = k, |U| = n$ )

| Balls<br>(domain) | Urns<br>(codomain) | unrestricted<br>(any function)   | $\leq 1$<br>(injective) | $\geq 1$<br>(surjective)   |
|-------------------|--------------------|--|-------------------------|--|
| labeled           | labeled            | $n^k$  | $n^k$                   | $n! \left\{ \begin{smallmatrix} k \\ n \end{smallmatrix} \right\}$ |
| unlabeled         | labeled            | $\binom{n}{k}$   | $\binom{n}{k}$          | $\binom{\binom{n}{k}}{k-n}$  |
| labeled           | unlabeled          | $\sum_{i=1}^n \left\{ \begin{smallmatrix} k \\ i \end{smallmatrix} \right\}$ | $[k \leq n]$            | $\left\{ \begin{smallmatrix} k \\ n \end{smallmatrix} \right\}$    |
| unlabeled         | unlabeled          | $\sum_{i=1}^n p_i(k)$  | $[k \leq n]$            | $p_n(k)$   |

- ▶  $n^k = (n)_k = P(n, k) = P_k^n$
- ▶  $\left\{ \begin{smallmatrix} k \\ n \end{smallmatrix} \right\} = \# \text{ partition of } [k] \text{ into } n \text{ parts.}$
- ▶  $\binom{n}{k} = C(n, k) = C_k^n$
- ▶  $\binom{\binom{n}{k}}{k-n} = \binom{n+k-1}{k}$
- ▶  $p_n(k) = \# \text{ partition of } k \text{ into } n \text{ parts.}$
- ▶  $[k \leq n]$ : Iverson bracket
- ▶  $\left\{ \begin{smallmatrix} k \\ n \end{smallmatrix} \right\}$

## Physics Digression

Maxwell-Boltzmann statistics (applicable to no known particles)

- ▶  $k$  distinguishable particles,  $n$  distinguishable cells.
- ▶ Different arrangements with equal probability  $1/n^k$ .

Bose-Einstein statistics (bosons, e.g., photons, nuclei, atoms, spin-1 particles)

- ▶  $k$  **indistinguishable** particles,  $n$  distinguishable cells.
- ▶ Different arrangements with equal probability  $1/\binom{n}{k}$ .

Fermi-Dirac statistics (fermions, e.g., electrons, neutrons, protons, spin- $\frac{1}{2}$  particles)

- ▶  $k$  **indistinguishable** particles,  $n$  distinguishable cells.
- ▶ No two or more particles can be in the same cell.
- ▶ Different arrangements with equal probability  $1/\binom{n}{k}$ .

# Permutations

*k*-permutation of *n*

Number of ways of arranging *k* elements from a set of size *n* (**order matters**) is given by

$$\begin{aligned} n^k &= P(n, k) = P_k^n = \frac{n!}{(n - k)!} \\ &= \underbrace{n \cdot (n - 1) \cdot (n - 2) \cdots (n - k + 1)}_{k \text{ terms}} \end{aligned}$$

which is obviously an integer. Note:  $0! = 1$ .

# Finite Calculus

## Finite Calculus

- $\Delta f(x) = f(x+1) - f(x)$
- If  $g = \Delta G$ , then

$$\begin{aligned}\sum_a^b g(x) \delta x &= G(x) \Big|_a^b \\ &= G(b) - G(a)\end{aligned}$$

## Infinite Calculus

- $Df(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$
- If  $g = DG$ , then

$$\begin{aligned}\int_a^b g(x) dx &= G(x) \Big|_a^b \\ &= G(b) - G(a)\end{aligned}$$

For integers  $b \geq a$ , we should put

$$\sum_a^b g(x) \delta x = \sum_{k=a}^{b-1} g(k) = \sum_{a \leq k < b} g(k)$$

Then for integers  $m, n \geq 0$ , we have

- $\Delta(x^m) = mx^{m-1}$
- $\sum_{0 \leq k < n} k^m = \frac{k^{\frac{m+1}{m+1}}}{m+1} \Big|_0^n = \frac{n^{\frac{m+1}{m+1}}}{m+1}$

# Finite Calculus

## Example

►  $\sum_{0 \leq k < n} k = \frac{n^2}{2} = \frac{n(n - 1)}{2}$ .

► Note that  $k^2 = k^{\underline{2}} + k^{\underline{1}}$ , then

$$\sum_{0 \leq k < n} k^2 = \frac{n^3}{3} + \frac{n^2}{2} = \frac{1}{3}n(n - 1)(n - 2 + \frac{3}{2}) = \frac{1}{3}n(n - \frac{1}{2})(n - 1)$$

► Note that  $k^3 = k^{\underline{3}} + 3k^{\underline{2}} + k^{\underline{1}}$ , then

$$\sum_{a \leq k < b} k^3 = \frac{n^4}{4} + n^3 + \frac{n^2}{2} \Big|_a^b$$

# Stirling Number of the Second Kind

We can always express ordinary powers using factorial powers via Stirling numbers, i.e., for  $n \geq 0$ ,

$$x^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^k$$

## Definition

For  $n, k \in \mathbb{N}$ ,  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$  is the number of ways to partition a set with  $n$  elements into  $k$  disjoint, nonempty subsets.  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$  is called a **Stirling number of the second kind**. Reads “ $n$  subset  $k$ ”.

## Remark

The power conversion above can be proved by induction based on the following recurrence relation

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\} + k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\}$$

# Partition of A Set and Partition of A Number

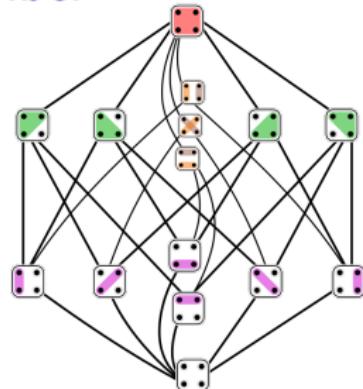
## Partition of A Set

►  $\binom{4}{3} = 6.$

12|3|4, 13|2|4, 14|2|3, 23|1|4, 24|1|3, 34|1|2.

►  $\binom{4}{2} = 7.$

123|4, 124|3, 134|2, 234|1, 12|34, 13|24, 14|23.



## Partition of An Integer

A partition of the number  $k$  is a tuple of **positive** integers

$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$  such that  $\sum_{i=1}^n \lambda_i = k$  and  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . We use  $p(k)$  for the number of partitions of  $k$ , and  $p_n(k)$  for the number of partitions of  $k$  into exactly  $n$  parts. Note that  $p(k) = \sum_{i=1}^k p_i(k)$ .

■  $p_2(7) = 3.$

$$7 = 6 + 1 = 5 + 2 = 4 + 3.$$

■  $p_3(7) = 4.$

$$\begin{aligned} 7 &= 5 + 1 + 1 = 4 + 2 + 1 \\ &= 3 + 3 + 1 = 3 + 2 + 2. \end{aligned}$$

## Particular Values of Stirling Number of the Second Kind

- $\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = 0$  if  $k > n$ .
- $\left\{ \begin{matrix} n \\ 0 \end{matrix} \right\} = 0$ .
- $\left\{ \begin{matrix} n \\ 1 \end{matrix} \right\} = 1$ .
- $\left\{ \begin{matrix} n \\ n \end{matrix} \right\} = 1$ .
- $\left\{ \begin{matrix} n \\ n-1 \end{matrix} \right\} = \binom{n}{2}$ .
- $\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 2^{n-1} - 1$ .

| $n$ | $\left\{ \begin{matrix} n \\ 0 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 1 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 3 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 4 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 5 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 6 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 7 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 8 \end{matrix} \right\}$ | $\left\{ \begin{matrix} n \\ 9 \end{matrix} \right\}$ |
|-----|---|---|---|---|---|---|---|---|---|---|
| 0   | 1   |   |   |   |   |   |   |   |   |   |
| 1   | 0   | 1   |   |   |   |   |   |   |   |   |
| 2   | 0   | 1   | 1   |   |   |   |   |   |   |   |
| 3   | 0   | 1   | 3   | 1   |   |   |   |   |   |   |
| 4   | 0   | 1   | 7   | 6   | 1   |   |   |   |   |   |
| 5   | 0   | 1   | 15  | 25  | 10  | 1   |   |   |   |   |
| 6   | 0   | 1   | 31  | 90  | 65  | 15  | 1   |   |   |   |
| 7   | 0   | 1   | 63  | 301   | 350   | 140   | 21  | 1   |   |   |
| 8   | 0   | 1   | 127   | 966   | 1701  | 1050  | 266   | 28  | 1   |   |
| 9   | 0   | 1   | 255   | 3025  | 7770  | 6951  | 2646  | 462   | 36  | 1   |

By inclusion-exclusion principle (on this later),

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \frac{1}{k!} \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} j^n$$

# Stirling Number of the Second Kind

## Theorem

For  $n \geq k \geq 1$ ,

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\} + k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\}$$

## Proof by double counting.

- ▶ By definition, the LHS counts the ways of partitioning of an  $n$ -element set into  $k$  subsets.
- ▶ Consider whether the element  $n$  is alone in its own set.
  - ▶ If yes, then there are  $\left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\}$  ways of partition the remaining  $n-1$  elements into  $k-1$  subsets.
  - ▶ If no, then there are  $\left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\}$  ways of partition the remaining  $n-1$  elements into  $k$  subsets, and there are  $k$  choices to insert the element  $n$  into any of the  $k$  subsets.



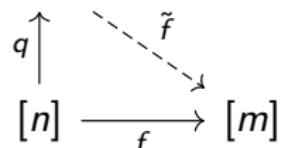
# Stirling Number of the Second Kind

## Theorem

For  $m, n \geq 0$ ,

$$m^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} m^k$$

$$[k] \cong [n]/\sim$$



Proof by double counting (Method I).

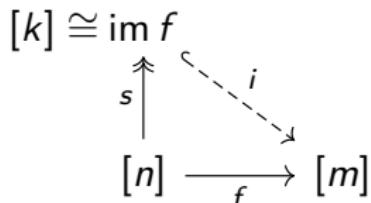
- ▶ By definition, the LHS counts the number of functions from  $[n] = \{1, \dots, n\}$  to  $[m] = \{1, \dots, m\}$ .
- ▶ Consider the partition of the domain induced by the function  $f : [n] \rightarrow [m]$ , the size of the partition ranges from 0 to  $n$ . Now for each fixed partition of size  $k$ , the induced function  $\tilde{f} : [n]/\sim \rightarrow [m]$  is injective. There are  $m^k$  such injections, with  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$  choices of different domains of size  $k$  (Note that  $q$  is surjective). □

# Stirling Number of the Second Kind

## Theorem

For  $m, n \geq 0$ ,

$$m^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} m^k$$



Proof by double counting (Method II).

- ▶ By definition, the LHS counts the number of functions from  $[n] = \{1, \dots, n\}$  to  $[m] = \{1, \dots, m\}$ .
- ▶ For each function  $f : [n] \rightarrow [m]$ , we can write  $f = i \circ s$ , where  $s : [n] \twoheadrightarrow \text{im } f$  is surjective, and  $i : \text{im } f \hookrightarrow [m]$  is injective (called an inclusion map). For each  $\text{im } f$  of fixed size  $k$ , the number of surjection  $s$  is given by  $k! \left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ , and the choice of  $\text{im } f \subset [m]$  is given by  $\binom{m}{k} = m^k/k!$ . □

# Combinations

## Definition

The number of ways to choose  $k$  elements from a set of  $n$  (**order does NOT matter**) is denoted

$$C(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{P(n, k)}{k!} = \frac{n^k}{k!}$$

Reads “ $n$  choose  $k$ ”.

## Basic Properties

- ▶  $\binom{n}{k} = \binom{n}{n-k}$
- ▶  $\binom{n}{0} = \binom{n}{n} = 1$
- ▶  $\binom{n}{1} = n$

# Combinations

Why is  $\binom{n}{k}$  an integer?

## Method I: Counting Prime Factors.

Note that  $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ , it is sufficient to show that all factors of the denominator are cancelled by factors in the numerator. Let  $\mu_p(x)$  be the number of the prime factor  $p$  in  $x$ . According to Legendre's theorem, for  $N \in \mathbb{N}$ , we have

$$\mu_p(N!) = \sum_{k \geq 1} \left\lfloor \frac{N}{p^k} \right\rfloor$$

We want to show that  $\mu_p(n!) \geq \mu_p(k!) + \mu_p((n-k)!)$  for all  $p \in \mathbb{P}$ .  
Indeed,

$$\mu_p(n!) - \mu_p(k!) - \mu_p((n-k)!) = \sum_{k \geq 1} \left( \left\lfloor \frac{n}{p^k} \right\rfloor - \left\lfloor \frac{k}{p^k} \right\rfloor - \left\lfloor \frac{n-k}{p^k} \right\rfloor \right)$$

The rest follows by noticing that for  $x, y \in \mathbb{R}$ , either  $\lfloor x+y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$ , or  $\lfloor x+y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor + 1$ .



## Combinations

### Method II: Using Lagrange's Theorem.

Consider the symmetric group  $G = S_n$ , then  $|G| = n!$ . Next consider the subgroup  $H \leq G$  given by

$$H := \left\{ f \in S_n \mid \begin{array}{l} f(\{1, \dots, k\}) = \{1, \dots, k\} \\ f(\{k+1, \dots, n\}) = \{k+1, \dots, n\} \end{array} \right\}$$

then  $H \cong S_k \times S_{n-k}$ , and  $|H| = k!(n-k)!$ . It follows by Lagrange's theorem that  $|H|$  divides  $|G|$ . □

### Method III: Using Induction.

The induction procedure follows by the recursive identity

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

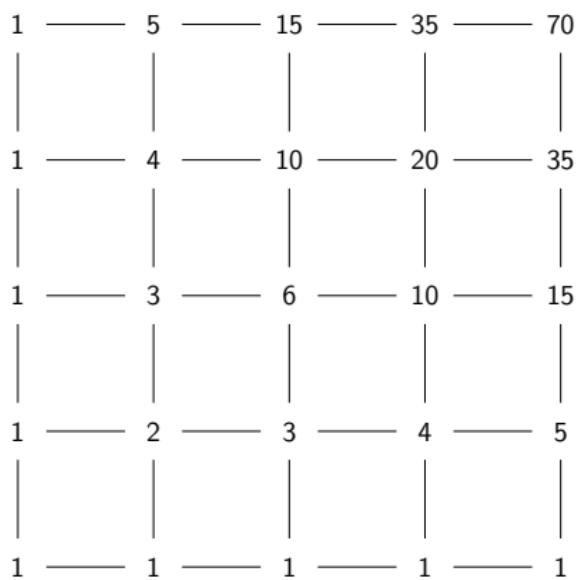


# Recursive Identity for Binomial Coefficients

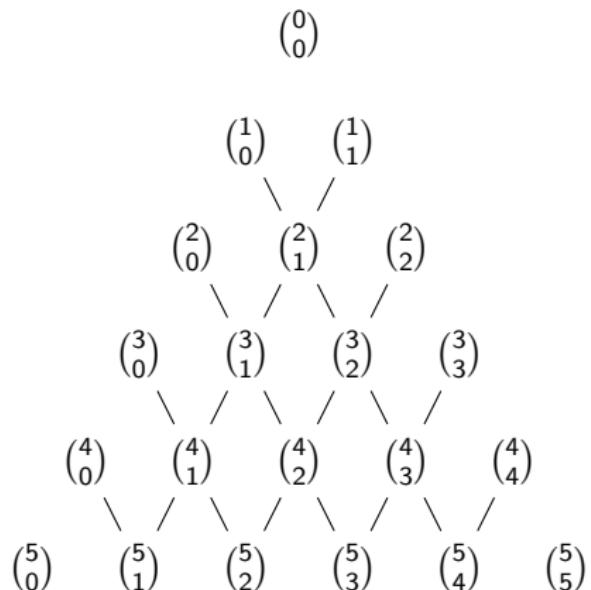
## Theorem

For all  $n > 0$  and  $0 < k < n$ ,  $k, n \in \mathbb{N}$ ,  $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ .

## Lattice Paths

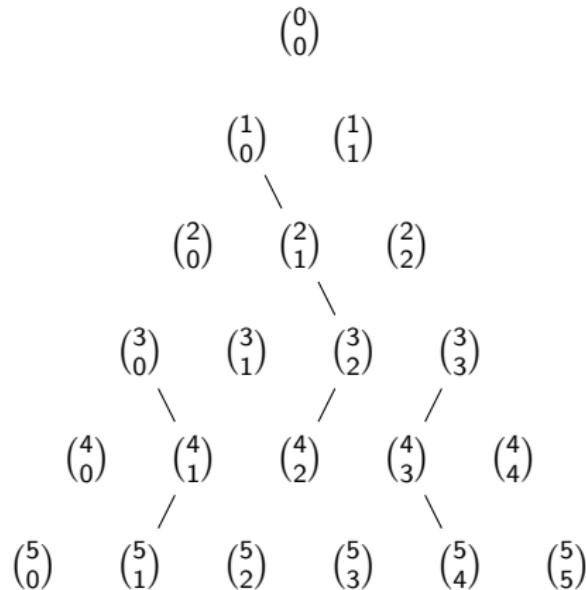


## Pascal Triangle



## Recursive Identity for Binomial Coefficients

$$\sum_{k=0}^n \binom{r+k}{k} = \binom{r}{0} + \binom{r+1}{1} + \cdots + \binom{r+n}{n} = \binom{r+n+1}{n}$$

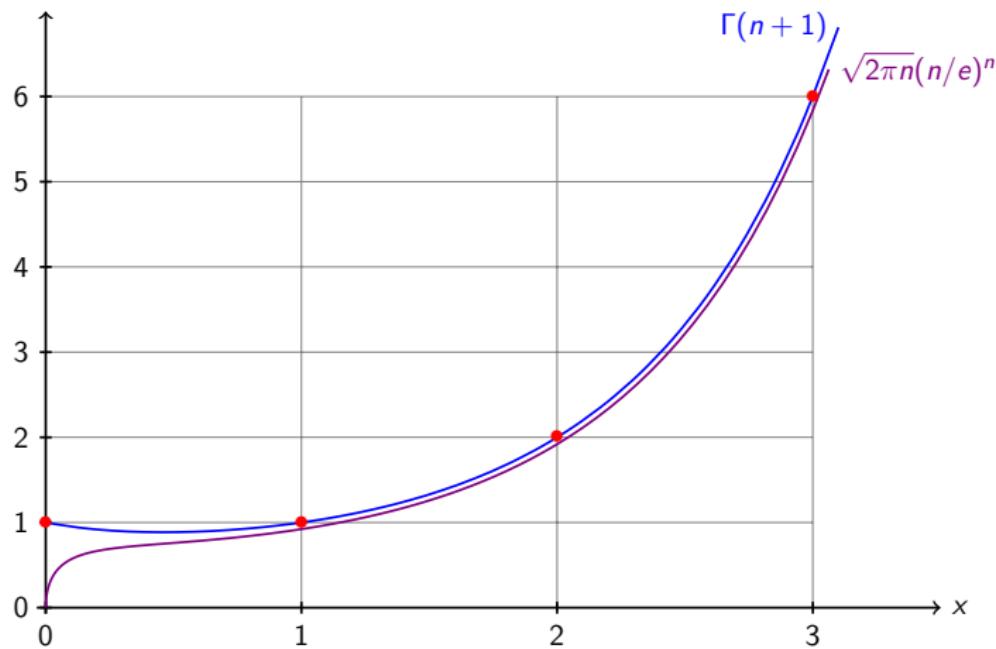


# Stirling Approximation

How to calculate  $\binom{n}{k}$  if  $n, k$  are really large?

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

It works extremely good even for small integers.

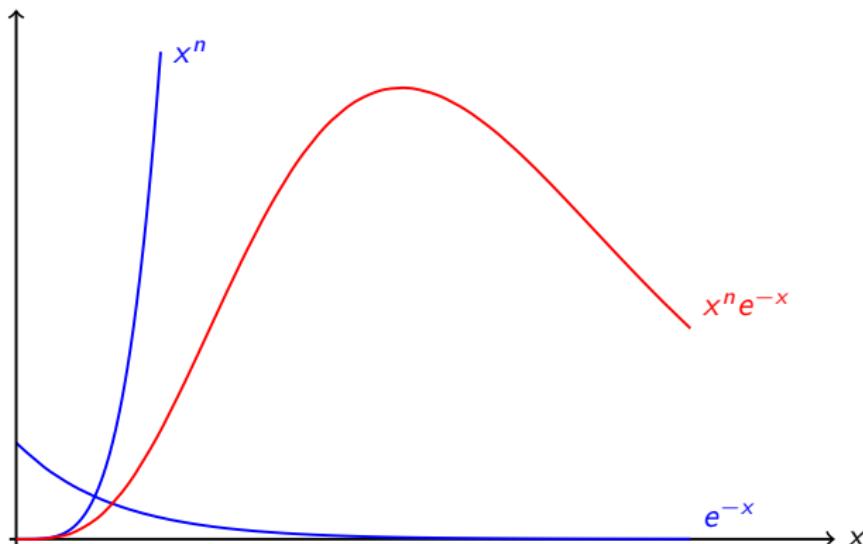


# Stirling Approximation

The continuous version of the factorial is given by the Gamma function

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

we can approximate the integrand using a bell-shaped curve (Gaussian normal distribution).



## Stirling Approximation

Consider the Gamma integral representation of factorial, then

$$n! = \int_0^\infty x^n e^{-x} dx = \int_0^\infty e^{-(x-n \ln x)} dx$$

Let the exponent be  $f(x) := x - n \ln x$ , which we will approximate using a quadratic function. This is also known as Laplace's method. Since

$$f'(x) = 1 - \frac{n}{x} = 0 \Rightarrow x = n \quad \text{and} \quad f''(x) = \frac{n}{x^2} \Big|_{x=n} = \frac{1}{n}$$

thus  $-f$  attains minimum at  $x = n$ , thus by Taylor expansion, approximately near  $x = n$ ,

$$f(x) \sim n - n \ln n + \frac{1}{2n}(x - n)^2$$

therefore

$$\begin{aligned} \int_0^\infty e^{-(x-n \ln x)} dx &\sim \int_0^\infty e^{-(n-n \ln n)} e^{-(x-n)^2/(2n)} dx \\ &\sim e^{-(n-n \ln n)} \int_{\mathbb{R}} e^{-(x-n)^2/(2n)} dx = n^n e^{-n} \sqrt{2\pi n} \end{aligned}$$

# Stirling Approximation

By stirling approximation, we have

$$\binom{n}{k} \sim \sqrt{\frac{n}{2\pi k(n-k)}} \cdot \frac{n^n}{k^k(n-k)^{n-k}}$$

In particular,  $\binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}}$ .

## Example

Consider flipping  $2n$  fair coins, what is the probability that exactly half are heads and the other half are tails?

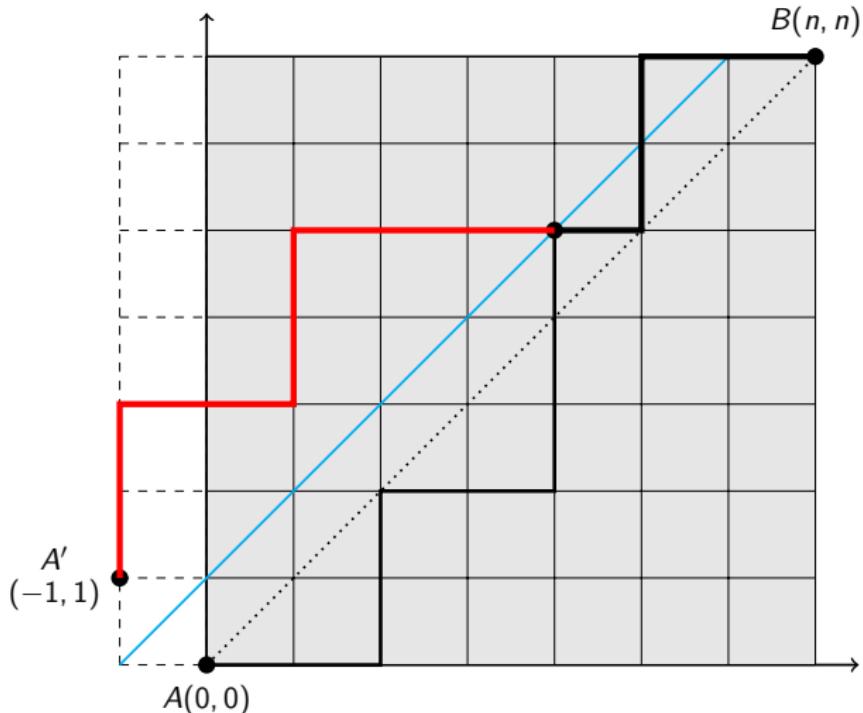
Let  $X$  be a random variable following a binomial distribution, i.e.,  $X \sim \text{Binomial}(2n, \frac{1}{2})$ , then

$$\Pr(X = n) = \binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \sim \frac{1}{\sqrt{\pi n}}$$

If  $2n = 100$ , then probability is approximately  $\frac{1}{\sqrt{50\pi}} = \frac{1}{5\sqrt{2\pi}} \approx 0.08$ . Note that  $\sqrt{2\pi} \approx e \approx 2.5$ , (recall  $1! \sim \sqrt{2\pi \cdot 1} (1/e)^1$ ).

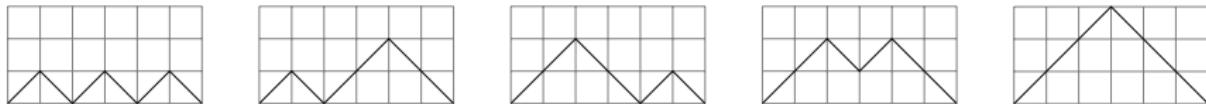
## Catalan Numbers

How many lattice paths from  $(0, 0)$  to  $(n, n)$  that never go above the diagonal?  $C_n = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n}$ .

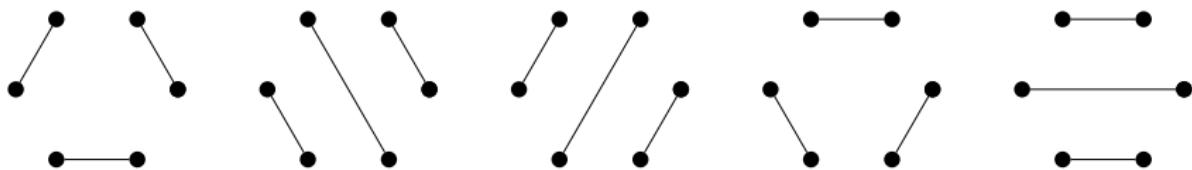


# Catalan Numbers

## Mountain Range/Dyck Paths



## Noncrossing Handshakes

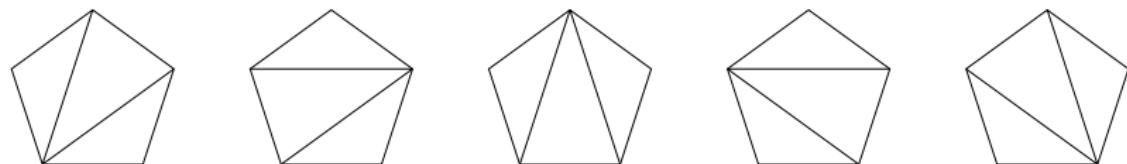


## Paired Parentheses

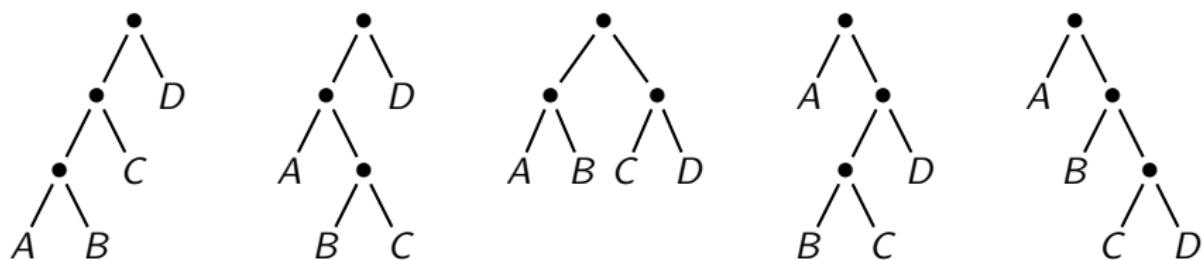
- ()()
  - (( ))()
  - ()(())
  - (( ))()
  - ((( )))

# Catalan Numbers

## Polygon Triangulation



## Full Binary Trees



## Matrix Chain Multiplication

- $((AB)C)D$
- $(A(BC))D$
- $(AB)(CD)$
- $A((BC)D)$
- $A(B(CD))$

## Catalan Numbers

### Segner's recurrence relation

We can establish the following recurrence relation starting with  $C_0 = 1$ , and

$$C_n = \sum_{k=0}^n C_k C_{n-k} \text{ for } n \geq 0,$$

We recognize the RHS is a convolution. Now consider the following generating function

$$c(x) := \sum_{n=0}^{\infty} C_n x^n$$

then  $c(x) = 1 + xc(x)^2$ , and

$$c(x) = \frac{1 \pm \sqrt{1 - 4x}}{2x} = \frac{1 \pm (1 - 2x - 2x^2 + \dots)}{2x}$$

We want the solution to be a (formal) power series, take the minus sign.

# Catalan Numbers

Segner's recurrence relation (Cont.)

Note that

$$c(x) = \frac{1 - \sqrt{1 - 4x}}{2x}$$

which we expand as

$$\begin{aligned} c(x) &= \frac{1}{2x}(1 - \sqrt{1 - 4x}) = \frac{1}{2x} \cdot 2 \sum_{n=1}^{\infty} \frac{(-1)^n}{4^n} \binom{2n-2}{n-1} \frac{(-4x)^n}{n} \\ &= \sum_{n=1}^{\infty} \binom{2n-2}{n-1} \frac{x^{n-1}}{n} \\ &= \sum_{n=0}^{\infty} \binom{2n}{n} \frac{x^n}{n+1} = \sum_{n=0}^{\infty} C_n x^n. \end{aligned}$$

## Top 10 Binomial Coefficient Identities

factorial expansion

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

integers,  
 $n \geq k \geq 0$ .

symmetry

$$\binom{n}{k} = \binom{n}{n-k}$$

integer  $n \geq 0$ ,  
integer  $k$ .

absorption/extraction

$$\binom{r}{k} = \frac{r}{k} \binom{r-1}{k-1}$$

integer  $k \neq 0$ .

addition/induction

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}$$

integer  $k$ .

upper negation

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}$$

integer  $k$ .

## Top 10 Binomial Coefficient Identities (Cont.)

trinomial  
revision

$$\binom{r}{m} \binom{m}{k} = \binom{r}{k} \binom{r-k}{m-k}$$

integers  $m, k$ .

binomial  
theorem

$$\sum_k \binom{r}{k} x^k y^{r-k} = (x+y)^r$$

integers  $r \geq 0$ ,  
or  $|x/y| < 1$ .

parallel  
summation

$$\sum_{k \leq n} \binom{r+k}{k} = \binom{r+n+1}{n}$$

integer  $n$ .

upper  
summation

$$\sum_{0 \leq k \leq n} \binom{k}{m} = \binom{n+1}{m+1}$$

integers  
 $m, n \geq 0$ .

Vandermonde  
convolution

$$\sum_k \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}$$

integer  $n$ .

# Multinomial Coefficients

## Multinomial Coefficients

For all  $n, m, k_1, \dots, k_m \in \mathbb{N}$ , with  $k_1 + \dots + k_m = n$  and  $m \geq 2$ , we have the **multinomial coefficients**

$$\binom{n}{k_1, k_2, \dots, k_m} = \frac{n!}{k_1! k_2! \cdots k_m!} = \frac{(\sum_{i=1}^m k_i)!}{\prod_{i=1}^m k_i!}$$

which counts the number of ways of splitting a set of  $n$  elements into an **ordered** sequence of  $m$  disjoint subsets. Relating to binomial coefficients

$$\binom{n}{k_1, k_2, \dots, k_m} = \binom{n}{k_1} \binom{n - k_1}{k_2} \binom{n - k_1 - k_2}{k_3} \cdots \binom{n - \sum_{i=1}^{m-1} k_i}{k_m}$$

## Example

Count distinct permutations of the word MISSISSIPPI.

$$\frac{11!}{1!4!4!2!}$$

## Multinomial Formula

Theorem (Gallier, Prop. 4.10, p.216-7)

For all  $m, n \in \mathbb{N}$  with  $m \geq 2$ , for all pairwise commuting variables  $a_1, \dots, a_m$ , we have

$$(a_1 + \cdots + a_m)^n = \sum_{\substack{k_1, \dots, k_m \geq 0 \\ k_1 + \cdots + k_m = n}} \binom{n}{k_1, k_2, \dots, k_m} a_1^{k_1} a_2^{k_2} \cdots a_m^{k_m}$$

Question: How many terms occur on the RHS of the multinomial formula?

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

# Notation

## Definition

Let  $\binom{n}{k}$  be the number of  $k$ -element multisets on an  $n$ -element set. Reads “ $n$  multichoose  $k$ ”.

## Remark

If  $k > n \geq 0$ ,  $n, k \in \mathbb{N}$

- ▶  $\binom{n}{k} = 0$ . (pigeonhole principle.)
- ▶  $\binom{n}{k} \neq 0$ .

## Remark

$\binom{n}{k}$  counts the ways to select  $k$  objects from a set of  $n$  elements, where order is not important, but repetition is allowed.

# Counting Multisets

## Proposition

The number of  $k$ -element multisets on an  $n$ -element set is

$$\binom{n}{k} = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

Proof by double counting.

- ▶ Definition.
- ▶ We want to divide  $k$  identical **stars** by  $n - 1$  **bars**. We arrange everything in  $n + k - 1$  positions. Choose  $k$  positions for the stars, and the rest for bars, or vice versa. □

Example: 10-element multisets from a 4-element set

|   |   |   |   |   |   |   |   |   |    |    |    |    |   |
|---|---|---|---|---|---|---|---|---|----|----|----|----|---|
| * | * | * |   | * | * |   |   | * | *  |    | *  | *  | * |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |   |

# A Multisets Identity

## Theorem

Given  $n, k \geq 1, n, k \in \mathbb{N}$ ,

$$\binom{n}{k} = \binom{k+1}{n-1}$$

Proof by double counting.

- ▶ Number of ways to arrange  $k$  stars and  $n - 1$  bars.
- ▶ Number of ways to arrange  $n - 1$  stars and  $k$  bars.

The above two maps to each other by switching bars and stars.

□



## What Multiset Counts

### Example

The quantity  $\binom{n}{k}$  counts,

- ▶ the number of ways to put  $k$  identical balls into urns  $B_1, \dots, B_n$ .
- ▶ the number of ways to distribute  $k$  candy bars to  $n$  people.
- ▶ the number of ways to buy  $k$  drinks from a vending machine with  $n$  varieties.
- ▶ The number of nonnegative integer solutions to  $x_1 + x_2 + \dots + x_n = k$ .
- ▶ The number of positive integer sequences  $a_1, a_2, \dots, a_k$  where  $1 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq n$ .

# Counting Integer Solutions

## Example

Consider

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \leq 538$$

What are the number of integer solutions if

1.  $x_i > 0$  and = holds;
2.  $x_i \geq 0$  and = holds;
3.  $x_i > 0$  and < holds;
4.  $x_i \geq 0$  and < holds;
5.  $x_i \geq 0$ .

## Remark

- $x_i > 0 \Rightarrow x_i \geq 1$ .

# Counting Integer Solutions

## Example

How many nonnegative integer solutions to the equation

$$x_1 + x_2 + x_3 + x_4 = 63$$

such that  $x_1, x_2 \geq 0$ ,  $2 \leq x_3 \leq 5$ ,  $x_4 > 0$ .

Consider the following solution sets, where  $A \supset B$ .      Answer:  $|A| - |B|$ .

- ▶  $A$ : such that  $x_1, x_2 \geq 0$ ,  $x_3 \geq 2$ ,  $x_4 > 0$ , i.e.,  $x_3 - 2 \geq 0$ ,  $x_4 - 1 \geq 0$ , and

$$x_1 + x_2 + (x_3 - 2) + (x_4 - 1) = 60$$

We have  $|A| = \binom{60+3}{3}$ .

- ▶  $B$ : such that  $x_1, x_2 \geq 0$ ,  $x_3 > 5$ ,  $x_4 > 0$ , i.e.,  $x_3 - 6 \geq 0$ ,  $x_4 - 1 \geq 0$ , and

$$x_1 + x_2 + (x_3 - 6) + (x_4 - 1) = 56$$

We have  $|B| = \binom{56+3}{3}$ .

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

# Inclusion-Exclusion Principle

## Example

► 2 sets:  $|A \cup B| = |A| + |B| - |A \cap B|$

► 3 sets:  $|A \cup B \cup C| = |A| + |B| + |C|$

$$\begin{aligned} & - |A \cap B| - |B \cap C| - |C \cap A| \\ & + |A \cap B \cap C| \end{aligned}$$



A 10x10 grid titled "Sieve of Eratosthenes". The first column contains a small illustration of a sieve. The numbers 2 through 100 are listed in a 10x10 grid. Prime numbers are circled in red. Composite numbers are colored according to their smallest prime factor: 2 (blue), 3 (orange), 4 (green), 5 (red), 6 (blue), 7 (orange), 8 (green), 9 (red), 10 (blue), 11 (red), 12 (green), 13 (blue), 14 (orange), 15 (red), 16 (green), 17 (orange), 18 (green), 19 (red), 20 (blue), 21 (red), 22 (green), 23 (blue), 24 (orange), 25 (red), 26 (green), 27 (red), 28 (orange), 29 (red), 30 (blue), 31 (red), 32 (green), 33 (red), 34 (green), 35 (blue), 36 (orange), 37 (red), 38 (green), 39 (red), 40 (blue), 41 (red), 42 (orange), 43 (blue), 44 (green), 45 (red), 46 (green), 47 (blue), 48 (orange), 49 (red), 50 (blue), 51 (red), 52 (green), 53 (blue), 54 (orange), 55 (red), 56 (blue), 57 (red), 58 (green), 59 (blue), 60 (red), 61 (red), 62 (green), 63 (orange), 64 (green), 65 (blue), 66 (red), 67 (blue), 68 (green), 69 (red), 70 (blue), 71 (red), 72 (green), 73 (blue), 74 (red), 75 (orange), 76 (green), 77 (red), 78 (orange), 79 (red), 80 (blue), 81 (red), 82 (green), 83 (blue), 84 (orange), 85 (red), 86 (green), 87 (red), 88 (green), 89 (blue), 90 (red), 91 (blue), 92 (green), 93 (red), 94 (green), 95 (blue), 96 (red), 97 (blue), 98 (orange), 99 (red), 100 (blue).

## Applications

- Sieve of Eratosthenes
- Euler's totient function
- ...

# Inclusion-Exclusion Principle

## Notation

Given  $I \subset \{1, \dots, n\}$ , we let

$$A_I := \bigcap_{i \in I} A_i,$$

where  $A_i \in X$  for all  $i \in I$ . For example,  $A_{\{1,2,4\}} = A_1 \cap A_2 \cap A_4$ . In particular,  $A_\emptyset = X$ .

## Theorem (Inclusion-Exclusion Principle)

Let  $A_1, \dots, A_n$  be subsets of  $X$ . Then the number of elements of  $X$  which lie in none of the subsets  $A_i$  is

$$\sum_{I \subset \{1, \dots, n\}} (-1)^{|I|} |A_I|$$

# Inclusion-Exclusion Principle

Proof (Not by induction).

Re-write the sum as

$$\sum_{I \subset [n]} (-1)^{|I|} |A_I| = \sum_{I \subset [n]} \sum_{x \in A_I} (-1)^{|I|} = \sum_x \sum_{I: x \in A_I} (-1)^{|I|}$$

- ▶ If  $x \in X \setminus \cup_{i=1}^n A_i$ , then  $I = \emptyset$ . Since now  $x$  is fixed, thus  $\sum_{x \in A_\emptyset} (-1)^{|\emptyset|} = \sum_{x \in X} (-1)^0 = 1$ .
- ▶ Otherwise, the set  $J = \{i \mid x \in A_i\} \neq \emptyset$ . Note that  $x \in A_I$  iff  $I \subset J$ , thus

$$\sum_{I \subset J} (-1)^{|I|} = \sum_{i=0}^{|J|} \binom{|J|}{i} (-1)^i = (1 - 1)^{|J|} = 0$$

Sum the terms and we are done.



# Inclusion-Exclusion Principle

## Corollary

Let  $A_1, \dots, A_n$  be a sequence of (not necessarily distinct) sets, then

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{\emptyset \neq I \subset \{1, \dots, n\}} (-1)^{|I|+1} |A_I|.$$

## Proof.

Take the complement of both sides of previous theorem within the set  $X = A_\emptyset$ , that is,

$$\begin{aligned} |A_1 \cup \dots \cup A_n| &= |A_\emptyset| - \sum_{I \subset \{1, \dots, n\}} (-1)^{|I|+1} |A_I| \\ &= \sum_{\emptyset \neq I \subset \{1, \dots, n\}} (-1)^{|I|+1} |A_I| \end{aligned}$$

□

# Inclusion-Exclusion Principle

## Special Case

The formula is a lot simpler when

$$|I| = |J| \Rightarrow |A_I| = |A_J|,$$

that is,  $|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|$  depends only on  $k$ , where  $I = \{i_1, i_2, \dots, i_k\}$ .

Now the formula becomes

$$|A_1 \cup \cdots \cup A_n| = \sum_{|I|=1}^n (-1)^{|I|+1} \binom{n}{|I|} |A_I|$$

# Derangement

## Definition

A permutation  $\sigma \in S_n$  over the set  $\{1, 2, \dots, n\}$  is called a derangement if  $\sigma(i) \neq i$  for all  $i = 1, \dots, n$ .

## Theorem

The number of derangements of the set  $\{1, 2, \dots, n\}$  is given by

$$d_n = \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)! = n! \sum_{i=0}^n \frac{(-1)^i}{i!}$$

## Proof.

Take  $A_i := \{\sigma \in S_n \mid \sigma(i) = i\}$ , thus  $|A_i| = (n-1)!$ . Note that for general set  $I \subset \{1, 2, \dots, n\}$ ,  $|A_I| = (n - |I|)!$ . The rest follows by inclusion-exclusion principle and

$$\binom{n}{i} (n-i)! = \frac{n!}{i!(n-i)!} (n-i)! = \frac{n!}{i!}$$

□

# Derangement

## Asymptotics

Assume that each  $\sigma \in S_n$  happens equally likely, what is the probability that  $\sigma$  is a derangement?

Note that  $|S_n| = n!$ , thus

$$\lim_{n \rightarrow \infty} \frac{d_n}{n!} = \lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{(-1)^i}{i!} = e^{-1} = \frac{1}{e} \approx \frac{1}{3}$$

# Counting Surjections

## Theorem

Let  $k \geq n$ . The number of surjections  $f : \{1, \dots, k\} \rightarrow \{1, \dots, n\}$  is given by

$$S_{k,n} = \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} (n-i)^k$$

## Proof.

Take  $A_i = \{f \mid f(j) \neq i \text{ for all } j\} = \{f \mid i \notin \text{im } f\}$ .

□

# Counting Surjections

## Example

What is  $S_{5,3} = |\{f : \{1, 2, 3, 4, 5\} \rightarrow \{1, 2, 3\} \mid f \text{ surjective}\}|$ ?

## Method I.

$$S_{5,3} = \binom{3}{0}(3-0)^5 - \binom{3}{1}(3-1)^5 + \binom{3}{2}(3-2)^5 - \underbrace{\binom{3}{3}(3-3)^5}_{=0}$$

□

## Method II.

We first calculate that

$$\left\{ \begin{matrix} 5 \\ 3 \end{matrix} \right\} = \binom{5}{3} + 3 \binom{5}{4} = 25$$

thus  $S_{5,3} = 3! \left\{ \begin{matrix} 5 \\ 3 \end{matrix} \right\} = 150$ .

□

## Dimension of Vector Spaces

Given finite dimensional vector spaces  $U$ ,  $V$ , and  $W$ , are the following identities correct?

- ▶  $\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$
- ▶ 
$$\begin{aligned} \dim(U + V + W) \\ = \dim U + \dim V + \dim W \\ - \dim(U \cap V) - \dim(U \cap W) - \dim(V \cap W) \\ + \dim(U \cap V \cap W). \end{aligned}$$

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

# Matrix Chain Multiplication

## Multiplying Two Matrices

- Given  $A \in \mathbb{R}^{p \times q}$  and  $B \in \mathbb{R}^{q \times r}$ , then  $C = AB \in \mathbb{R}^{p \times r}$ , with

$$C_{ij} = \sum_{k=1}^q A_{ik} B_{kj}$$

- Cost:  $pqr$  scalar multiplications. (Additions does not cost as much)

$$\begin{bmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1r} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ir} \\ \vdots & & \vdots & & \vdots \\ c_{p1} & \cdots & c_{pj} & \cdots & c_{pr} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1q} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{iq} \\ \vdots & & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pj} & \cdots & a_{pq} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1r} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kj} & \cdots & b_{kr} \\ \vdots & & \vdots & & \vdots \\ b_{q1} & \cdots & b_{qj} & \cdots & b_{qr} \end{bmatrix}$$

# Matrix Chain Multiplication

## Multiplying More Than Two Matrices

Given  $A_1 \in \mathbb{R}^{p_0 \times p_1}$ ,  $A_2 \in \mathbb{R}^{p_1 \times p_2}$ ,  $A_3 \in \mathbb{R}^{p_2 \times p_3}$ , by associativity of matrix multiplication

$$(A_1 A_2) A_3 = A_1 (A_2 A_3)$$

but with different cost of scalar multiplications.

### Example

Let  $A_1 \in \mathbb{R}^{10 \times 5}$ ,  $A_2 \in \mathbb{R}^{5 \times 10}$ ,  $A_3 \in \mathbb{R}^{10 \times 5}$ , i.e.,

$(p_0, p_1, p_2, p_3) = (10, 5, 10, 5)$ .

- ▶ cost of  $(A_1 A_2) A_3 = 10 \cdot 5 \cdot 10 + 10 \cdot 10 \cdot 5 = 1000$
- ▶ cost of  $A_1 (A_2 A_3) = 5 \cdot 10 \cdot 5 + 10 \cdot 5 \cdot 5 = 500$

# Matrix Chain Multiplication

## Problem Statement

- ▶ **Input:** A sequence (chain) of matrices  $(A_1, A_2, \dots, A_n)$ , where  $A_i \in \mathbb{R}^{p_{i-1} \times p_i}$ . Or equivalently, a vector of matrix dimensions  $p = (p_0, p_1, \dots, p_n)$ .
- ▶ **Output:** Full parenthesization (ordering) for the product  $A_1 \times A_2 \times \dots \times A_n$  that minimizes the number of (scalar) multiplications.

## Counting the Number of Orderings

- ▶ Let  $P(n)$  be the number of orderings for a chain of  $n$  matrices.
- ▶ Then  $P(1) = 1$  and for  $n \geq 2$ ,  $P(n) = C_{n-1}$ , the Catalan number,

$$\begin{aligned} P(n) &= P(1)P(n-1) + P(2)P(n-2) + \dots + P(n-1)P(1) \\ &= \sum_{k=1}^{n-1} P(k)P(n-k) \end{aligned}$$

# Matrix Chain Multiplication

## Structure of An Optimal Ordering

- ▶ An optimal ordering of the product  $A_1 A_2 \cdots A_n$  splits the product between  $A_k$  and  $A_{k+1}$  for some  $k$ :

$$A_1 A_2 \cdots A_n = (\textcolor{blue}{A_1} \cdots \textcolor{red}{A_k}) \cdot (\textcolor{red}{A_{k+1}} \cdots A_n)$$

- ▶ The (“local”) ordering of  $\textcolor{blue}{A_1} \cdots \textcolor{blue}{A_k}$  within this (“global”) optimal ordering must be an optimal ordering of (the sub-product)  $\textcolor{blue}{A_1} \cdots \textcolor{blue}{A_k}$ .
- ▶ Same for  $A_{k+1} \cdots A_n$ .

## Optimal Substructure

A problem is said to have **optimal substructure** if an optimal solution can be constructed from optimal solutions of its subproblems. This property is used to determine the usefulness of **dynamic programming** and **greedy algorithms** for a problem. (Cormen, Leiserson, Rivest, and Stein, Introduction to Algorithms, or CLRS for short.)

# Matrix Chain Multiplication

## A Recursive Solution (Top-Down)

Let  $m[i, j]$  be the minimum number of scalar multiplications needed to compute the matrix  $A_{i..j}$ ; for the full problem, the lowest-cost way to compute  $A_{1..n}$  is thus given by  $m[1, n]$ .

- ▶ Splitting  $A_i \cdots A_j$  into  $A_{i..k} \cdot A_{k+1..j}$  for some  $k$ , where  $i \leq k < j$ , yields  $m[i, j] = m[i, k] + m[k + 1, j] + p_{i-1}p_k p_j$ .
- ▶ The minimum cost is obtained by minimizing over all possible splittings:

$$m[i, j] = \begin{cases} 0, & \text{if } i = j \\ \min_{i \leq k < j} \{m[i, k] + m[k + 1, j] + p_{i-1}p_k p_j\}, & \text{if } i < j \end{cases}$$

- ▶ Use  $s[i, j]$  to keep track of the splitting index  $k$ , i.e.,

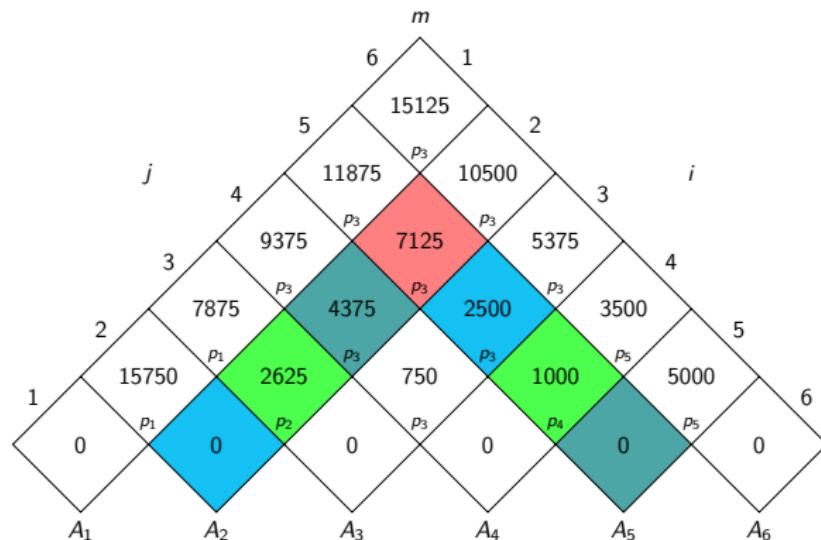
$$s[i, j] = \arg \min_{i \leq k < j} \{m[i, k] + m[k + 1, j] + p_{i-1}p_k p_j\}$$

# Matrix Chain Multiplication

## Example (Bottom-Up)

Given  $p = (30, 35, 15, 5, 10, 20, 25)$ , i.e.,

| matrix    | $A_1$          | $A_2$          | $A_3$         | $A_4$         | $A_5$          | $A_6$          |
|-----------|----------------|----------------|---------------|---------------|----------------|----------------|
| dimension | $30 \times 35$ | $35 \times 15$ | $15 \times 5$ | $5 \times 10$ | $10 \times 20$ | $20 \times 25$ |



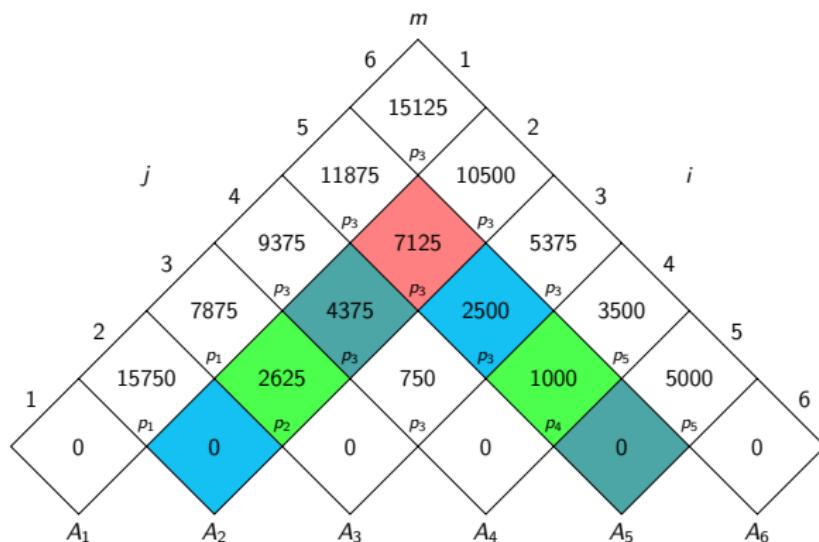
# Matrix Chain Multiplication

## Example (Bottom-Up) Cont.

$m[2, 5]$

$$= \min \begin{cases} m[2, 2] + m[3, 5] + p_1 p_2 p_5 = 0 + 2500 + 35 \cdot 15 \cdot 20 = 13000 \\ m[2, 3] + m[4, 5] + p_1 p_3 p_5 = 2625 + 1000 + 35 \cdot 5 \cdot 20 = 7125 \\ m[2, 4] + m[5, 5] + p_1 p_4 p_5 = 4375 + 0 + 35 \cdot 10 \cdot 20 = 11375 \end{cases}$$

$= 7125$



## Matrix Chain Multiplication

### Example (Bottom-Up) Cont.

Given  $p = (30, 35, 15, 5, 10, 20, 25)$ , we build

$$[m_{ij}] = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \left[ \begin{matrix} 0 & 15750 & 7875 & 9375 & 11875 & 15125 \\ 0 & 0 & 2625 & 4375 & 7125 & 10500 \\ 0 & 0 & 0 & 750 & 2500 & 5375 \\ 0 & 0 & 0 & 0 & 1000 & 3500 \\ 0 & 0 & 0 & 0 & 0 & 5000 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \right] \end{matrix}$$

Based on the matrix  $[m_{ij}]$ , **the** minimum number of multiplication is given by

$$m[1, 6] = 15125$$

# Matrix Chain Multiplication

## Example (Bottom-Up) Cont.

We also build

$$[s_{ij}] = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \left[ \begin{array}{cccc|cc} 0 & 1 & 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 3 & 3 & 3 \\ 0 & 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{matrix}$$

Based on the matrix  $[s_{ij}]$ , **an** optimal parenthesization (ordering) of the matrix chain multiplication is given by

$$(A_1(A_2A_3))((A_4A_5)A_6)$$

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

# Fibonacci Sequence

## Definition

The Fibonacci Sequence is given by  $f_0 = f_1 = 1$ , and for  $n \geq 2$ ,

$$f_n = f_{n-1} + f_{n-2}$$

- ▶ Tiling of  $2 \times n$  board by  $1 \times 2$  blocks.
- ▶ Spirals on pineapple/sunflower.
- ▶ growth of an idealized (biologically unrealistic) rabbit population.



# Regions in a Plane Divided by $n$ Straight Lines

## Example

$r_1 = 2, r_2 = 4, r_3 = 7, \dots$  In general,

$$r_{n+1} = r_n + n + 1$$

which admits an explicit solution as

$$r_n = \binom{n}{2} + n + 1$$

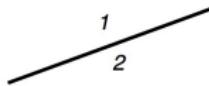
See Lovász, Sec. 11.2.1 for a combinatorial proof.

0 lines

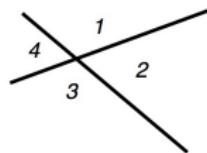
1

1 region

1 line

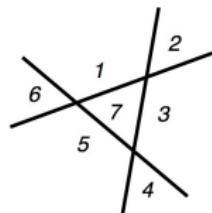


2 lines



4 regions

3 lines



7 regions

## Second Order Homogeneous Equations

### Example

Find the general solution to

$$f_{n+2} + f_{n+1} - 6f_n = 0$$

that is,

$$(A^2 + A - 6)f_n = 0$$

where  $A$  is the **advancement/shift operator** given by  $Af_n = f_{n+1}$ .

Naturally,  $A^k f_n = f(n+k)$ ,  $k, n \in \mathbb{N}$ .

We solve the difference equation by trying ansatz  $f_n = r^n$ , thus

$$(A^2 + A - 6)f_n = r^{n+2} + r^n - 6r^n = r^n(r^2 + r - 6) = 0$$

thus  $r^2 + r - 6 = 0$ , i.e.,  $r_{1,2} = -3, 2$ . Hence  $(-3)^n$  and  $2^n$  are solutions to the equation. By linearity, the general solution is thus given by

$$f_n = c_1(-3)^n + c_22^n, \quad c_1, c_2 \in \mathbb{C}$$

## Second Order Homogeneous Equations

### Example (Cont.)

The equation  $r^2 + r - 6 = 0$  called the characteristic equation for the difference equation.

Alternatively, we observe that

$$(A + 3)[(A - 2)f_n] = 0 \Rightarrow (A - 2)f_n = f_{n+1} - 2f_n = c_1(-3)^n$$

$$(A - 2)[(A + 3)f_n] = 0 \Rightarrow (A + 3)f_n = f_{n+1} + 3f_n = c_22^n$$

We can take the difference and cancel  $f_{n+1}$ , hence

$$5f_n = c_22^n - c_1(-3)^n \quad c_1, c_2 \in \mathbb{C}$$

or by relabeling the constant coefficients, we have

$$f_n = \tilde{c}_1(-3)^n + \tilde{c}_22^n, \quad \tilde{c}_1, \tilde{c}_2 \in \mathbb{C}$$

## Second Order Homogeneous Equations

### Example

Find the general solution to

$$f_{n+2} + 4f_{n+1} + 4f_n = 0$$

i.e.,

$$(A^2 + 4A + 4)f_n = 0$$

Solve the characteristic equation  $r^2 + 4r + 4 = 0$ , thus  $r_{1,2} = -2$  (repeated roots). The general solution is thus given by

$$f_n = c_1(-2)^n + c_2 n(-2)^n = (-2)^n(c_1 + c_2 n)$$

where  $c_1, c_2 \in \mathbb{C}$ .

## Second Order Homogeneous Equations

### Example (Cont.)

To add a little more detail, note that  $(A^2 + 4A + 4)f_n = (A + 2)^2 f_n = 0$ , which implies that  $(A + 2)f_n = f_{n+1} + 2f_n = c_1(-2)^n$ . Therefore

$$f_n + 2f_{n-1} = c_1(-2)^{n-1}$$

$$(-2)f_{n-1} + 2(-2)f_{n-2} = c_1(-2)^{n-2} \cdot (-2)$$

$$(-2)^2 f_{n-2} + 2(-2)^2 f_{n-1} = c_1(-2)^{n-3} \cdot (-2)^2$$

⋮

$$(-2)^{n-1} f_1 + 2(-2)^{n-1} f_0 = c_1(-2)^0 \cdot (-2)^{n-1}$$

Add together, we have  $f_n - (-2)^n f_0 = c_1 n (-2)^{n-1}$ . Relabeling, we have

$$f_n = \tilde{c}_1 n (-2)^n + \tilde{c}_2 (-2)^n = (\tilde{c}_1 n + \tilde{c}_2) (-2)^n$$

where  $\tilde{c}_1, \tilde{c}_2 \in \mathbb{C}$ .

## General Homogeneous Equations

### Theorem (Distinct Roots)

Consider the following linear homogeneous recurrence equation

$$p(A)f_n = (A - r_1)(A - r_2) \cdots (A - r_k)f_n = 0$$

with  $r_1, r_2, \dots, r_k$  distinct non-zero constants. The general solution is given by

$$f_n = c_1 r_1^n + c_2 r_2^n + \cdots + c_k r_k^n$$

with constants  $c_1, c_2, \dots, c_k$ .

## General Homogeneous Equations

### Theorem (Repeated Roots)

Let  $k \geq 1$  and consider the recurrence equation

$$(A - r)^k f_n = 0$$

Then the general solution is given by

$$\begin{aligned}f_n &= c_1 r^n + c_2 n r^n + c_3 n^2 r^n + c_4 n^3 r^n + \cdots + c_k n^{k-1} r^n \\&= (c_1 + c_2 n + c_3 n^2 + c_4 n^3 + \cdots + c_k n^{k-1}) r^n\end{aligned}$$

with constants  $c_1, c_2, \dots, c_k$ .

## General Homogeneous Equations

### Example

Consider

$$(A - 1)^5(A + 1)^3(A - 3)^2(A + 8)(A - 9)^4 f_n = 0$$

The general solution is given by

$$\begin{aligned}f_n = & c_1 + c_2 n + c_3 n^2 + c_4 n^3 + c_5 n^4 \\& + (c_6 + c_7 n + c_8 n^2)(-1)^n \\& + (c_9 + c_{10} n)3^n \\& + c_{11}(-8)^n \\& + (c_{12} + c_{13} n + c_{14} n^2 + c_{15} n^3)9^n\end{aligned}$$

with constants  $c_1, c_2, \dots, c_{15}$ .

# Inhomogeneous/Nonhomogeneous Equations

## General Strategy

Homogeneous solution + particular solution

## Example

Find the general solution to

$$(A+2)(A-6)f_n = 3^n$$

- ▶ Homogeneous solution:  $f_n^{(h)} = c_1(-1)^n + c_26^n.$
- ▶ Particular solution: Try  $f_n^{(p)} = d3^n.$  ( $\Rightarrow d = -1/15$ )

General solution

$$f_n = c_1(-1)^n + c_26^n - \frac{1}{15}3^n$$

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

## Overview

- ▶ Study a way to describe the growth of functions in the limit — asymptotic efficiency
- ▶ Focus on what's important (leading factor) by abstracting lower-order terms and constant factors
- ▶ Indicate running times of algorithms
- ▶ A way to compare “sizes” of functions

$$O \approx \leq$$

$$\Omega \approx \geq$$

$$\Theta \approx =$$

In addition,

$$o \approx <$$

$$\omega \approx >$$

# Big “Oh” Notation

## Definition

A function  $g(n)$  is an **asymptotic upper bound** for  $f(n)$ , denoted by

$$f(n) = O(g(n))$$

if there exist positive constants  $c$  and  $n_0$  such that

$$0 \leq f(n) \leq cg(n) \text{ for all } n \geq n_0$$

i.e.,

$$\limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty$$

## Example

Show that  $2n + 10 = O(n^2)$ .

**Proof 1.** Since  $2n + 10 \leq n^2$  for  $n \geq 5$ , we can choose  $c = 1$  and  $n_0 = 5$ .

**Proof 2.** Observe that  $2n + 10 \leq 2n^2 + 10n^2 = 12n^2$  for  $n \geq 1$ , we can choose  $c = 12$  and  $n_0 = 1$ .

## Big “Oh” Notation

- ▶  $O(g(n))$  is a **set** of functions

$$O(g(n)) = \{f(n) \mid \exists c, n_0 > 0 \text{ s.t. } 0 \leq f(n) \leq cg(n) \text{ for } n \geq n_0\}$$

We write  $f(n) = O(g(n))$  or  $f(n) \in O(g(n))$ .

- ▶ Examples of functions in  $O(n^2)$ :

- ▶  $n^2 + n$
- ▶  $n^2 + 1000n$
- ▶  $1000n^2 + 1000n$
- ▶  $n/1000$
- ▶  $n^2/\lg n$

## $\Omega$ Notation

### Definition

A function  $g(n)$  is an **asymptotic lower bound** for  $f(n)$ , denoted by

$$f(n) = \Omega(g(n))$$

if there exist positive constants  $c$  and  $n_0$  such that

$$0 \leq cg(n) \leq f(n) \text{ for all } n \geq n_0$$

i.e.,

$$\liminf_{n \rightarrow \infty} \frac{f(n)}{g(n)} > 0$$

### Example

$\sqrt{n} = \Omega(\lg n)$ . We can choose  $c = 1$  and  $n_0 = 16$ .

## $\Omega$ Notation

- ▶  $\Omega(g(n))$  is a **set** of functions

$$\Omega(g(n)) = \{f(n) \mid \exists c, n_0 > 0 \text{ s.t. } 0 \leq cg(n) \leq f(n) \text{ for } n \geq n_0\}$$

- ▶ Examples of functions in  $O(n^2)$ :

- ▶  $n^2$
- ▶  $n^2 + n$
- ▶  $n^2 - n$
- ▶  $1000n^2 + 1000n$
- ▶  $1000n^2 - 1000n$
- ▶  $n^{2+\varepsilon}, \varepsilon > 0$
- ▶  $n^2 \lg n$
- ▶  $n^3$

## $\Theta$ Notation

### Definition

A function  $g(n)$  is an **asymptotic tight bound** for  $f(n)$ , denoted by

$$f(n) = \Theta(g(n))$$

if there exist constants  $c_1$ ,  $c_2$ , and  $n_0$  such that

$$0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n) \text{ for all } n \geq n_0$$

### Example

- ▶  $\frac{1}{2}n^2 - 2n = \Theta(n^2)$ . We can choose  $c_1 = \frac{1}{4}$ ,  $c_2 = \frac{1}{2}$ , and  $n_0 = 8$ .
- ▶ If  $p(n) = \sum_{i=1}^d a_i n^i$  and  $a_d > 0$ , then  $p(n) = \Theta(n^d)$ .

## $\Theta$ Notation

- $\Theta(g(n))$  is a **set** of functions

$$\Theta(g(n))$$

$$= \{f(n) \mid \exists c_1, c_2, n_0 \text{ s.t. } 0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n) \text{ for } n \geq n_0\}$$

- Examples of functions in  $O(n^2)$ :

- $n^2$
- $n^2 + n$
- $n^2 - n$
- $1000n^2 + 1000n$
- $1000n^2 - 1000n$

## Theorem

$$\Theta(g(n)) = O(g(n)) \cap \Omega(g(n)).$$

# Asymptotic Notation

## Example

The function grows faster as the list goes:

- $\log^* n$
- $n$
- $n^{\log \log n}$
- $n^n$
- $\log \log \log n$
- $n \log n$
- $n^{\log n}$
- $2^{2^n}$
- $\log \log n$
- $n^{3/2}$
- $2^n$
- $2^{2^{2^n}}$
- $\log n$
- $n^2$
- $(\log n)^n$
- $n^3$
- $n^{n/2}$
- $\sqrt{n}$

## Notation

The iterated logarithm, “log star”, is given by

$$\log^* n := \begin{cases} 0, & n \leq 1 \\ 1 + \log^*(\log n), & n > 1 \end{cases}$$

which is well defined if base is  $> e^{1/e} \approx 1.444667$ . We use  $\lg^*$  for binary iterated logarithm.

## Using Limits for Comparing Orders of Growth

Let

$$L := \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$$

We have several possibilities (given that  $f$  and  $g$  nonnegative.)

- ▶ If  $L = 0$ , then  $f(n) = O(g(n))$ ;
- ▶ If  $L = \infty$ , then  $f(n) = \Omega(g(n))$ ;
- ▶ If  $0 < L < \infty$ , then  $f(n) = \Theta(g(n))$ ;
- ▶ Limit does not exist: inconclusive.

### L'Hôpital's rule

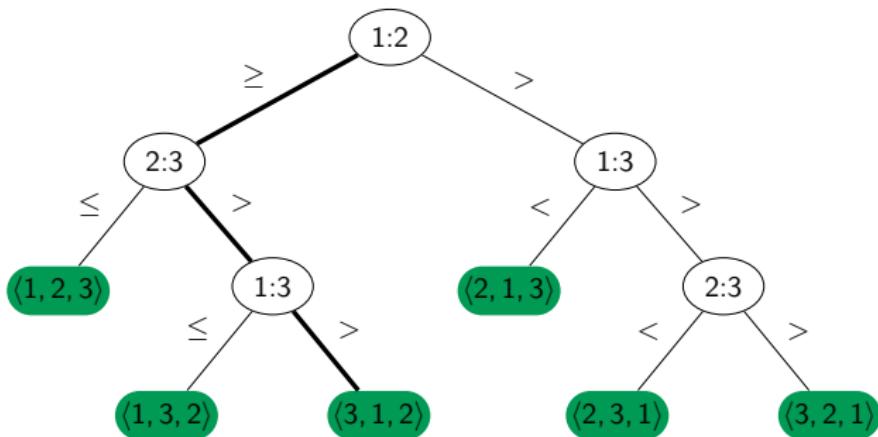
Let  $f$  and  $g$  be differentiable. If  $\lim_{x \rightarrow \infty} |f(x)| = \lim_{x \rightarrow \infty} |g(x)| = \infty$ , and  $\lim_{x \rightarrow \infty} g'(x) \neq 0$ , then

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}$$

provided that the limit on the RHS exists in  $\overline{\mathbb{R}}$ .

## Lower Bound for Sorting

In a **comparison sort**, we use only comparisons between elements to gain order information about an input sequence  $\langle a_1, a_2, \dots, a_n \rangle$ , and the output is given by permutation of the input as  $\langle a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(n)} \rangle$ ,  $\sigma \in S_n$ .



Because any correct sorting algorithm must be able to produce each permutation of its input, each of the  $n!$  permutations on  $n$  elements must appear as one of the leaves of the decision tree for a comparison sort to be correct.

# Lower Bound for Sorting

## Theorem

*Given  $n$  elements, any comparison sort algorithm requires  $\Omega(n \lg n)$  comparisons in the worst case.*

## Proof.

It suffices to determine the height of a decision tree in which each permutation appears as a reachable leaf. Consider a decision tree of height  $h$  with  $\ell$  **reachable leaves**. Because each of the  $n!$  permutations of the input appears as some leaf, we have  $n! < \ell$ . Since a binary tree of height  $h$  has no more than  $2^h$  leaves, we have

$$n! \leq \ell \leq 2^h$$

Take the logarithms, we have

$$h \geq \lg(n!) = \Omega(n \lg n)$$

where the last equality follows from Stirling's approximation formula. □

# Table of Contents

1. Binomial Coefficients
2. Multichoosing
3. Inclusion-Exclusion Principle
4. Matrix Chain Multiplication
5. Linear Recurrence Equations
6. Asymptotic Notations
7. Master Method

# Solving Recurrences

## Example

$T(n) = 4T(n/2) + n$ , which means

$$T(n) = \begin{cases} \Theta(1), & n = 1 \text{ (usually omit this part)} \\ 4(T/2) + n, & n > 1 \end{cases}$$

## General Methods

- ▶ Substitution Method (guess, say, by trial and error, and prove, say, by induction).
- ▶ Recursion-tree Method.
- ▶ Master Method.

# Solving Recurrences

## Example (Factorial)

Let  $T(n)$  denote the worst-case running time of `fact`, then

$$T(1) = d$$

$$T(n) = T(n - 1) + c$$

where  $c$  is a constant denoting the work of the comparison–conditional–multiplication–return, and  $d$  is a constant denoting the work of the comparison–conditional–return.

---

```
1 Function fact(n):
2   if n = 1 then
3     | return 1
4   else
5     | return n · fact(n - 1)
6   end
7 end
```

---

# Solving Recurrences

## Example (Merge Sort)

Let  $T(n)$  denote the worst-case running time of Merge Sort on an input array containing  $n$  elements. Then, for a constant  $c$ , we have:

$$T(1) = c$$

$$T(n) = T(\lfloor \frac{n}{2} \rfloor) + T(\lceil \frac{n}{2} \rceil) + cn$$

which is a typical “*divide-conquer-combine*” process.

---

```
1 Function mergeSort(A[1...n]):  
2   if n = 1 then  
3   |   return A  
4   else  
5   |   L ← mergeSort(1... $\lfloor \frac{n}{2} \rfloor$ )  
6   |   R ← mergeSort( $\lfloor \frac{n}{2} \rfloor + 1 \dots n$ )  
7   |   return merge(L, R)  
8   end  
9 end
```

---

# Solving Recurrences

## Example

Solve  $T(n) = 4T(n/2) + n$ .

It is sufficient to consider  $n = 2^m$ ,  $m \in \mathbb{N}$ . Thus  $n/2 = 2^{m-1}$ , and  $m = \log_2 n = \lg n$ . Now we have  $T(2^m) = 4T(2^{m-1}) + 2^m$ . Let  $\tilde{T}(m) := T(2^m)$ , then

$$\tilde{T}(m) = 4\tilde{T}(m-1) + 2^m$$

whose general solution is given by

$$\tilde{T}(m) = c \cdot 4^m + d \cdot 2^m, \quad c, d \text{ constants}$$

thus

$$\begin{aligned} T(n) &= c \cdot 4^{\log_2 n} + d \cdot 2^{\log_2 n} = c \cdot n^{\log_2 4} + d \cdot n^{\log_2 2} \\ &= c \cdot n^2 + d \cdot n = \Theta(n^2) \end{aligned}$$

## Master Theorem/Method

Theorem (Master Thoerem, cf., Corman, Leiserson, Rivest, & Stein.)

If  $T(n) = aT(n/b) + f(n)$  (for constants  $a \geq 1$ ,  $b > 1$ ,  $d \geq 0$ ), then

1.  $T(n) = \Theta(n^{\log_b a})$  if  $f(n) = O(n^{\log_b a - \varepsilon})$  for some constant  $\varepsilon > 0$ .
2.  $T(n) = \Theta(n^{\log_b a} \lg n)$  if  $f(n) = \Theta(n^{\log_b a})$ .
3.  $T(n) = \Theta(f(n))$ , if  $f(n) = \Omega(n^{\log_b a + \varepsilon})$  for some constant  $\varepsilon > 0$ , and if  $af(n/b) \leq cf(n)$  for some constant  $c < 1$  and all sufficiently large  $n$  (regularity condition).

# Master Theorem/Method

## Remark

1.  $n^{\log_b a}$  is polynomially larger than  $f(n)$ , e.g.,

$$T(n) = 7 \cdot T(n/2) + \Theta(n^2)$$

2.  $n^{\log_b a}$  and  $f(n)$  are on the same order, e.g.,

$$T(n) = 2 \cdot T(n/2) + \Theta(n)$$

3.  $f(n)$  is polynomially larger than  $n^{\log_b a}$ , and satisfies the **regularity condition**, e.g.,

$$T(n) = 4 \cdot T(n/2) + n^3$$

Note that the master theorem does not cover all possible cases (e.g., quick sort).

## Part IV

### Selected Topics in Graph Theory

# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Graphs

## Definition

A **graph**  $G$  consists of a set of **vertices**, denoted by  $V(G)$ , a set of edges, denoted by  $E(G)$ , and a relation called **incidence** so that each edge is incident with either one or two vertices, called **ends** (or **endpoints**). For convenience, we sometimes write  $G = (V, E)$  to indicate that  $G$  is a graph with **vertex set**  $V$  and **edge set**  $E$ .

## Definition

Two distinct vertices  $u, v$  in a graph  $G$  are **adjacent** if there is an edge with ends  $u, v$ . We also call  $u, v$  neighbors in  $G$ .

## Remark

- ▶ Vertices are also called nodes, points, locations, stations, etc.
- ▶ Edges are also called arcs, lines, links, pipes, connectors, etc.

# Loops, Parallel Edges, and Simple Graphs

## Definition

An edge with just one end is called a **loop**. Two distinct edges with the same ends are are **parallel** (called “parallel edges” or “multiple edges”). A graph without loops or parallel edges is called **simple**.

## Remark

We specify a simple graph  $(V, E)$  by its **vertex set**  $V$ , and **edge set**  $E$ , where  $E \subset \binom{V}{2}$ . We write  $e = uv$  or  $e = vu$  for an edge  $e \in E$  with ends  $u, v \in V$ . (That is,  $e = \{u, v\}$ .)

# Isomorphism

## Definition

An **isomorphism** from a simple graph  $G$  to a simple graph  $H$  is a bijection  $f : V(G) \rightarrow V(H)$  such that  $uv \in E(G)$  iff  $f(u)f(v) \in E(H)$ . We say “ $G$  is **isomorphic** to  $H$ ”, denoted  $G \cong H$ , if there is an isomorphism from  $G$  to  $H$ .

## Remark

The **relation isomorphism**, consisting of the set of ordered pairs  $(G, H)$  such that  $G$  is isomorphic to  $H$  is an equivalence relation on the set of simple graphs.

# Representing Graph

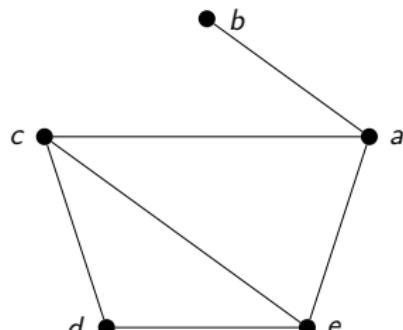
## Example

- ▶ By specifying the vertex and edge sets of the graph.
- ▶ By using database structure.
- ▶ By showing a “drawing” of the graph.

## Ajacency Tables

An adjacency table lists all the vertices of the graph and the vertices adjacent to them. Consider  $G = (V, E)$  with  $V = \{a, b, c, d, e\}$  and  $E = \{\{a, b\}, \{a, c\}, \{a, e\}, \{c, d\}, \{c, e\}, \{d, e\}\}$

| Vertex | Adjacent Vertices |
|--------|-------------------|
| a      | b, c, e           |
| b      | a                 |
| c      | a, d, e           |
| d      | c, e              |
| e      | a, c, d           |



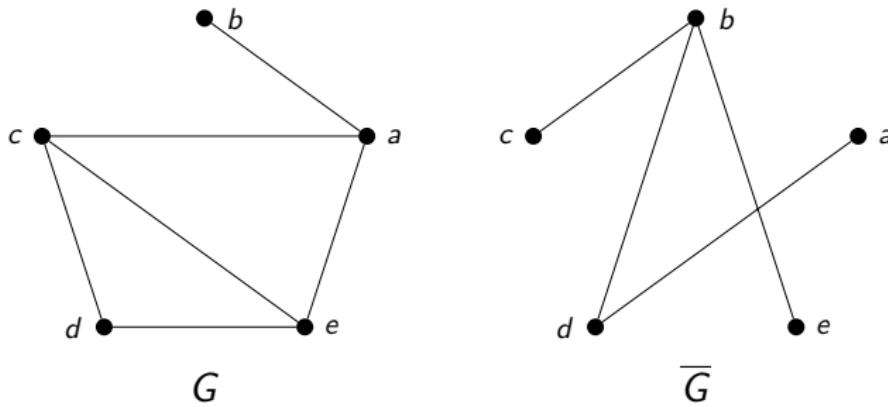
# Using Graphs as Models

## Example

- ▶ Acquaintance relations.

## Definition

The **complement**  $\overline{G}$  of a simple graph  $G$  is the simple graph with vertex set  $V(G)$  defined by  $uv \in E(\overline{G})$  iff  $uv \notin E(G)$ . Note that given graph  $G = (V, E)$ , we have  $\overline{G} = (V, \binom{V}{2} - E)$ .



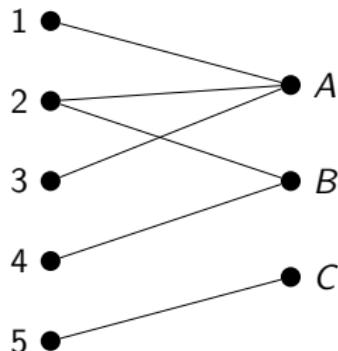
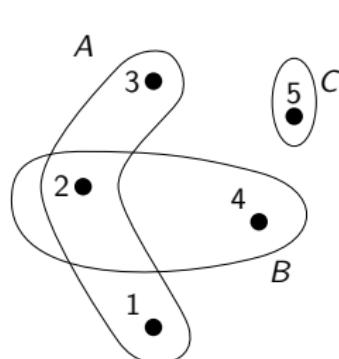
# Using Graphs as Models

## Example

- Job assignments.

## Definition

A graph (not necessarily simple) is **bipartite** if  $V(G)$  is the union of two disjoint (possibly empty) independent sets (i.e., a set of pairwise nonadjacent vertices), called **partite sets** of  $G$ .



|   | A | B | C |
|---|---|---|---|
| 1 | 1 | 0 | 0 |
| 2 | 1 | 1 | 0 |
| 3 | 1 | 0 | 0 |
| 4 | 0 | 1 | 0 |
| 5 | 0 | 0 | 1 |

# Using Graphs as Models

## Example

- ▶ Maps and coloring.
- ▶ Routes in road networks.
- ▶ ...

# Standard Graphs

## Definition

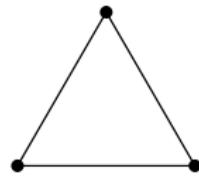
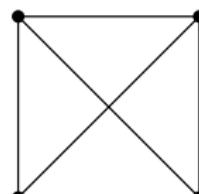
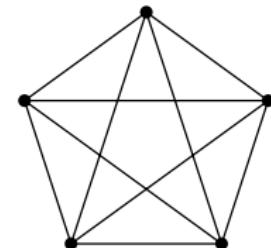
The **null graph** is the graph whose vertex set and edge set are empty.

## Definition

A graph  $G$  is **complete** if it is simple and all pairs of distinct vertices are adjacent. A complete graph on  $n$  vertices is denoted by  $K_n$ .

## Definition

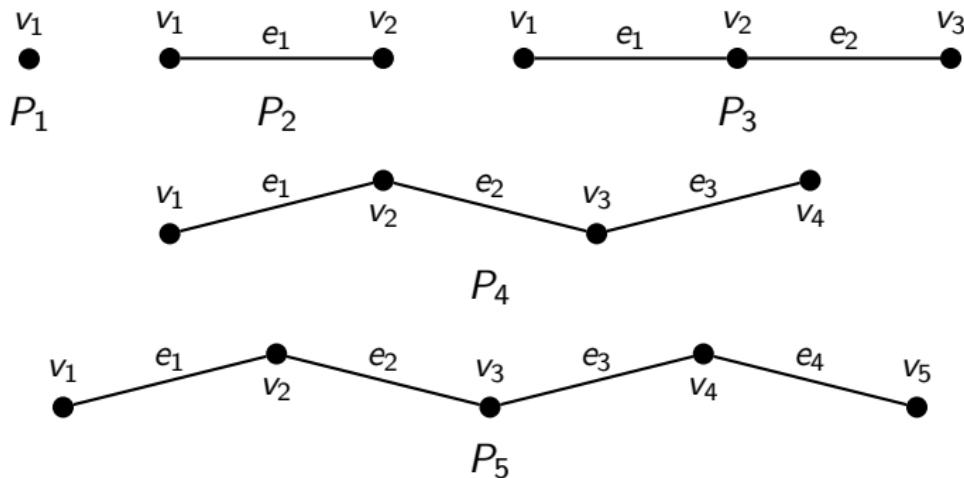
A **clique** in a graph is a set of pairwise adjacent vertices.

 $K_1$  $K_2$  $K_3$  $K_4$  $K_5$

# Standard Graphs

## Definition

A graph  $G$  is called a **path** if the vertices can be ordered as  $v_1, \dots, v_n$ , and edges can be ordered as  $e_1, \dots, e_{n-1}$  such that  $e_i = v_i v_{i+1}$ ,  $i = 1, \dots, n$ . A path on  $n$  vertices is denoted by  $P_n$ .



# Standard Graphs

## Definition

A graph  $G$  is a **cycle** if  $V(G)$  can be ordered as  $v_1, \dots, v_n$ , and  $E(V)$  can be ordered as  $e_1, \dots, e_n$ , where

$$e_i = \begin{cases} v_i v_{i+1}, & 1 \leq i \leq n-1 \\ v_n v_1, & i = n \end{cases}$$

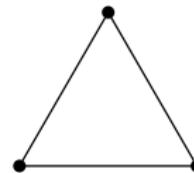
A cycle on  $n$  vertices is denoted by  $C_n$ .



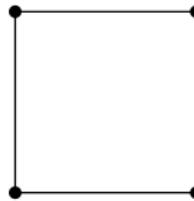
$C_1$



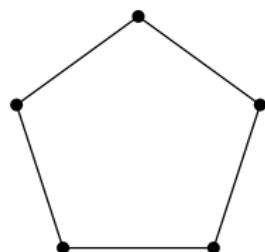
$C_2$



$C_3$



$C_4$



$C_5$

# Subgraphs

## Definition

If  $G, H$  have  $V(H) \subset V(G)$ , and  $E(H) \subset E(G)$  with incidence in  $H$  the same as  $G$ , then  $H$  is a **subgraph** of  $G$ , denoted by  $H \subset G$ .

Obviously, given  $H_1, H_2 \subset G$ , then

- ▶  $H_1 \cap H_2 \subset G$ , with

$$\begin{aligned}V(H_1 \cap H_2) &= V(H_1) \cap V(H_2) \\E(H_1 \cap H_2) &= E(H_1) \cap E(H_2)\end{aligned}$$

- ▶  $H_1 \cup H_2 \subset G$ , with

$$\begin{aligned}V(H_1 \cup H_2) &= V(H_1) \cup V(H_2) \\E(H_1 \cup H_2) &= E(H_1) \cup E(H_2)\end{aligned}$$

# Subgraphs

## Remark

When we name a graph without naming its vertices, we often mean its isomorphsim class. Technically, " $H$  is a subgraph of  $G$ " means that some subgraph of  $G$  is isomorphic to  $H$  (we say " $G$  contains a **copy** of  $G$ ").

## Example

- ▶  $C_3$  is a subgragh of  $K_5$ .
- ▶  $P_1$  is a subgragh of  $K_5$ .
- ▶  $K_5$  is a subgragh of  $K_6$ .

# Degree of Vertices

## Definition

The **degree** of a vertex  $v$  in a graph  $G$ , denoted  $\deg(v)$  is the number of incident edges (loops counted twice). We write  $\deg_G(v)$  if  $G$  is not clear (i.e., we are not sure if  $v \in V(G)$ ).

## Theorem

For all  $G = (V, E)$ ,

$$\sum_{v \in V} \deg(v) = 2|E|$$

## Proof.

By double counting. □

Corollary (Handshaking lemma/degree sum formula)

Every graph has an even number of odd degree vertices.

# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Walks

## Definition

A **walk**  $W$  in a graph  $G$  is a sequence  $v_0, e_1, v_1, \dots, e_n, v_n$  such that every  $e_i$  has ends  $v_{i-1}$  and  $v_i$ . If  $v_0 = v_n$ , we say that  $W$  is closed.

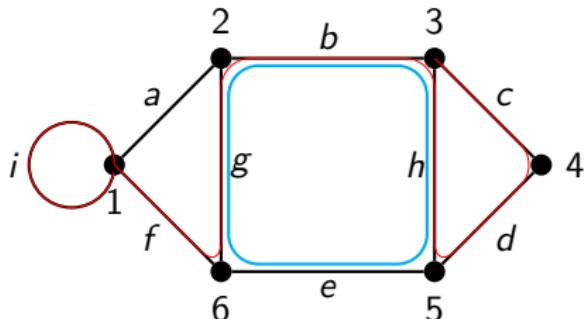
## Definition

The **length** of a walk, path, or cycle is its number of edges. A walk is **closed** if its ends are the same.

## Remark

- ▶ A walk is **NOT** a graph in general.
- ▶ A path is a graph.
- ▶ If  $v_0, \dots, v_n$  in a walk are distinct, we also call this walk a path.
- ▶ A walk with only 1 vertex has length 0.

## Walks



### Example

- ▶  $W = 3, c, 4, d, 5, h, 3, b, 2, g, 6, f, 1, i, 1$ .  $W$  is NOT a closed walk (b/c 3 is not the same vertex as 1). The length of  $W$  is 7.
- ▶  $W' = 3, b, 2, g, 6, e, 5, h, 3$ .  $W'$  is a closed walk, with length 4. Note that  $W' \cong C_4$ .

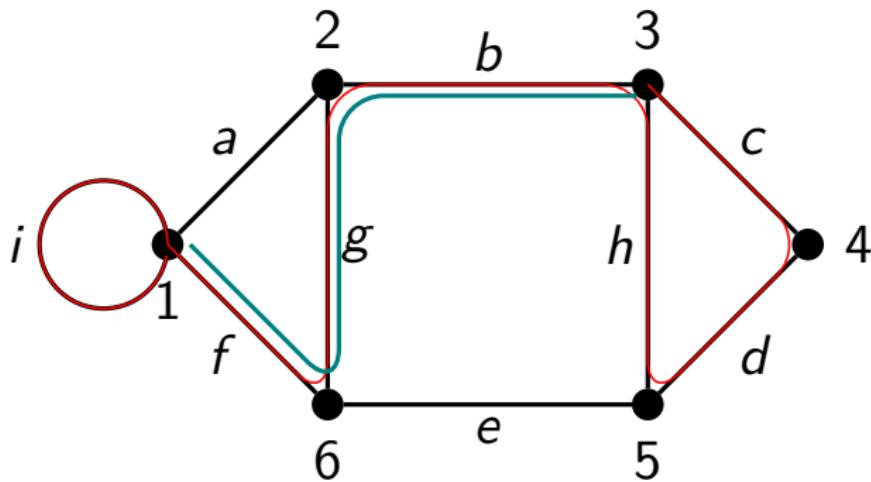
# Connected Graph

## Definition

A graph  $G$  is **connected** if for all  $u, v \in V(G)$ , there is a walk from  $u$  to  $v$  (also called a  $u, v$ -walk). Otherwise,  $G$  is **disconnected**.

## Theorem

*If there is a walk from  $u$  to  $v$ , then there is a path from  $u$  to  $v$ .*



## Connected Graph

### Proof.

Claim: The path from  $u$  to  $v$  is the shortest walk from  $u$  to  $v$  (i.e., the walk of **minimum** length.)

Indeed. Let  $W$  be a walk of minimum length from  $u$  to  $v$ , say

$v_0 e_1 v_1 e_2 v_2 \cdots e_n v_n$ , with  $u = v_0$  and  $v = v_n$ .

Suppose this is **NOT** a path, then there exists  $v_i = v_j$  such that  $0 \leq i < j \leq n$ . Therefore  $v_0 e_1 v_1 \cdots v_i e_{j+1} v_{j+1} \cdots e_n v_n$  is a shorter walk from  $u$  to  $v$ , which is a contradiction.



## Connected Graph

### Theorem

$G$  is disconnected iff there is a partition  $\{X, Y\}$  of  $V(G)$  such that no edge has an end in  $X$  and an end in  $Y$ .

### Proof.

( $\Leftarrow$ ) True by definition of connectivity.

( $\Rightarrow$ ) Choose  $x, y \in V(G)$  such that no walk from  $x$  to  $y$  exists, define

$$X := \{z \mid \exists \text{ a walk from } x \text{ to } z\}$$

$$Y := V(G) \setminus X$$

Claim: no edge has an end in  $X$  and an end in  $Y$ , which is obvious. □

# Connected Graph

## Theorem

Given  $H_1, H_2 \subset G$ ,  $H_1, H_2$  connected graphs, and  $V(H_1) \cap V(H_2) \neq \emptyset$ , then  $H_1 \cup H_2$  is connected.

## Proof.

Let  $u, v \in V(H_1 \cup H_2)$ . Choose  $w \in V(H_1 \cap H_2)$  ( $\neq \emptyset$ ), note that  $u, v$  is either in  $H_1$  or  $H_2$ , w.l.o.g., let  $u \in V(H_1)$ ,  $v \in V(H_2)$ . For  $i = 1, 2$ ,  $H_i$  is connected, so there is a  $u, w$ -walk  $W_i$ . Now concatenate  $W_1$  and  $W_2$ , we have a  $u, v$ -walk. Since  $u, v$  are arbitrary, therefore  $H_1 \cup H_2$  is connected.



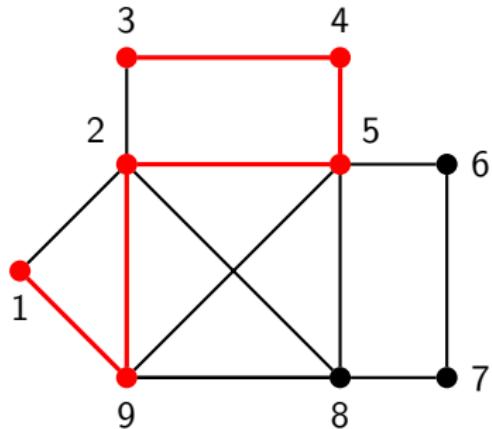
# Connected Graph

## Definition

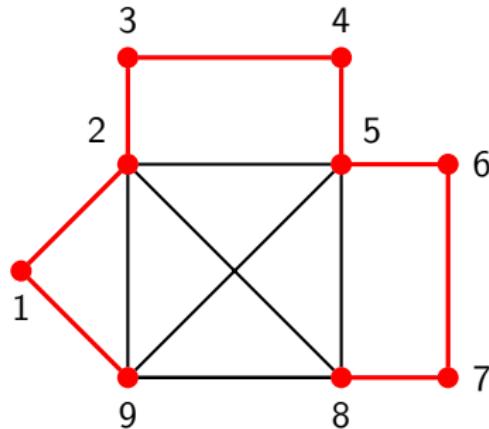
A **maximal** connected subgraph of  $G$  is a subgraph that is connected and is **not** contained in any other connected subgraph of  $G$ .

## Remark

A path/subgraph in  $G$  is **maximal** if it cannot be enlarged.



a path that is maximal

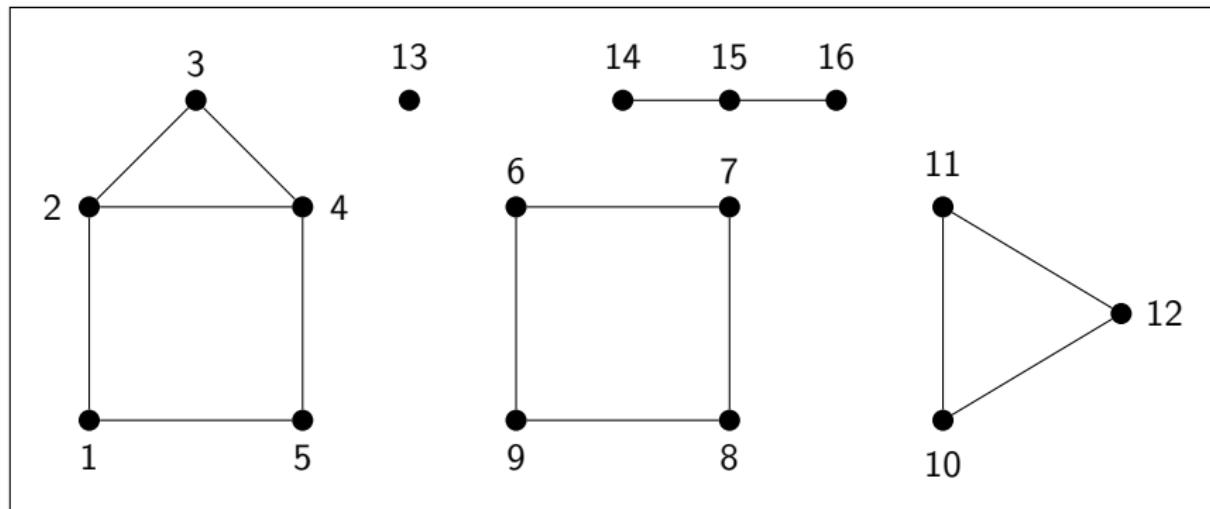


a path that is both  
maximal and maximum

# Connected Graph

## Definition

A **component** of a graph  $G$  is a **maximal** non-empty connected subgraph of  $G$ . The number of components of  $G$  is denoted  $\text{comp}(G)$ .



$$\text{comp}(G) = 5$$

# Connected Graph

## Theorem

*Every vertex is in a **unique** component.*

## Proof.

Let  $v \in V(G)$ . Note that  $v$  is in a connected subgraph  $(\{v\}, \emptyset)$ , which consists of only  $v$  and no other vertices or edges. If  $H_1$  and  $H_2$  are connected subgraphs containing  $v$ , then  $H_1 \cap H_2 \neq \emptyset$ , thus  $H_1 \cup H_2$  is connected. Therefore  $v$  is in a unique component. □

## Remark

- ▶ Components are pairwise disjoint;
- ▶ No two components share a vertex;
- ▶ Adding an edge with endpoints in distinct components combine the two components into one.
- ▶ Adding/Deleting an edge decreases/increases the number of components by at most 1.

# Connected Graph

## Deleting Edges

Given graph  $G$ ,  $S \subset E(G)$ , then  $G - S$  is the graph obtained from  $G$  by deleting  $S$ .

## Deleting Vertices

Given graph  $G$ ,  $X \subset V(G)$ , then  $G - X$  is the graph obtained from  $G$  by deleting every vertex in  $X$  **and every edge incident to a vertex in  $X$** .

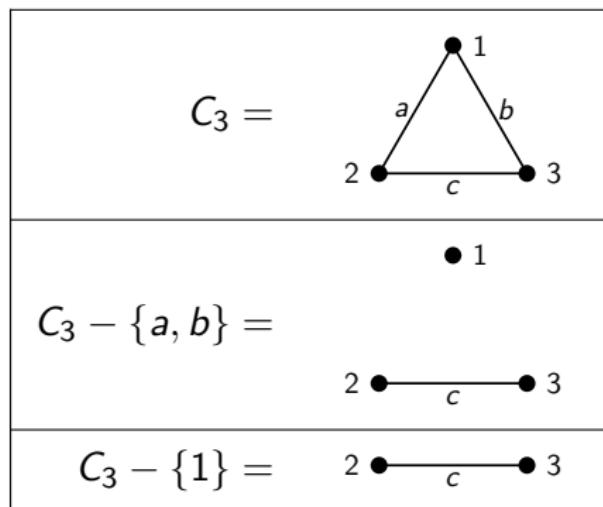
## Notation

If  $e \in E(G)$  or  $v \in V(G)$ , we define

- ▶  $G - e := G - \{e\}$ ;
- ▶  $G - v := G - \{v\}$ .

For example,

$$G - v - w = G - \{v, w\}.$$



# Connected Graph

## Definition

An edge  $e \in E(G)$  is called a ***cut-edge*** or ***bridge*** if no cycle contains  $e$ .

## Theorem

Given graph  $G$  and  $e \in E(G)$ , then

- ▶ either  $e$  is a cut-edge and  $\text{comp}(G - e) = \text{comp}(G) + 1$ ;
- ▶ or  $e$  is NOT a cut-edge and  $\text{comp}(G - e) = \text{comp}(G)$ .

## Proof.

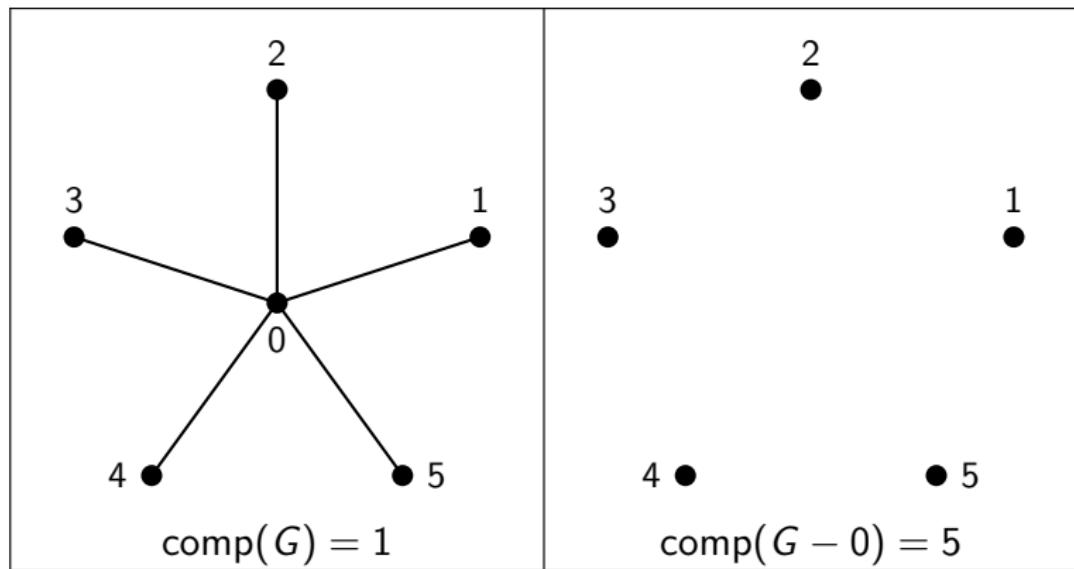
Let  $u, v$  be the ends of  $e$  ( $u = v$  if  $e$  is a loop). Note that  $G$  has a cycle containing  $e$ , iff  $G - e$  contains a path from  $u$  to  $v$ , iff  $u, v$  are in the same component of  $G - e$ . Now

- ▶ If  $u, v$  are in the same component of  $G - e$ , then  $H + e$  is a component of  $G$ , so  $\text{comp}(G - e) = \text{comp}(G)$ .
- ▶ If  $u, v$  are in distinct components, say  $H_1, H_2$  of  $G - e$ , then  $H_1 \cup H_2 + e$  is a component of  $G$ , so  $\text{comp}(G - e) = \text{comp}(G) + 1$ . □

# Connected Graph

## Definition

A vertex  $v \in V(G)$  is called a ***cut-vertex*** whose deletion increases the number of components.



# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

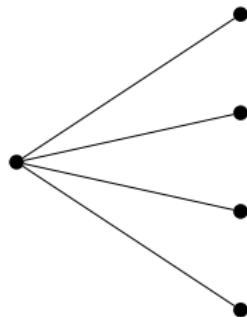
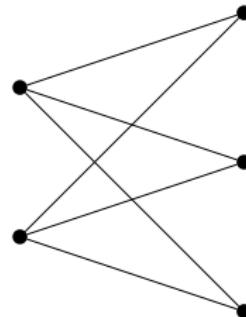
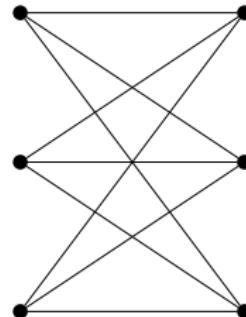
# Bipartition

## Definition

A **bipartition** of a graph  $G$  is a pair  $(A, B)$  where  $A, B \subset V(G)$  with  $A \cap B = \emptyset$ ,  $A \cup B = V(G)$  such that every edge has an end in  $A$  and an end in  $B$ .  $G$  is **bipartite** if it admits a bipartition.

## Definition

A **complete bipartite graph** or **biclique**, denoted  $K_{m,n}$ , is a simple bipartite graph with bipartition  $(A, B)$  with  $|A| = m$  and  $|B| = n$  such that every vertex in  $A$  is adjacent to every vertex in  $B$ .

 $K_{1,4}$  $K_{2,3}$  $K_{3,3}$

# Bipartition

## Theorem

For every graph  $G$ , TFAE

- (i)  $G$  is bipartite.
- (ii)  $G$  has no cycle of odd length.
- (iii)  $G$  has no closed walk of odd length.

## Proof.

**(i) $\Rightarrow$ (ii):** Assume that  $G = (A \cup B, E)$  is bipartite and let  $C \subset G$  be a cycle. Then every other vertex of  $C$  is in  $A$  and every other vertex is in  $B$ , hence  $C$  must have even length. (It takes even number of steps in a bipartite graph to return to the starting point.)

## Bipartition

### Proof (Cont.)

**(ii)  $\Rightarrow$  (iii):** We show the contrapositive, i.e.,  $\neg(\text{iii}) \Rightarrow \neg(\text{ii})$ . Let  $G$  have a closed walk of odd length, and choose such a walk  $v_0, e_1, v_1, \dots, v_n$  of **minimum** length. If there exist  $1 \leq i < j \leq n$  with  $v_i = v_j$ , then

- ▶ either  $j - i$  is odd and  $v_i, e_i, \dots, v_j$  is a shorter closed walk of odd length,
- ▶ or  $j - i$  is even and  $v_0, e_1 \dots v_i, e_{j+1}, v_{j+1}, \dots, v_n$  is a shorter closed walk of odd length.

It follows that  $v_1, \dots, v_n$  must be distinct ( $v_0 = v_n$ ), hence  $(\{v_1, \dots, v_n\}, \{e_1, \dots, e_n\})$  is an odd cycle.

## Bipartition

### Proof (Cont.)

(iii)  $\Rightarrow$  (i): Let  $G$  be a graph with no closed walk of odd length, w.l.o.g., we may assume that  $G$  is connected. Choose a “base point”  $u \in V(G)$ , observe that for every vertex  $v \in V(G)$ ,

- ▶ either all  $u, v$ -walks have even length,
- ▶ or all  $u, v$ -walks have odd length.

(Note that otherwise can concatenate an odd and an even walk to form a closed walk of odd length.) Now define

$$A := \{v \in V(G) \mid \exists u, v\text{-walk of even length}\}$$

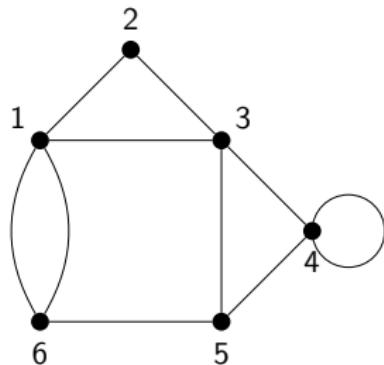
$$B := \{v \in V(G) \mid \exists u, v\text{-walk of odd length}\}$$

It follows that  $A \cap B = \emptyset$ . Since  $G$  is connected, we have  $A \cup B = V(G)$ . It follows that  $(A, B)$  is a bipartition of  $G$ , hence (i) is satisfied. □

# Induced Subgraph

## Definition

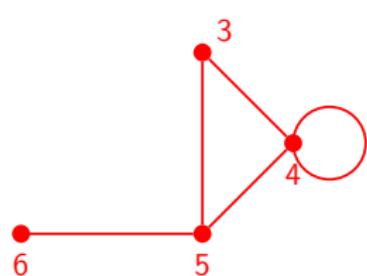
A subgraph  $H \subset G$  is **induced** if every edge of  $G$  with both ends in  $V(H)$  is in  $E(H)$ . Equivalently,  $H$  is induced if  $H = G - (V(G) \setminus V(H))$ .



$G$



induced in  $G$



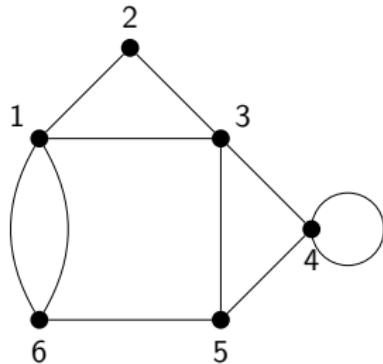
induced in  $G$

## Remark

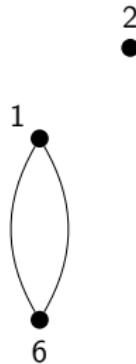
- An **induced path** is sometimes called a **snake**.
- An **induced cycle** is sometimes called a **chordless cycle** or a **hole**.

# Induced Subgraph

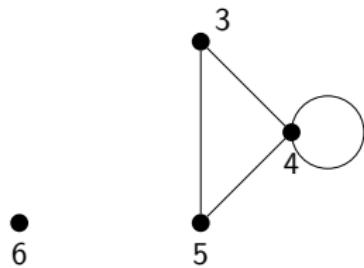
## NOT induced subgraph



$G$



NOT induced in  $G$



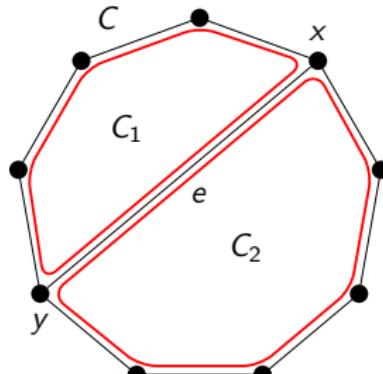
NOT induced in  $G$

# Bipartition

## Theorem

For every graph  $G$ , TFAE

- (i)  $G$  is bipartite.
- (ii)  $G$  has no cycle of odd length.
- (iii)  $G$  has no closed walk of odd length.
- (iv)  $G$  has no induced cycle of odd length.



## Proof.

(ii)  $\Rightarrow$  (iv). Immediate.

(iv)  $\Rightarrow$  (ii). We show the contrapositive, i.e.,  $\neg(\text{ii}) \Rightarrow \neg(\text{iv})$ . Suppose  $G$  has a cycle of odd length, choose a shortest cycle  $C \subset G$ . Note that  $C$  is induced, otherwise  $\exists e \in E(G) \setminus E(C)$ , with ends  $x, y$ . But now either  $C_1$  or  $C_2$  is an odd cycle of shorter length, contradiction.  $\square$

## Remark

The **girth** of a graph with a cycle is the length of its shortest cycle. A graph with no cycle has infinite girth.

# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Matching

## Definition

A **matching** in a graph  $G = (V, E)$  is a subset of edges  $M$  such that  $M$  does not contain a loop and no two edges in  $M$  are incident with a common vertex. (i.e., the graph  $(V, M)$  has all vertices of degree  $< 2$ )

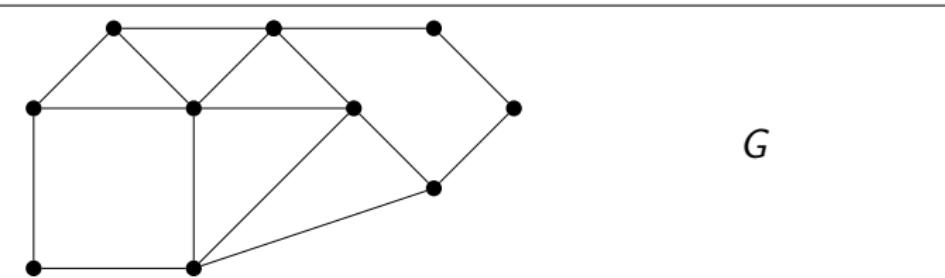
## Definition

- ▶ A matching  $M$  is **maximal** if there is no matching  $M'$  such that  $M \subsetneq M'$ .
- ▶ A matching  $M$  is **maximum** if there is no matching  $M'$  such that  $|M| < |M'|$ .
- ▶ A **perfect matching** is a matching  $M$  such that every vertex of  $G$  is incident with an edge in  $M$ .

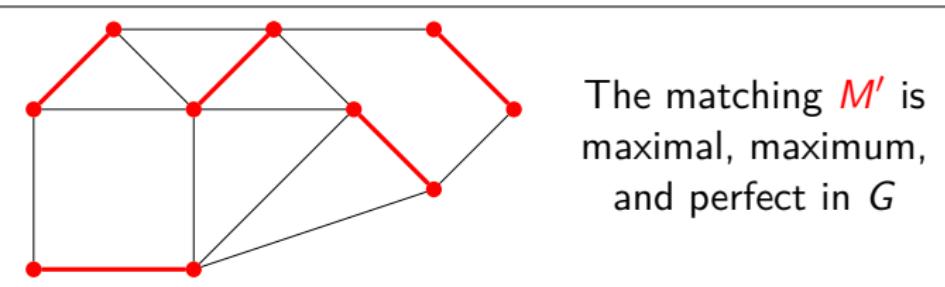
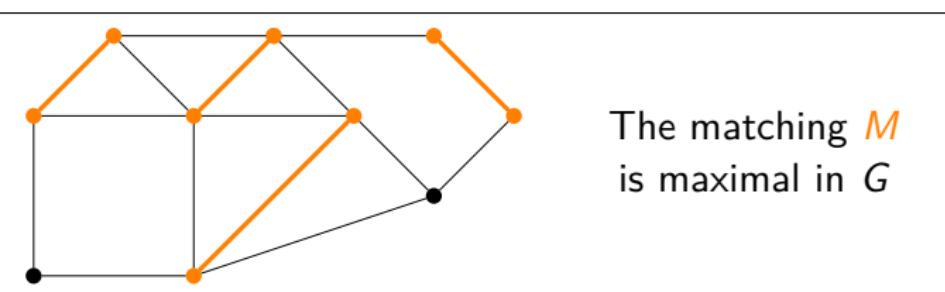
## Example

- ▶  $K_{n,n}$  has  $n!$  perfect matchings.
- ▶  $K_{2n+1}$  has 0 perfect matchings.
- ▶  $K_{2n}$  has  $(2n - 1)(2n - 3) \cdots (3)(1) = (2n - 1)!!$  perfect matchings.

# Matching



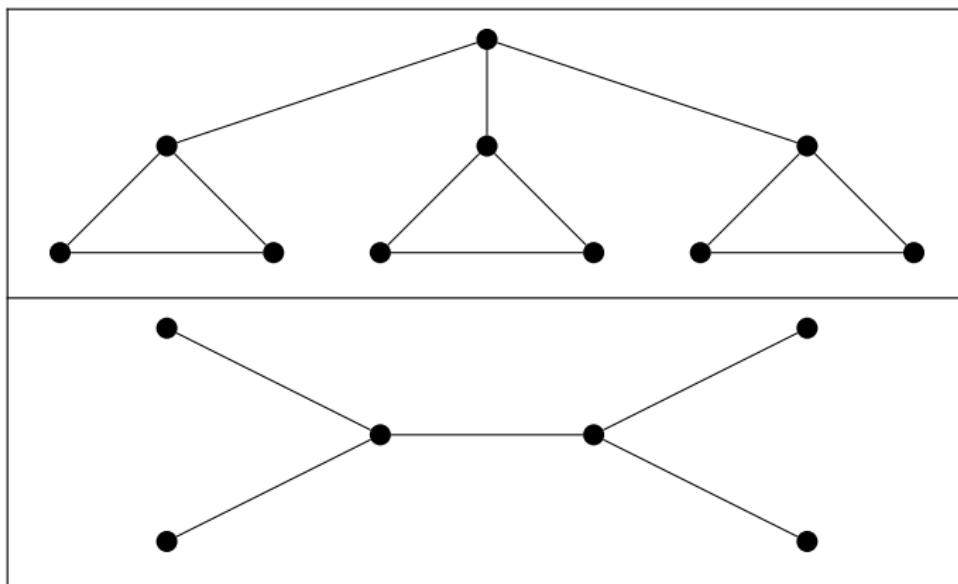
$G$



# Matching

## Remark

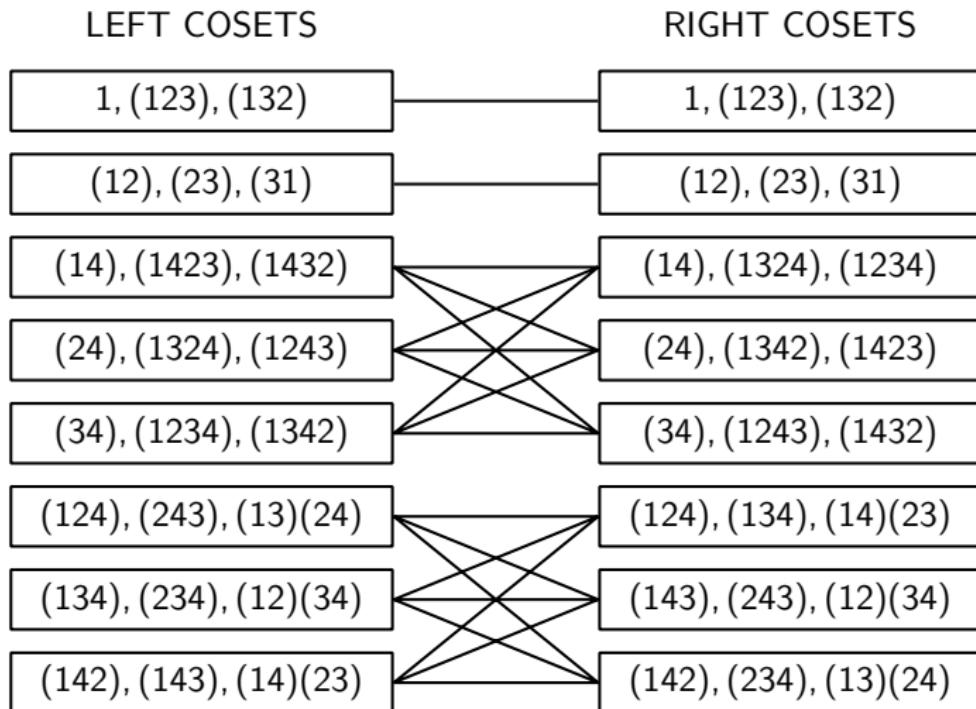
A necessary but not sufficient condition for a graph  $G$  to have a perfect matching is that  $|V(G)|$  is even.



# Group Transversals

## Example

Consider  $G = S_4$ , and subgroup  $H = \langle (123) \rangle \cong C_3$ .



# Group Transversals

## Definition

Let  $H, K \leq G$ . We define the coset intersection graph  $\Gamma_{H,K}^G$  to be a graph with vertex set consisting of all left cosets of  $H$ , i.e.,  $\{\ell_i H\}_{i \in I}$ , together with all right cosets of  $K$ , i.e.,  $\{K r_j\}_{j \in J}$ , where  $I, J$  are index sets. If a left coset of  $H$  and right coset of  $K$  correspond, they are included twice. Edges (undirected) are included whenever any two of these cosets have non-empty intersection, and an edge  $aH - Kb$  corresponds to the nonempty set  $aH \cap Kb$ .

## Theorem

*The coset intersection graph  $\Gamma_{H,K}^G$  is always a disjoint union of complete bipartite graphs.*

## Remark

The common transversals are given by a perfect matching in  $\Gamma_{H,K}^G$ .

## Group Transversals

Proof (Button, Choido, and Laris).

We first show that for  $a, b, c, d \in G$ , if  $aH-Kb-cH-Kd$  is a path in  $\Gamma_{H,K}^G$ , then there is an edge  $aH-Kd$ . Note that there exist  $h_1, h_2, h_3 \in H$  and  $k_1, k_2, k_3 \in K$  such that  $ah_1 = k_1b$ ,  $k_2b = ch_2$ ,  $ch_3 = k_3d$ . Rearranging yields  $c = k_3dh_3^{-1}$ , so  $b = k_2^{-1}k_3dh_3^{-1}h_2$ , so  $a = k_1k_2^{-1}k_3dh_3^{-1}h_2h_1^{-1}$ , and thus  $ah_1h_1^{-1}h_3 = k_1k_2^{-1}k_3d$ . Hence  $aH-Kd$  as required.

Take any  $\ell_iH$  and some  $Kr_j$  in the connected component of  $\ell_iH$  in  $\Gamma_{H,K}^G$ . There must be at least one such  $Kr_j$ , we show that  $\ell_iH$  and  $Kr_j$  are connected by an edge. For if not, then there must be at least one finite path of length  $> 1$  connecting them; take a minimal such path  $\gamma$  from  $\ell_iH$  to  $Kr_j$ . Then  $\gamma$  begins with  $\ell_iH-Ka-bH-Kc-\dots$ , where  $Ka \neq Kr_j$ . But we have showed previously that  $\ell_iH$  and  $Kr_j$  must be joined by an edge, contradicting the minimality of  $\gamma$ . So  $\ell_iH$  and  $Kr_j$  are joined by an edge for every  $Kr_j$  in the connected component of  $\ell_iH$ . □

## Group Transversals

Now suppose that  $G$  is finite with  $|H| = |K| = m$ . Since the connected component is isomorphic to  $K_{s,t}$  for some  $s, t \in \mathbb{N}$ , and note that the cosets of  $H$  (or  $K$ ) are disjoint and have the same size  $|H|$  (or  $|K|$ ), we have both  $s|H| \leq t|K|$  and  $t|H| \leq s|K|$ , hence  $s = t$ . Therefore a perfect matching exists for each component of  $\Gamma_{H,K}^G$ . Each perfect matching corresponds to a set of common transversals. The results also follow if  $H = K$ .

# Hall's Theorem

## Definition

If  $X \subset V(G)$ , the **neighbors** of  $X$  is

$$N(X) := \{v \in V(G) \setminus X \mid v \text{ is adjacent to a vertex in } X\}$$

For simplicity, we write  $N(x) := N(\{x\})$ .

## Definition

The edges  $S \subset E(G)$  **covers**  $X \subset V(G)$  if every  $x \in X$  is incident to some  $e \in S$ .

## Definition

The vertices  $X \subset V(G)$  **covers**  $S \subset E(G)$  if every  $e \in S$  is incident to some  $v \in X$ .

## Theorem (Hall)

Let  $G$  be a bipartite graph with bipartition  $(A, B)$ . There exists a matching covering  $A$  iff there does not exist  $X \subset A$  with  $|N(X)| < |X|$ .

# Hall's Theorem

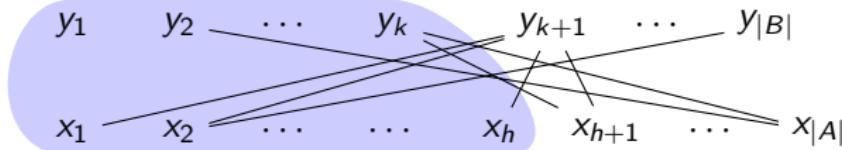
Proof via Dilworth's Theorem.

**Necessity.** Immediate by pigeonhole principle.

**Sufficiency.** Let  $G = (A \cup B, E)$  be a bipartite graph satisfying Hall's condition that  $|N(X)| \geq |X|$  for all  $X \subset A$ . Define a poset  $(P, \leq)$  by letting  $P = A \cup B$ , and  $x < y$  if  $x \in A$ ,  $y \in B$ , and  $xy \in E$ . Suppose that the largest antichain is  $S = \{x_1, \dots, x_h, y_1, \dots, y_k\}$ , then

$$N(\{x_1, \dots, x_k\}) \subset B \setminus \{y_1, \dots, y_k\}$$

(for otherwise  $S$  would not be an antichain if  $y \in \{y_1, \dots, y_k\}$  were the neighbor of some  $x \in \{x_1, \dots, x_k\}$ .) Thus Hall's condition implies  $|B| - k \geq h$ , i.e.,  $|B| \geq k + h$ .



## Hall's Theorem

### Proof (Cont.)

By Dilworth's theorem,  $P$  can be partitioned into  $k + h$ , denote the matching by  $M$ , then

$$|M| + (|A| - |M|) + (|B| - |M|) = k + h \leq |B|$$

that is,

$$|A| + |B| - |M| \leq |B|$$

thus

$$|M| \geq |A|$$

i.e., there is a matching  $M$  covering  $A$ .

□

# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Trees

## Definition

A **forest** is a graph with no cycles. A **tree** is a connected forest.

## Theorem

If  $G$  is a forest, then  $\text{comp}(G) = |V(G)| - |E(G)|$ . In particular, if  $T$  is a tree, then  $|V(T)| = |E(T)| + 1$ .

## Proof.

Induction on  $|E(G)|$ .

**Base case.** If  $|E(G)| = 0$ ,  $G$  has no edge, thus  $\text{comp}(G) = |V(G)|$ .

**Inductive case.**  $|E(G)| > 0$ . Choose  $e \in E(G)$ , since  $G$  has no cycle, then  $e$  is a cut-edge, thus

$$\begin{aligned}\text{comp}(G) &= \text{comp}(G - e) - 1 \\ &= |V(G - e)| - |E(G - e)| - 1 \quad (\text{by IH}) \\ &= |V(G)| - |E(G)|\end{aligned}$$



# Trees

## Definition

A **leaf** is a vertex of degree 1.

## Theorem

Let  $T$  be a tree with  $|V(T)| \geq 2$ , then  $T$  has at least 2 leaves, and if there are only 2 leaves, then  $T$  is a path.

## Proof.

Note that  $2 = 2|V(T)| - 2|E(T)| = \sum_{v \in V(T)} (2 - \deg(v))$ . Since  $T$  is connected, and  $|V(T)| \geq 2$ , all vertices have degree  $> 0$ . This means there are at least 2 leaves.

Further if there are exactly 2 leaves, then all other vertices have degree 2, therefore  $T$  is a path. (Take any maximal path in  $T$ , note that any extra edge would increase the degree of interior vertices to 3, or increases the degree of leaves to 2) □

# Trees

## Lemma

If  $T$  is a tree and  $v$  is a leaf, then  $T - v$  is a tree.

### Proof.

Observe that  $T - v$  has no cycle and is connected. □

## Theorem

If  $T$  is a tree, and  $u, v \in V(T)$ , then there is a unique  $u, v$ -path.

### Proof.

Induction on  $|V(T)|$ .

**Base case.**  $|V(T)| = 1$ . We have  $u = v$ .

**Inductive case.**

- ▶ If there is a leaf  $w \neq u, v$ , apply induction to  $T - w$ ;
- ▶ Otherwise  $T$  is a path with ends  $u, v$ , which is unique by IH. □

# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Spanning Trees

## Definition

If  $T$  is a subgraph of a graph  $G$ , and  $T$  is a tree with  $V(T) = V(G)$ , then we call  $T$  a **spanning tree** of  $G$ .

## Notation

If  $G$  is a graph,  $H \subset G$  and  $e \in E(G)$ , then we define  $H + e$  to be the subgraph of  $G$  obtained from  $H$  by adding  $e$  and its ends.

# Spanning Trees

## Theorem

Let  $G$  be a connected graph with  $|V(G)| \geq 2$ . If  $H$  is a subgraph satisfying

- (i) either  $H$  is minimal such that  $V(H) = V(G)$  and  $H$  is connected,
- (ii) or  $H$  is maximal such that  $H$  has no cycles,

then  $H$  is a spanning tree of  $G$ .

## Proof.

- (i) It suffices to show that  $H$  is a tree. Suppose that  $H$  has a cycle  $C$ , choose  $e \in E(C)$ . Now  $H - e$  is connected (b/c  $e$  is not a cut-edge), but this contradicts that  $H$  is minimal.
- (ii) Note that  $V(H) = V(G)$  by maximality of  $H$ . It remains to show that  $H$  is connected. Suppose not, choose a partition  $\{X, Y\}$  of  $V(H) = V(G)$  such that no edge of  $H$  has one end in  $X$  and the other in  $Y$ . Choose  $e \in E(G)$  such that  $e$  has one end in  $X$  and the other in  $Y$  (b/c  $G$  connected), but now  $H + e$  contradicts that  $H$  is maximal.



# Trees

## Theorem

If  $|V(G)| = |E(G)| + 1$ , and

- (i) either  $G$  has no cycles,
- (ii) or  $G$  is connected,

then  $G$  is a tree.

## Proof.

- (i) Since  $G$  is a forest, then  $1 = |V(G)| - |E(G)| = \text{comp}(G)$ .
- (ii) Choose a spanning tree  $T$  of  $G$  (possible b/c  $G$  connected), then

$$|E(G)| = |V(G)| - 1 = |V(T)| - 1 = |E(T)|$$

thus  $G = T$ , so  $G$  is a tree.



# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Weighted Graph

## Definition

A **weighted graph** is a graph  $G$  with a weight function  $w : E(G) \rightarrow \mathbb{R}$ . A **minimum-cost tree** (or **minimum-weight spanning tree**) of  $G$  is a spanning tree  $T$  for which

$$\sum_{e \in E(T)} w(e)$$

is minimum.

# Kruskal's Algorithm

## Kruskal's Algorithm

- ▶ Input: A connected weighted graph  $G = (V, E)$ .
- ▶ Output: A minimum-cost tree  $T$ .
- ▶ Procedure: Choose a sequence of edges  $e_1, e_2, \dots, e_m$  according to the rule that  $e_i$  is an edge of minimum weight in  $E(G) \setminus \{e_1, \dots, e_{i-1}\}$  so that  $\{e_1, \dots, e_{i-1}\}$  does not contain the edge set of a cycle. When no such edge exists, stop and return the subgraph  $T = (V, \{e_1, \dots, e_m\})$ .

# Table of Contents

1. Basic Graph Theory

2. Connectivity

3. Bipartite Graph

4. Matching

5. Trees

6. Spanning Trees

7. Kruskal's Algorithm

8. Dijkstra's Algorithm

# Distance Function

## Definition

Given graph  $G$ , and  $u, v \in V(G)$ , the distance from  $u$  to  $v$ , denoted  $\text{dist}(u, v)$ , is the shortest length of a walk from  $u$  to  $v$  in  $G$ .

## Remark

For  $u, v, w \in V(G)$ , the triangle inequality holds,

$$\text{dist}(u, v) + \text{dist}(v, w) \geq \text{dist}(u, w).$$

## Weighted Distance

### Definition

Given a simple connected graph with weight function  $w : E(G) \rightarrow \mathbb{R}_{\geq 0}$ , the length of a walk  $v_1 e_1 v_2 e_2 \cdots e_k v_{k+1}$  is given by

$$w(e_1) + w(e_2) + \cdots + w(e_k).$$

Then the distance from  $u$  to  $v$  is the length of the shortest walk from  $u$  to  $v$ .

### Definition

Given graph  $G$ , and  $r \in V(G)$ , a tree  $T \subset G$  with  $r \in V(T)$  is a **shortest path tree** for  $r$  if

$$\text{dist}_G(r, v) = \text{dist}_T(r, v)$$

for every  $v \in V(T)$ .

# Dijkstra's Algorithm

## Dijkstra's Algorithm

- ▶ Input: A simple connected graph  $G = (V, E)$  with root vertex  $r$  and nonnegative weight function  $w : E(G) \rightarrow \mathbb{R}_{\geq 0}$ .
- ▶ Output: A shortest path spanning tree for  $r$ .
- ▶ Procedure:
  1.  $i = 1$ . Set  $T_1$  to be the tree consisting of only the root vertex  $r$ .
  2.  $i \geq 2$ . Choose an edge  $uv$  such that  $u \in V(T_{i-1})$ ,  
 $v \in V(G) \setminus V(T_{i-1})$ , and  $\text{dist}_T(r, u) + w(uv)$  is minimum. Let  
 $T_i := T_{i-1} + uv$ . If no such choice is possible, return the present  
tree.