# UMA205: Introduction to Algebraic Structures

Naman Mishra

January 2024

# Contents

# The Course

**Lecture 01.**
Wed 03 Jan '24

**Instructor:** Prof. Arvind Ayyer
**Office:** X-15
**Office hours:** TBD
**Lecture hours:** MWF 11:00–11:50
**Tutorial hours:** Tue 9:00–9:50

80% attendance is mandatory.

**Prerequisites:** UMA101 and UMA102 **Texts:** Several

- *Analysis I*, 3rd edition, Terence Tao.
- *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory*, 3rd edition, Miklos Bona.

## Grading

(20%) Quizzes on alternate Tuesdays, worst dropped. No makeup quizzes, but if a quiz is missed for a medical reason (with certificate), that quiz will be dropped.

(30%) Midterm

(50%) Final

Homeworks after every class, ungraded. Exams are closed book and closed notes, with no electronic devices allowed.

## Aims of the Course

- Deal with formal mathematical structures.
- Learning the axiomatic method.

- See how more complicated structures arise from simpler ones.

# Chapter I

# Peano's Axioms

We try to formulate two fundamental quantities: 0 and the successor function $n \mapsto n_{++}$.

(P1) 0 is a natural number.

(P2) If $n$ is a natural number, so is $n_{++}$.

(P3) 0 is not the successor of any natural number.

(P4) Different natural numbers have different successors.

(P5) (Principle of mathematical induction) Let $P(n)$ be any "property" for a natural number $n$. Suppose that $P(0)$ is true, and that $P(n_{++})$ is true whenever $P(n)$ is true. Then $P$ is true for all natural numbers.

Denote *the* set of natural numbers by $\mathbb{N}$. (Any two sets satisfying the Peano axioms are isomorphic.) Note that $\mathbb{N}$ is itself infinite, but all of its elements are finite.

*Proof.* 0 is finite. If $n$ is finite, then $n_{++}$ is finite. Thus, by induction, all natural numbers are finite. (But wtf is a finite number?) □

*Remark.*

- There exist number systems which admit infinite numbers. For example, cardinals, ordinals, etc.
- This way of thinking is *axiomatic*, but not constructive.

**Lecture 02.**
Mon 08 Jan '24

## I.1 Addition

**Definition I.1.1** (Addition)**.** Suppose $m, n \in \mathbb{N}$. We define the binary operation $+$ by setting $0 + m = m$. Suppose we have defined $n + m$. Then we inductively define $n_{++} + m = (n + m)_{++}$.

For example, note that $1 + m = (0 + m)_{++} = m_{++}$.

**Lemma I.1.2.** *For $n \in \mathbb{N}$, we have $n + 0 = n$.*

*Proof.* $0 + 0 = 0$. If $n + 0 = n$, then $n_{++} + 0 = (n + 0)_{++} = n_{++}$. □

**Lemma I.1.3.** *For $m, n \in \mathbb{N}$, we have $n + m_{++} = (n + m)_{++}$.*

*Proof.* Fix $m$ and induct on $n$. For $n = 0$, we have $0 + m_{++} = m_{++} = (0 + m)_{++}$. Suppose $n + m_{++} = (n + m)_{++}$. Then

$$\begin{aligned} n_{++} + m_{++} &= (n + m_{++})_{++} && \text{(definition)} \\ &= ((n + m)_{++})_{++} && \text{(hypothesis)} \\ &= (n_{++} + m)_{++} && \text{(definition)} \end{aligned}$$

as desired. □

**Exercise I.1.4.** *(Commutativity) For $m, n \in \mathbb{N}$, we have $n + m = m + n$.*

*Proof.* Fix $m$ and induct on $n$. For $n = 0$, we have $0 + m = m = m + 0$. Suppose $n + m = m + n$. Then

$$\begin{aligned} n_{++} + m &= (n + m)_{++} && \text{(definition)} \\ &= (m + n)_{++} && \text{(hypothesis)} \\ &= m + n_{++} \end{aligned}$$

by the previous lemma. □

**Exercise I.1.5.** *(Associativity) For $m, n, p \in \mathbb{N}$, we have $(m + n) + p = m + (n + p)$.*

*Proof.* Induct on $m$. $(0 + n) + p = n + p = 0 + (n + p)$.
Suppose $(m + n) + p = m + (n + p)$. Then

$$\begin{aligned} (m_{++} + n) + p &= (m + n)_{++} + p && \text{(definition)} \\ &= ((m + n) + p)_{++} && \text{(definition)} \\ &= (m + (n + p))_{++} && \text{(hypothesis)} \\ &= m_{++} + (n + p). && \text{(definition)} \end{aligned}$$

This closes the induction. □

**Exercise I.1.6.** *(Cancellation) For $m, n, p \in \mathbb{N}$, if $m + n = m + p$, then $n = p$.*

*Proof.* Induct on $m$. $0 + n = 0 + p$ implies $n = p$.

Suppose $m + n = m + p$ implies $n = p$. Then $m_{++} + n = m_{++} + p$ implies $(m + n)_{++} = (m + p)_{++}$ and so $m + n = m + p$ by (P4). By the inductive hypothesis, $n = p$. □

**Definition I.1.7** (Positive)**.** A natural number is positive if it is not 0.

**Proposition I.1.8.** *If $a$ is positive and $b \in \mathbb{N}$, then $a + b$ is positive.*

*Proof.* Induct on $b$. $a + 0 = a$ is positive. $a + b_{++} = (a + b)_{++}$ is positive since 0 is not the successor of any natural number. □

**Exercise I.1.9.** *If $m, n$ in $\mathbb{N}$ with $m + n = 0$, then $m = n = 0$.*

*Proof.* Contrapositive of the previous proposition, with commutativity. □

**Exercise I.1.10.** *Let $a$ be positive. Then there exists a unique $b \in \mathbb{N}$ such that $a = b_{++}$.*

*Proof.* Let $P(n)$ be that $n$ is zero or there exists a unique $b \in \mathbb{N}$ such that $n = b_{++}$. $P(0)$ is true.

Suppose $P(n)$ is true. $n_{++}$ is non-zero, successor of $n$ and only $n$, by (P3) and (P4). Thus $P(n_{++})$ is true. □

**Definition I.1.11** (Order)**.** Let $m, n \in \mathbb{N}$. We say that $n$ is greater than or equal to $m$, written $n \geq m$ or $m \leq n$, if $n = m + a$ for some $a \in \mathbb{N}$.

Similarly, we say that $n$ is (strictly) greater than $m$, written $n > m$ or $m < n$, if $n \geq m$ and $n \neq m$.

Note that $n_{++} > n$, so there is no largest natural number.

**Proposition I.1.12.** *Let $a, b, c \in \mathbb{N}$. Then*

*1) $a \geq a$ (reflexivity),*

*2) $a \geq b$ and $b \geq a$ implies $a = b$ (antisymmetry),*

*3) $a \geq b$ and $b \geq c$ implies $a \geq c$ (transitivity),* *4)* $a \geq b \iff a + c \geq b + c$*,*

*5)* $a > b \iff a \geq b_{++}$*,*

*6)* $a > b \iff a = b + c$ *for some positive* $c$*.*

*Proof.*

1) $a = a + 0$.

2) $a = b + c$ and $b = a + d$ implies $a = a + (c + d)$. By cancellation, $c + d = 0$ and so $c = d = 0$.

3) $a = b + m$ and $b = c + n$ implies $a = c + (m + n)$.

4) $a = b + m \iff (a + c) = (b + c) + m$.

5) From 6), $a > b \iff a = b + c$ for some positive $c$, iff $a = b + d_{++} = b_{++} + d$.

6) $a > b \iff a = b + c$ but $a \neq b$. Since $a \neq b$, $c$ cannot be zero. Conversely, if $c$ is positive, $a \neq b$. □

**Proposition I.1.13** (Trichotomy)**.** *Let* $a, b \in \mathbb{N}$*. Then exactly one of the following holds:* $a > b$*,* $a = b$*, or* $a < b$*.*

*Proof.* We first prove that no more than one of the three holds. $a = b$ cannot hold simultaneously with $a > b$ or $a < b$ by their definitions. Suppose $a > b$ and $a < b$. Then $a = b + c$ and $b = a + d$ for some positive $c$ and $d$. Thus $a = a + (c + d)$ and so $c + d = 0$, a contradiction.

We now prove that at least one of the three holds by induction on $a$. Since $b = 0 + b$, either $0 = b$ or $b > 0$. Suppose at least one of $a \geq b$ and $a < b$ holds. If $a = b + c$, then $a_{++} = b + (c_{++})$ and so $a_{++} > b$. If $a < b$, then by proposition I.1.12(5), $a_{++} \leq b$. This completes the induction. □

**Proposition I.1.14** (Strong induction)**.** *Let* $m_0 \in \mathbb{N}$ *and let* $P(m)$ *be a property for all* $m \in \mathbb{N}$*. Suppose for all* $m \geq m_0$*, we have the following: if* $P(m')$ *holds for all* $m_0 \leq m' < m$*, then* $P(m)$ *holds. Then* $P(m)$ *holds for all* $m \geq m_0$*.*

Note that the inductive step is vacuously true for $m = m_0$.

*Proof.* Define $Q(m)$ to be "$P(m')$ holds for all $m_0 \leq m' < m$". $Q(0)$ holds vacuously, since there are no $m' < 0$.

Suppose $Q(m)$ holds. If $m < m_0$, then $Q(m_{++})$ holds vacuously, since $m_{++} \leq m_0$ and so no $m'$ satisfies $m_0 \leq m' < m_{++} \leq m_0$.

Now if $m \geq m_0$, then $Q(m)$ and the proposition imply $P(m)$. Thus $P(m')$ holds for all $m_0 \leq m' \leq m \iff m_0 \leq m' < m_{++}$. Thus $Q(m_{++})$ holds. □

**Exercise I.1.15** (Backwards induction)**.** *Let $m_0 \in \mathbb{N}$, and let $P(m)$ be a property pertaining to the natural numbers such that whenever $P(m_{++})$ is true, then $P(m)$ is true. Suppose that $P(m_0)$ is also true. Prove that $P(m)$ is true for all natural numbers $m \leq m_0$.*

*Proof.* Define $Q(m)$ to be "if $P(m)$ is true, then $P(m')$ is true for all $m' \leq m$". $Q(0)$ holds vacuously, since $m' \leq 0$ implies $m' = 0$.

Suppose $Q(m)$ holds. Then if $P(m_{++})$ is true, so is $P(m)$, and by the inductive hypothesis, $P(m')$ is true for all $m' \leq m$. Thus $Q(m_{++})$ holds. Thus $Q(m)$ holds for all $m \in \mathbb{N}$.

In particular, $Q(m_0)$ holds, and so $P(m')$ is true for all $m' \leq m_0$. □

From now on, we will assume the usual laws of addition.

## I.2 Multiplication

**Definition I.2.1** (Multiplication)**.** Let $m \in \mathbb{N}$. The binary operation multiplication, denoted by $*$, is defined as follows. Set $0 * m = 0$. Then define it inductively as follows. If we know $n * m$, set $n_{++} * m = (n * m) + m$.

**Lemma I.2.2.** *Let $m, n \in \mathbb{N}$, Then $m * n = n * m$.*

*Proof.* First note that $m * 0 = 0$, since $0 * 0 = 0$ and $m_{++} * 0 = m * 0 + 0 = m * 0$.

Next note that $n * m_{++} = (n * m) + n$, since $0 * m_{++} = 0 = (0 * m) + 0$, and $n_{++} * m_{++} = (n * m_{++}) + m_{++}$ which is equal to $(n * m) + n + m_{++} = (n * m) + m + n_{++} = (n_{++} * m) + n_{++}$ by the inductive hypothesis.

Finally, $0 * n = n * 0$, and $m_{++} * n = n * m_{++}$ gives $m * n = n * m$ by induction on $m$. □

We use the notation $mn$ for $m * n$ and also employ the usual convention for precedence, so that $mn + p$ means $(m * n) + p$ and not $m * (n + p)$.

**Lemma I.2.3.** *Let $m, n \in \mathbb{N}$. Then $mn = 0$ iff at least one of $m$ and $n$ is $0$.*

*Proof.* The 'if' direction is clear. Suppose $m$, $n$ are positive. Then $m = \tilde{m}_{++}$ for some $\tilde{m} \in \mathbb{N}$.

$$\begin{aligned} mn &= (\tilde{m}_{++})n \\ &= (\tilde{m}n) + n \end{aligned}$$

which is positive since $n$ is positive. □

**Proposition I.2.4** (Distributivity)**.** *For $a, b, c \in \mathbb{N}$, we have $a(b + c) = ab + ac$ and $(b + c)a = ba + ca$.*

*Proof.* Prove the first by induction on $a$. $0(b + c) = 0 = 0 + 0 = 0b + 0c$.

Suppose $a(b + c) = ab + ac$. Then

$$\begin{aligned} a_{++}(b + c) &= a(b + c) + (b + c) && \text{(definition)} \\ &= (ab + ac) + (b + c) && \text{(hypothesis)} \\ &= (ab + b) + (ac + c) \\ &= a_{++}b + a_{++}c. && \text{(definition)} \end{aligned}$$

The second equality follows from the first by commutativity. □

**Exercise I.2.5.** *(Associativity) For $a, b, c \in \mathbb{N}$, we have $(ab)c = a(bc)$.*

*Proof.* Induct on $a$. $(0b)c = 0c = 0 = 0(bc)$.

Suppose $(ab)c = a(bc)$. Then

$$\begin{aligned} (a_{++}b)c &= (ab + b)c && \text{(definition)} \\ &= (ab)c + bc && \text{(distributivity)} \\ &= a(bc) + bc && \text{(hypothesis)} \\ &= a_{++}(bc) \end{aligned}$$

by definition. □

**Exercise I.2.6.** *(Order preservation) For $a, b, c \in \mathbb{N}$ with $a < b$ and $c \neq 0$, we have $ac < bc$.*

*Proof.* Induct on $c$ with base case $c = 1$. If $ac < bc$, then $ac + a < bc + a$ but $bc+a < bc+b$, both by order preservation under addition. By transitivity, $ac+a < bc+b$ and so $ac_{++} < bc_{++}$. □

**Exercise I.2.7.** *(Cancellation) For $a, b, c \in \mathbb{N}$ with $ac = bc$ and $c \neq 0$, we have $a = b$.*

*Proof.* From trichotomy and order preservation. □

**Proposition I.2.8** (Euclidean algorithm)**.** *Let $n \in \mathbb{N}$ and $m$ be positive. Then there exist unique $q, r \in \mathbb{N}$ such that $n = qm + r$ and $r < m$. We call $q$ the quotient and $r$ the remainder.*

*Proof.* We first prove uniqueness. Suppose $n = qm + r = q'm + r'$. If $q < q' \iff q_{++} \leq q'$, then $qm+r < qm+m = q_{++}m \leq q'm \leq q'm+r'$, a contradiction. Similarly, $q' < q$ is impossible. This leaves $q = q'$. Then $qm + r = q'm + r'$ gives $r = r'$ by cancellation.

For existence, we induct on $n$. $0 = 0m + 0$. Suppose $n = qm + r$. Then $n_{++} = qm+r_{++}$. If $r_{++} < m$, we are done. Otherwise, $r_{++} = m$ (since $r < m \iff r_{++} \leq m$) and so $n_{++} = (q_{++})m + 0$. □

This proposition allows us to divide.

**Definition I.2.9** (Exponentiation)**.** Let $m$ be positive. The binary operation exponentiation can be defined inductively as $m^0 = 1$ and $m^{n_{++}} = m^n m$. We further define $0^k = 0$ for all positive $k$.

# Chapter II

# From Sets to the Rationals

**Lecture 03.**
Mon 08 Jan '24

## II.1 Axioms of Set Theory (ZFC)

**Definition II.1.1** (Set)**.** A set is a well-defined collection of objects, which we call elements. We will write $x \in A$ to say that $x$ is an element of $A$.

Well-defined means that given any object, we can state without ambiguity whether it is an element of the set or not.

**Axiom II.1.** Sets are themselves objects. If $A$ and $B$ are sets, it is meaningful to ask whether $A$ is an element of $B$.

**Axiom II.2** (Extensionality)**.** Two sets $A$ and $B$ are equal, written $A = B$, if every element of $A$ is a member of $B$ and vice versa.

**Axiom II.3** (Existence)**.** There exists a set, denoted by $\varnothing$ or $\{\}$, known as the empty set, which does not contain any elements, *i.e.*, $x \notin \varnothing$ for all objects $x$.

**Problem 0.1.** $\varnothing$ *is unique.*

*Proof.* Suppose $\varnothing$ and $\varnothing'$ are both empty sets. Then $x \in \varnothing \iff x \in \varnothing'$ since both are always false. □

**Lemma II.1.2** (Single choice)**.** *Let $A$ be a non-empty set. Then there exists an object $x$ such that $x \in A$.*

*Proof.* If not, then $x \notin A$ for all objects $x$ and so $A = \varnothing$. □

Thus, we can choose an element of $A$ to certify its non-emptiness.

**Axiom II.4** (Pairing)**.** If $a$ is an object, there exists a set, denoted $\{a\}$, whose only element is $a$. Similarly, if $a$ and $b$ are objects, there exists a set, denoted $\{a, b\}$, whose only elements are $a$ and $b$.

For example, we can now construct $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$, $\{\varnothing, \{\varnothing\}\}$, etc, all of which are distinct.

**Axiom II.5** (Pairwise union)**.** Given sets $A$ and $B$, there exists a set, denoted $A \cup B$, called the union of $A$ and $B$, which consists of exactly the elements in $A$, $B$, or both.

**Problem 0.2.** $A \cup B = B \cup A$.

*Proof.* By commutativity of $\vee$. □

**Problem 0.3.** $(A \cup B) \cup C = A \cup (B \cup C)$.

*Proof.* By associativity of $\vee$. □

**Definition II.1.3** (Subset)**.** $A$ is a subset of $B$ if every element of $A$ is alaso an element of $B$, denoted $A \subseteq B$.

**Axiom II.6** (Specification)**.** (also called Separation). Let $A$ be a set and let $P(x)$ be a property for every $x \in A$. Then there exists a set $S = \{x \in A \mid P(x)\}$ where $x \in S$ iff $x \in A$ and $P(x)$ is true.

We can now define the intersection, $A \cap B$, and difference, $A \setminus B$, of sets $A$ and $B$.

**Definition II.1.4.** Let $A$ and $B$ be sets. we define the intersection $A \cap B = \{x \in A \mid x \in B\}$ and the difference $A \setminus B = \{x \in A \mid x \notin B\}$.

$A$ and $B$ are said to be disjoint if $A \cap B = \varnothing$.

Recall that sets form a Boolean algebra under the operations $\cup$, $\cap$, and $\setminus$. For example, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, de Morgan's laws, etc.

**Axiom II.7** (Replacement)**.** Let $A$ be a set and let $P(x, y)$ be a property for every $x \in A$ and every object $y$, such that for every $x \in A$ there is at most one $y$ for which $P(x, y)$ is true. Then there exists a set $S = \{y \mid P(x, y) \text{ is true for some } x \in A\}$. That is, $y \in S$ iff $P(x, y)$ is true for some $x \in A$.

*Examples.*

- Let $A = \{7, 9, 22\}$ and $P(x, y) \equiv y = x_{++}$. Then $S = \{8, 10, 23\}$.
- Let $A = \{7, 9, 22\}$ and $P(x, y) \equiv y = 1$. Then $S = \{1\}$.

**Axiom II.8** (Infinity)**.** There exists a set, denoted $\mathbb{N}$, whose objects are called natural numbers, *i.e.*, an object $0 \in \mathbb{N}$, and $n_{++}$ for every $n \in \mathbb{N}$, such that the Peano axioms hold.

**Axiom II.9** (Foundation)**.** (also called Regularity). If $A$ is a non-empty set, then there exists at least one $x \in A$ which is either not a set or is disjoint from $A$.

For example, if $A = \{\{1, 2\}, \{1, 2, \{1, 2\}\}\}$, then $\{1, 2\}$ is an element of $A$ which is disjoint from $A$.

**Lecture 04.**
Wed 10 Jan '24

**Definition II.1.5** (Cartesian product)**.** Let $A$ and $B$ be sets. Then $A \times B = \{(a, b) \mid a \in A, b \in B\}$ is called the *Cartesian product* of $A$ and $B$.

This exists by virtue of the axiom of powers (II.10).
We recall the definition of a relation and some properties.

**Definition II.1.6** (Relation)**.** Let $A$ and $B$ be sets. Then a subset $R$ of $A \times B$ is called a (binary) *relation* from $A$ to $B$. If $B = A$, we say that $R$ is a relation on $A$.

**Definition II.1.7.** Let $R$ be a relation on a set $A$. We say that $R$ is

(i) **reflexive** if $(a, a) \in R$ for all $a \in A$,

(ii) **symmetric** if $(a, b) \in R \implies (b, a) \in R$,

(iii) **antisymmetric** if $(a, b) \in R \wedge (b, a) \in R \implies a = b$,

(iv) **transitive** if $(a, b) \in R \wedge (b, c) \in R \implies (a, c) \in R$.

If $R$ satisfies (i), (ii) and (iv), it is said to be an *equivalence relation.* We write $a \sim_R b$ for $(a, b) \in R$.

If $R$ satisfies (i), (iii) and (iv), it is a partial order. We write $a \leq_R b$ or $a \geq_R b$ for $(a, b) \in R$.

**Definition II.1.8** (Equivalence class)**.** Let $X$ be a set and $\sim_R$ an equivalence relation on $X$. The equivalence class associated with $x \in X$ is

$$[x] = \{y \in X \mid y \sim_R x\}.$$

**Definition II.1.9** (Partition)**.** A (set) *partition* of a set $X$ is a family $\{X_\alpha \mid \alpha \in I\}$, where $I$ is some indexing set, such that,

(i) $X_\alpha \cap X_\beta = \varnothing$ for all $\alpha \neq \beta \in I$,

(ii) $\bigcup_{\alpha \in I} X_\alpha = X$.

This is also written as simply

$$\bigsqcup_{\alpha \in I} X_\alpha = X.$$

**Proposition II.1.10** (Fundamental theorem of equivalence relations)**.** *Let $X$ be a set and $\sim_R$ an equivalence relation on $X$. Then the family of equivalence classes $\{[x] \mid x \in X\}$ forms a partition of $X$. Conversely, every partition arises from an equivalence relation.*

*Proof.* Exercise for the reader. □

**Definition II.1.11.** Let $X$ be a set and $\sim_R$ an equivalence relation on $X$. Then the set $X/\sim_R = \{[x] \mid x \in X\}$ is called the *quotient set* of $X$ by $R$.

*Examples.*

- Consider $\mathbb{N}$ with the relation $a \sim_R b \iff a \equiv b \pmod 3$. The quotient set $\mathbb{N}/R$ is $\{[0],[1],[2]\}$, which is morally the same as $\{0,1,2\}$.
- For any set $A$ with the equality relation $=$, the quotient set $A/\!=$ is (morally) the same as $A$.
- Consider $\mathbb{R}^2$ with $(x,y) \sim (z,w)$ if $x^2+y^2 = z^2+w^2$. Then $\mathbb{R}^2/\!\sim\, = \{[(r,0)] \mid r \in \mathbb{R}\}$ which is morally just the set of non-negative reals.

**Definition II.1.12** (Function)**.** Let $A$ and $B$ be sets. A relation $f$ from $A$ to $B$ is said to be a *function* if for all $a \in A$, there exists a unique $b \in B$ such that $(a,b) \in f$.

$A$ is said to be the *domain*, $B$ is said to be the *range* or *codomain* of $f$. For a subset $C \subseteq A$, the image of $C$ under $f$ is $f(C) = \{f(a) \mid a \in C\}$.

For a subset $D \subseteq B$, the *preimage* or *inverse image* of $D$ under $f$ is $f^{-1}(D) = \{a \in A \mid f(a) \in D\}$.

Note that $f(C)$ exists by the axiom of replacement.

*Examples.*

- $A = B = \mathbb{N}$, $f(a) = a_{++}$. Then $f(\mathbb{N}) = \mathbb{N} \setminus \{0\}$.

$$f^{-1}(\{a\}) = \begin{cases} \{a-1\} & \text{if } a > 0 \\ \varnothing & \text{if } a = 0 \end{cases}$$

**Definition II.1.13.** Two functions $f$ and $g$ with the same domain $X$ and range $Y$ are equal if $f(x) = g(x)$ for all $x \in X$.

**Definition II.1.14** (Composition)**.** If $f : X \to Y$ and $g : Y \to Z$, then the *composition* $g \circ f$ is a function $g \circ f \colon X \to Z$ given by

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

**Lecture 05.** Fri 12 Jan '24

**Definition II.1.15.** A function $f \colon A \to B$ is said to be

- *injective*, if $f(x) = f(y)$ implies $x = y$,
- *surjective*, if $f(A) = B$,

- *bijective*, if it is both injective and surjective.
- an *involution*, if $f(f(x)) = x$ for all $x \in A$.

**Exercise II.1.16.** *Let $f\colon A \to B$ be an involution. Show that $f$ is bijective.*

*Solution.* $f$ is surjective since everything is in the range. Injective since $f(x) = f(y) \implies f(f(x)) = f(f(y)) \implies x = y$. ■

A function is bijective iff for any $b \in B$ there is a unique $a \in A$ such that $f(a) = b$.

**Definition II.1.17.** Let $f\colon A \to B$ be bijective. The *inverse* of $f$ is the function $f^{-1}\colon B \to A$ where $f^{-1}(b)$ is the unique $a \in A$ such that $f(a) = b$.

**Axiom II.10** (Powers)**.** Let $X$ and $Y$ be sets. Then there exists a set, denoted $Y^X$, consisting of all functions from $X \to Y$.

**Exercise II.1.18.** *Let $X$ be a set. Then $\{Y \mid Y \subseteq X\}$, called the* power set *of $X$, is a set.*

*Solution.* The property $P(F, X_F)$ given by

$$P(F, X_F) \iff F \in 2^X \wedge X_F = \{x \in X \mid F(x) = 1\}$$

is satisfied by at most one $X_F$ for any $F$. Thus applying the axiom of replacement on $2^S$ gives the desired set. ■

**Axiom II.11** (Unions)**.** Let $A$ be a set whose elements are also sets. Then there exists a set, denoted $\bigcup A$, whose elements are the elements of the elements of $A$. Thus $x \in \bigcup A \iff x \in S$ for some $S \in A$.

*Remark.* This axiom implies axiom II.5.

Let $I$ be a set such that $A_\alpha$ is a set for all $\alpha \in I$. Then $\{A_\alpha \mid \alpha \in I\}$ is a set by the axiom of replacement. Thus $\bigcup_{\alpha \in I} A_\alpha$ is a set.

**Definition II.1.19.** Two sets $X$ and $Y$ are said to have the same *cardinality* if there exists a bijection $f\colon X \to Y$.

Let $n \in \mathbb{N}$. If a set $X$ has the same cardinality as $\{0, 1, \ldots, n-1\}$, then $X$ is said to be *finite* and have cardinality $n$.

**Definition II.1.20.** A set $X$ is *countably infinite* or *countable* if it has the same cardinality as $\mathbb{N}$, is *at most countable* if it is finite or countable, and is *uncountable* otherwise.

**Exercise II.1.21.** *Let $m < n$ be naturals. Show that there is*

*(i) no surjection from $[m]$ to $[n]$*[1].

*(ii) no injection from $[n]$ to $[m]$.*

*(iii) a bijection from $[a]$ to $[b]$ iff $a = b$.*

**Exercise II.1.22** (Properties of countable sets)**.**

*(i) If $X$ and $Y$ are countable, then so is $X \cup Y$.*

*(ii) The set $\{(n, m) \in \mathbb{N} \times \mathbb{N} \mid 0 \leq m \leq n\}$ is countable.*

*(iii) $\mathbb{N} \times \mathbb{N}$ is countable.*

**Theorem II.1.23.** *Let $X$ be an arbitrary set. Then $X$ and $2^X$ cannot have the same cardinality.*

*Proof.* Let $f\colon X \to 2^X$. Consider $A = \{x \in X \mid x \notin f(x)\} \subseteq X$. So $A \in 2^X$. Since for any $x \in X$, $x \in A \iff x \notin f(x)$, we have $f(x) \neq A$ for all $x \in X$. Thus $f$ is not surjective. □

**Lecture 06.**
Mon 15 Jan '24

**News:** Quiz 1 tomorrow. Material upto and including lecture 6.

**Definition II.1.24.** Let $I$ be a possibly infinite indexing set and for all $\alpha \in I$ let $X_\alpha$ be a set. Then its (possibly infinite) Cartesian product is defined as

$$\prod_{\alpha \in I} X_\alpha = \left\{ (x_\alpha)_{\alpha \in I} \in \Big( \bigcup_{\beta \in I} X_\beta \Big)^I \;\middle|\; x_\alpha \in X_\alpha \text{ for all } \alpha \in I \right\}$$

**Exercise II.1.25.** *For any sets $I$ and $X$, $\prod_{\alpha \in I} X = X^I$.*

[1] $[n] = \{1, \ldots, n\}$

**Axiom II.12** (Choice)**.** Let $I$ be a set and for all $\alpha \in I$ let $X_\alpha \neq \varnothing$. Then $\prod_{\alpha \in I} X_\alpha$ is non-empty.

**Definition II.1.26.** A *choice function* on $X$ is a function $f\colon 2^X \setminus \varnothing \to X$ such that for all non-empty $S \subseteq X$, $f(S) \in S$.

**Fact II.1.27.** *The existence of a choice function for every $X$ is equivalent to the axiom of choice.*

*Remark.* A variant of AoC is the *axiom of countable choice*, which requires $I$ to be at most countable.

**Lemma II.1.28.** *Let $E$ be a bounded above non-empty subset of $\mathbb{R}$. Then there exists a sequence $(a_n)_{n\in\mathbb{N}}$ such that $a_n \in E$ for all $n$ and $\lim_{n\to\infty} a_n = \sup E$.*

*Proof.* Let $X_n = \left\{x \in E \mid \sup E - \frac{1}{n} \leq x \leq \sup E\right\}$. Each $X_n$ is non-empty. By AoC, there exists a sequence $(a_n)_{n\in\mathbb{N}}$ such that for all $n$, $a_n \in X_n$. Thus $a_n \in E$ for all $n$ and $\lim_{n\to\infty} a_n = \sup E$. □

**Definition II.1.29.** Let $(P, \leq)$ be a poset. A subset $Y \subseteq P$ is called a *chain* or *totally ordered* if for any $y, y' \in Y$, either $y \leq y'$ or $y' \leq y$.

**Definition II.1.30.** Let $(P, \leq)$ be a poset and $Y \subseteq P$. We say that $y$ is a *minimal* (resp. *maximal*) element of $Y$ if there is no $y' \in Y$ such that $y' < y$ (resp. $y' > y$).

**Definition II.1.31.** Let $(P, \leq)$ be a poset and $Y \subseteq P$ be a chain. We say that $Y$ is *well-ordered* if every non-empty subset of $Y$ has a minimal element.

**Axiom II.12** (Well-ordering principle)**.** Given any set $X$, there exists a well-ordering on $X$.

**Axiom II.12** (Zorn's lemma)**.** Let $(X, \leq)$ be a non-empty poset such that every chain $Y$ of $X$ has an upper bound (there exists an $x \in X$ such that $y \leq x$ for all $y \in Y$). Then $X$ has a maximal element.

**Fact II.1.32.** *The axiom of choice, well-ordering principle, and Zorn's lemma are equivalent.*

*Proof.* **Zorn $\implies$ AoC.** Let $X \neq \varnothing$ and let $P$ be the set of ordered pairs $(Y, f)$ where $Y \subseteq X$ and $f$ is a choice function on $Y$. Define $(Y, f) \leq (Y', f')$ if $Y \subseteq Y'$ and $f'|_Y = f$. $P$ is non-empty because $\{x\} \subseteq X$ has a choice function for all $x \in X$.

Let $C$ be a chain in $P$. Then let $\overline{Y} = \bigcup_{(Y,f)\in C} Y$ and define $\overline{f}$ by setting $\overline{f}(S) = f(S)$ for any $f$ for which $f(S)$ is defined. Then $(\overline{Y}, \overline{f})$ is an upper bound for $C$.

By Zorn's lemma, there exists a maximal element of $P$, say $(Y, f)$. If $x \in X \setminus Y$, we can extend $f$ to $Y \cup \{x\}$ by defining $f(S) = x$ for any $S$ containing $x$. This contradicts the maximality of $(Y, f)$. Thus $X \setminus Y$ must be empty, and so $f$ is a choice function on $X$.

**AoC $\implies$ Zorn.** Let $P$ be a poset whose every chain has an upper bound. Suppose $P$ has no maximal element. Pick $x_0 \in P$ using a choice function. Since $x_0$ is not maximal, there exists an $x_1$ larger than $x_0$, and $x_2$ larger than $x_1$, and so on. This gives a chain $x_0 < x_1 < x_2 < \dots$. But then $x_\omega$ is an upper bound for this chain. This gives another chain $x_\omega < x_{\omega+1} < \dots$. But then $x_{2\omega}$ is an upper bound for this chain.

Continuing in this way, we get a chain which is "larger" than $P$ itself, a contradiction.

□

## II.2 Integers

Lecture 07.
Wed 17 Jan '24

**Definition II.2.1.** An *integer* is an expression of the form $a \setminus b$, where $a, b \in \mathbb{N}$. Two integers are said to be equal, $a \setminus b = c \setminus d$, if $a + d = b + c$. Let $\mathbb{Z}$ denote the set of all integers.

**Exercise II.2.2.** *On $\mathbb{N} \times \mathbb{N}$, $(a, b) \sim (c, d)$ defined by $a + d = b + c$ is an equivalence relation.*

**Definition II.2.3.** The *sum* and *product* of two integers $a \setminus b$ and $c \setminus d$ is defined to be

$$(a \setminus b) + (c \setminus d) \coloneqq (a + c) \setminus (b + d)$$
$$(a \setminus b) \times (c \setminus d) \coloneqq (ac + bd) \setminus (ad + bc)$$

**Proposition II.2.4.** *These operations are well-defined.*

Since $n \setminus 0$ behaves like $n$, we identify $\mathbb{N}$ with the set of all such integers.

**Definition II.2.5.** If $a \setminus b \in \mathbb{Z}$, its *negation*, denoted $-(a \setminus b)$, is defined to be $b \setminus a$. In particular, if $n \in \mathbb{N}$, then $-n = 0 \setminus n$.

**Definition II.2.6.** We define $(a \setminus b) \leq (c \setminus d)$ if there exists an $n \in \mathbb{N}$ such that $(a \setminus b) + n = (c \setminus d)$.

**Lemma II.2.7** (Trichotomy)**.**

**Definition II.2.8** (Ring)**.** A *ring* is a set $S$ with two binary operations $+$ and $\cdot$ such that for all $a, b, c \in S$,

(R1) addition is associative,

(R2) addition is commutative,

(R3) there exists an additive identity 0,

(R4) there exists an additive inverse $-a$,

(R5) multiplication is associative,

(R6) there exists a multiplicative identity 1,

(R7) multiplication is distributive over addition (on both sides).

If the multiplicative identity does not exist, we call $S$ a *rng*.

For a *commutative ring*, we require additionally that

(CR1) multiplication is commutative.

Note that inverses are unique, since if $a + b = 0$ and $a + b' = 0$, then $b = (b' + a) + b = b' + (a + b) = b'$.

**Definition II.2.9** (Ordered Ring)**.** An *ordered ring* is a ring $S$ with a total order $\leq$ such that for all $a, b, c \in S$,

(OR1) $a \leq b$ implies $a + c \leq b + c$,

(OR2) $0 \leq a$ and $0 \leq b$ implies $0 \leq ab$.

**Theorem II.2.10** (Algebra of $\mathbb{Z}$)**.** *$\mathbb{Z}$ is an ordered commutative ring.*

## II.3 Rationals

**Definition II.3.1.** A *rational number* is an expression of the form $a//b$ where $a \in \mathbb{Z}$ and $b \in \mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$. $a//0$ is not a rational number for any $a \in \mathbb{Z}$. Two rational numbers are said to be equal, $a//b = c//d$, if $ad = bc$.

$\mathbb{Q}$ denotes the set of all rational numbers.

**Definition II.3.2.** The *sum* and *product* of two rational numbers $a//b$ and $c//d$ is defined to be

$$(a//b) + (c//d) \coloneqq (ad + bc)//bd$$
$$(a//b) * (c//d) \coloneqq (ac)//(bd)$$

Since $z//1$ behaves like $z$, we identify $\mathbb{Z}$ with the set of all such rational numbers.

**Definition II.3.3.** The *reciprocal* of a non-zero rational $a//b$, denoted $(a//b)^{-1}$, is defined to be $b//a$.

*Remark.* The numerator and denominator of a rational number are not well-defined.

**Definition II.3.4.** A field is a set $F$ with 2 operations $+ : F \times F \to F$ and $\times : F \times F \to F$ such that

(F1) $+$ & $\times$ are commutative on $F$.

(F2) $+$ & $\times$ are associative on $F$.

(F3) $+$ & $\times$ satisfy distributivity on $F$, *i.e.*, $a \times (b + c) = a \times b + a \times c$ and $(b + c) \times a = a \times b + a \times c$ for all $a, b, c \in F$.

(F4) There exist 2 *distinct* elements, called 0 and 1 such that for all $x \in F$,

$$x + 0 = x$$
$$x \times 1 = 1 \times x = x$$

(F5) For every $x \in F$, there is a $y \in F$ such that $x + y = 0$.

(F6) For every $x \in F^*$, there is a $z \in F$ such that $xz = zx = 1$.

If multiplication is not commutative, we call $F$ a *division ring*.

**Definition II.3.5.** An *ordered field* is a set that admits two operations $+$ and $\cdot$ and relation $<$ so that $(F, +, \cdot)$ is a field and $(F, <)$ is an ordered set and:

(OF1) For $x, y, z \in F$, if $x < y$ then $x + z < y + z$.

(OF2) For $x, y \in F$, if $0 < x$ and $0 < y$ then $0 < x \cdot y$.

**Theorem II.3.6** (Algebra of $\mathbb{Q}$)**.** *$\mathbb{Q}$ is an ordered field.*

**Definition II.3.7.** The quotient of $x \in \mathbb{Q}$, $y \in \mathbb{Q}^*$, denoted $x/y$, is defined to be $xy^{-1}$.

This gives $a//b = a/b$ for all $a \in \mathbb{Z}$ and $b \in \mathbb{Z}^*$.

**Definition II.3.8.** $x \in \mathbb{Q}$ is a positive (resp. negative) rational number if $x = a/b$ for $a, b \in \mathbb{N}^*$ (resp. $x = -y$ for some positive $y$).

**Proposition II.3.9.** *Let $x \in \mathbb{Q}$. Then there exists a unique $n \in \mathbb{Z}$ such that $n \leq x < n + 1$. In particular, there exists an $N \in \mathbb{Z}$ such that $x < N$. This unique $n$ is denoted $\lfloor x \rfloor$.*

**Proposition II.3.10.** *If $x < y$ are rational numbers, then there is a rational $z$ such that $x < z < y$.*

**Fact II.3.11.** *There is no rational number whose square is 2.*

# Chapter III

# Combinatorics

**Lecture 08.** Fri 19 Jan '24

**Theorem III.0.1** (Pigeonhole principle)**.** *Let $n, k \in \mathbb{N}^*$ and $n > k$. Suppose we place $n$ balls in $k$ boxes. Then there exists a box with more than one ball.*

*Proof.* Suppose not. Then each box has at most one ball. If $m$ boxes have 1 ball and the others have none, then the total number of balls is $m \leq k < n$, a contradiction. □

*Example.* Construct the sequence $(a_n)_{n \geq 1}$, where

$$a_n = \sum_{k=1}^{n-1} 7 \cdot 10^k = \underbrace{77 \ldots 7}_{n \text{ times}}.$$

Then there exists an element in this sequence which is divisible by 2023.

*Proof.* We will prove something stronger, namely that such an element can be found in the first 2023 elements. Let $(r_1, \ldots, r_{2023})$ be the sequence of remainders when divided by 2023. If any $r_i = 0$, we are done. Otherwise, by the pigeonhole principle, there exist $i < j$ such that $r_i = r_j$. Then $a_j - a_i$ is divisible by 2023. But $a_j - a_i = \sum_{k=i}^{j-1} 7 \cdot 10^k = a_{j-i} \cdot 10^i$. Since 2023 is coprime to 10, 2023 divides $a_{j-i}$. □

*Example.* A round robin tournament has $n$ players and all pair of players play exactly one game. Each game is played one after the other. Then at any given time, there are two players who have played the same number of games.

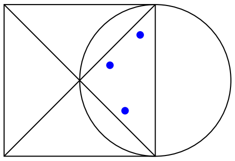
*Proof.* At any given time, let $(p_1, \ldots, p_2)$ be the sequence of number of games played by each player. Each $p_i$ is between 0 and $n-1$ (inclusive). If any $p_i = 0$, then no $p_i$ can be $n-1$, and vice versa. Thus there can be at most $n-1$ distinct values of $p_i$. By the pigeonhole principle, two of the $p_i$'s must be equal. □

**Theorem III.0.2** (Generalised PHP)**.** *Let $r, m, n \in \mathbb{N}^*$ such that $n > mr$. Suppose we place $n$ balls in $m$ boxes. Then there exists a box with at least $r+1$ balls.*

*Proof.* Suppose not. Then each box has at most $r$ balls. Then the total number of balls is at most $mr < n$, a contradiction. □

*Example.* Nine points are distributed in a unit square in some arrangement. Show that there exists a set of three points which can be covered by a disk of radius $\frac{1}{2}$.

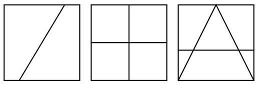
*Proof.* Partition the square by its diagonals into four triangles. By the pigeonhole principle, one of these triangles contains at least three points.



The circumradius of this triangle is $\frac{1}{2}$. □

## III.1 Mathematical Induction

*Example.* Let $f(m)$ be the maximum number of regions into which $m$ lines divide the plane.



Then $f(m) \leq \binom{m}{0} + \binom{m}{1} + \binom{m}{2}$ for all $m \geq 0$.

*Proof.* The base case $m = 1$ is true by observation.

Suppose the statement is true for some $m \geq 1$. Then there exists an arrangement $A$ of $m$ lines with $f(n)$ regions. Suppose $l$ is the $(m+1)$th line added to $A$, which intersects $k$ lines. Then it divides $k+1$ regions into two, so that the number of regions increases by $k + 1 \leq m + 1$. Thus

$$\begin{aligned} f(m+1) &\leq f(m) + m + 1 \\ &\leq \binom{m}{0} + \binom{m}{1} + \binom{m}{2} + \binom{m+1}{1} \\ &= \binom{m+1}{0} + \binom{m+1}{1} + \binom{m+1}{2}. \end{aligned}$$ □

*Example* (Wrong)*.* All cows are the same colour.

*"Proof".* We rephrase this as follows: for all $n \in \mathbb{N}^*$, all sets of $n$ cows are the same colour.

Trivially true for one cow. Suppose it is true for all sets of $n \geq 1$ cows. Consider a set of $n+1$ cows. Label them $c_1, \dots, c_{n+1}$. By the inductive hypothesis, $\{c_1, c_2, \dots, c_n\}$ are the same colour. Similarly, $\{c_2, \dots, c_n, c_{n+1}\}$ are the same colour. Thus $c_1$ and $c_{n+1}$ are the same colour. By induction, all cows are the same colour. □

**Lecture 09.**
Mon 23 Jan '24

Recall the statement of strong induction provided in proposition I.1.14.

**Theorem III.1.1** (Strong induction)**.** *Let $m_0 \in \mathbb{N}$ and let $P(m)$ be a property for all $m \in \mathbb{N}$. Suppose for all $m \geq m_0$, we have the following: if $P(m')$ holds for all $m_0 \leq m' < m$, then $P(m)$ holds. Then $P(m)$ holds for all $m \geq m_0$.*

*Example.* Consider the following reccurrence of order 2:

$$\begin{aligned} f(0) &= 1 \\ f(1) &= 2 \\ f(n+1) &= f(n-1) + 2f(n), \quad \text{for all } n \geq 1. \end{aligned}$$

Then, $f(n) \leq 3^n$ for all $n \in \mathbb{N}$.

*Proof.* The base cases $n = 0$ and $n = 1$ are true by observation. Now suppose $f(m) \leq 3^m$ for all $m < n$. Then,

$$\begin{aligned} f(n) &= f(n-1) + 2f(n-2) \\ &\leq 3^{n-1} + 2 \cdot 3^{n-2} \\ &= 7 \cdot 3^{n-2} \\ &< 3^n. \end{aligned}$$

□

## III.2 Permutations

**Definition III.2.1** (Permutation)**.** The arrangement of at most countable different objects in a linear order such that each object occurs exactly once is called a *permutation.*

*Example.* Permutations of the set $\{1, 2, 3\}$ are

$$123, 132, 213, 231, 312, 321.$$

**Proposition III.2.2.** *The number of permutations of $n$ objects, denoted $n!$ and read "n factorial", is equal to*

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1.$$

*where by convention,* $0! = 1$*.*

*Proof.* To arrange $n$ objects in a linear order, we have $n$ choices for the first object. Each choice leaves $n-1$ objects to be arranged, so we have $n! = n \cdot (n-1)!$ which gives the result by the base case $1! = 1$. □

We have several notations of a permutation.

- A bijection $\pi \colon [n] \to [n]$.
- Two line notation:
$$\begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}.$$
- One line notation:
$$\begin{pmatrix} \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$
- Cycle notation. For example, $(1347)(26)(5)$ is cycle notation for the permutation
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 6 & 4 & 7 & 5 & 2 & 1 \end{pmatrix}.$$

How fast does $n!$ grow?

**Fact III.2.3** (Stirling's formula)**.** $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$*, where* $a_n \sim b_n$ *means* $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 1$*.*

**Proposition III.2.4.** *The number of arrangements of length $k$ among $n$ objects where each of them appears at most once is* $\frac{n!}{(n-k)!}$*.*

*Proof.* We arrange the $n$ objects in a linear order. Then the first $k$ of them form an arrangement of length $k$ where each object appears at most once.

For each $k$-arrangement, we have $(n-k)!$ ways to arrange the remaining $n-k$ objects. Thus each $k$-arrangement is counted $(n-k)!$ times. Dividing by $(n-k)!$ gives the result. □

*Notation.* The *Pochhammer symbol* or *rising factorial* is defined as

$$x^{\overline{n}} = x(x+1)(x+2)\cdots(x+n-1).$$

Thus for example, $1^{\overline{n}} = n!$. Similarly, the *falling factorial* is defined as

$$x^{\underline{n}} = x(x-1)(x-2)\cdots(x-n+1).$$

**Definition III.2.5** (Multiset permutation)**.** Suppose we have $a_1$ objects of type 1, $a_2$ objects of type 2, and so on, up to $a_k$ objects of type $k$. Then a linear order of these objects is called a *multiset permutation*.

**Proposition III.2.6.** *The number of multiset permutations of $a_1$ objects of type 1, …, $a_k$ objects of type $k$, where $a_1 + a_2 + \cdots + a_k = n$, is given by the multinomial coefficient*

$$\binom{n}{a_1, a_2, \ldots, a_k} = \frac{n!}{a_1! a_2! \cdots a_k!}.$$

*For the special case of $k = 2$, we have*

$$\binom{n}{j, n-j} =: \binom{n}{j} = \frac{n!}{j!(n-j)!}.$$

*Proof.* We arrange the $n$ objects in a linear order. We have $n!$ ways to do so.

How many times is each multiset permutation counted? For each $i$, we have $a_i!$ ways to arrange the $a_i$ objects of type $i$. Dividing by $\prod a_i!$ gives the result. □

**Proposition III.2.7.** *The number of multisets of $k$ distinct objects with repetition allowed of length $n$ is $\binom{n+k-1}{n}$. This is also denoted as $\left(\!\binom{n}{k}\!\right)$, read "n multichoose k".*

**Proposition III.2.8.** *The number of $j$ element subsets of $[n]$ is $\binom{n}{j}$.*

**Corollary III.2.9.** $\sum_{j=0}^{n} \binom{n}{j} = 2^n$.

*Remarks.*

- We define $\binom{n}{j} = 0$ for $n \in \mathbb{N}$ but $j \in \mathbb{N} \setminus \{0, 1, \ldots, n\}$.
- If $n \in \mathbb{N}$, then we define $\binom{-n}{j}$ to be $(-1)^j \binom{n+j-1}{j}$.
- We leave $\binom{n}{j}$ undefined for $j \notin \mathbb{N}$.

- The binomial theorem states that
$$(1+x)^r = \sum_{j=0}^{\infty} \binom{r}{j} x^j.$$
This is a finite sum if $r \in \mathbb{N}$ by the first remark. If $r \in \mathbb{R} \setminus \mathbb{N}$, then sum is infinite, where
$$\binom{r}{k} =: \frac{r(r-1)\cdots(r-k+1)}{k!}.$$

**Proposition III.2.10.** $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ *for all* $n, k \in \mathbb{N}$.

The table of binomial coefficients is called Pascal's triangle.

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1
\end{array}$$

**Proposition III.2.11** (Identities)**.**

- $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0.$
- $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}.$
- $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$

*Proof.* Algebraic proofs:

- Consider the expansion of $(1-1)^n$.
- Consider the derivative of $(1+x)^n$, evaluated at $x = 1$.
- Consider the expansion of $(1+x)^n(1+x)^n$.

□

**Theorem III.2.12** (Zhu-Vandermonde theorem)**.** *For all* $m, n, r \in \mathbb{N}$, *we have*
$$\sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r}.$$

*Proof.* Consider a set of $m$ boys and $n$ girls. Then the number of ways to choose $r$ out of these is $\binom{m+n}{r}$, but also the LHS, where each summand is the number of ways of doing so while choosing exactly $k$ boys.

Or algebraically, compare coefficients of $x^r$ in $(1+x)^m(1+x)^n$ and $(1+x)^{m+n}$. □

**Lecture 10.**
Wed 24 Jan '24

**Theorem III.2.13** (Multinomial theorem)**.** *Let $n, k \in \mathbb{N}$ and $x_1, \ldots, x_k$ be indeterminates. Then*

$$\sum_{\substack{0 \leq a_1, \ldots, a_k \leq n \\ a_1 + \cdots + a_k = n}} \binom{n}{a_1, \ldots, a_k} x_1^{a_1} \ldots x_k^{a_k} = (x_1 + \cdots + x_k)^n$$

*Proof.* When the RHS is expanded, we get 1 term for each arrangement of $n$ objects out of $x_1, \ldots, x_k$, with repetition allowed. The coefficient of $x_1^{a_1} \ldots x_k^{a_k}$ is the number of such arrangements with $a_1$ of $x_1$, $a_2$ of $x_2$, etc. This is given by $\binom{n}{a_1, \ldots, a_k}$.

For $a_1 + \cdots + a_k \neq n$, no such term is obtained. □

**Exercise III.2.14.** *Compute $\binom{1/2}{n}$.*

*Solution.*

$$\begin{aligned}
\binom{1/2}{n} &= \frac{(1/2)(-1/2)\ldots(1/2 - n + 1)}{n} \\
&= \frac{(-1)^{n-1}(2n-3)!!}{2^n n!} \\
&= \frac{(-1)^{n-1}(2n-2)!}{2^n n!(2n-2)(2n-4)\ldots 2} \\
&= \frac{(-1)^{n-1}(2n-2)!}{2^n n! 2^{n-1}(n-1)!} \\
&= \frac{(-1)^{n-1}}{2^{2n-1}} \frac{(2n-2)!}{n!(n-1)!} \\
&= \frac{(-1)^{n-1}}{n 2^{2n-1}} \binom{2n-2}{n-1}.
\end{aligned}$$

■

## III.3 Compositions & Paritions

**Definition III.3.1.** A *weak composition* of $n \in \mathbb{N}$ is a sequence $(a_i)_{i=1}^k$ where $a_i \in \mathbb{N}$ and $a_1 + \cdots + a_k = n$. If each $a_i > 0$, then it is called a *(strict) composition.*

*Example.* For $n = 3$, its strict compositions are $(1, 1, 1)$, $(1, 2)$, $(2, 1)$ and $(3)$.

**Proposition III.3.2.** *The number of weak compositions of $n$ into $k$ parts is $\binom{n+k-1}{k-1}$.*

*Proof.* □

**Corollary III.3.3.** *The number of compositions of $n$ into $k$ parts is $\binom{n-1}{k-1}$.*

*Proof.* Each box must get at least one ball, so use proposition III.3.2 with $n \mapsto n - k$. □

**Corollary III.3.4.** *The total number of compositions is $2^{n-1}$.*

*Proof.* $\sum_{k=1}^{n} \binom{n-1}{k-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} = 2^{n-1}$. □

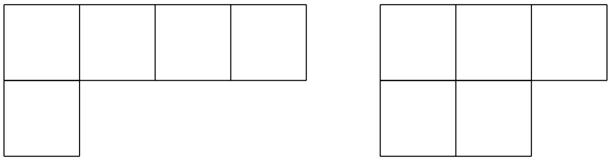
**Definition III.3.5** (Partitions)**.** An *(integer) partition* of $n \in \mathbb{N}$ is a sequence $\lambda = (\lambda_1, \ldots, \lambda_k)$ of weakly decreasing positive integers which sum to $n$. We write $\lambda \vdash n$. Each $\lambda_i$ is called a *part* and the number of parts is called the *length*, denoted $\ell(\lambda)$. We write $p(n)$ for the number of partitions of $n$.

*Example.* The partitions of 5 are $(5)$, $(4, 1)$, $(3, 2)$, $(3, 1, 1)$, $(2, 2, 1)$, $(2, 1, 1, 1)$ and $(1, 1, 1, 1, 1)$. Thus $p(5) = 7$.

**Proposition III.3.6.** *The number of partitions of $n$ into exactly (resp. at most) $k$ parts is the same as the number of partitions of $n$ with largest part exactly (resp. at most) $k$.*

**Definition III.3.7** (Young diagram)**.** The *Young/Ferrers diagram* of a partition is a left-justified array of boxes with $\lambda_i$ boxes in the $i$th row.

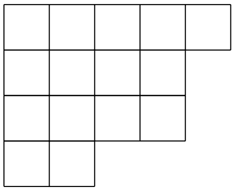
*Example.* The Young diagrams of $(4, 1)$ and $(3, 2)$ are



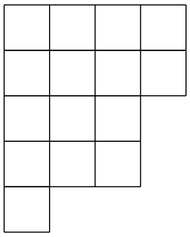
**Definition III.3.8** (Conjugate)**.** The *conjugate* of a partition $\lambda$, denoted $\lambda'$, is the partition whose Young diagram is the transpose of that of $\lambda$. That is,

$$\lambda'_i = \#\{j \in \mathbb{N} : \lambda_j \geq i\}$$

*Proof of proposition III.3.6.* If $\lambda$ has length $k$, then $\lambda'$ has largest part $k$. For example,



has conjugate



□

**Theorem III.3.9.** *The number of self-conjugate partitions of $n$ is equal to the number of partitions of $n$ into distinct odd parts.*

*Proof.* □

**Fact III.3.10** (Euler)**.** *The number of partitions of $n$ into odd parts is equal to the number of partitions of $n$ into distinct parts.*

**Fact III.3.11** (Hardy-Ramanujan Formula)**.**

$$p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{\frac{2n}{3}}}$$

**Definition III.3.12.** A *set partition* of $[n]$ is a collection of pairwise disjoint non-empty subsets/blocks whose union is $[n]$. The number of set partitions of $[n]$ into $k$ (non-empty) blocks is called the *Stirling number of the second kind* and denoted $\genfrac\{\}{0pt}{}{n}{k}$, read "$n$ set $k$".

*Example.* The set partitions of $[3]$ are 123, 12|3, 13|2, 1|23 and 1|2|3.

**Lecture 11.**
Mon 29 Jan '24

We define, by convention, $\genfrac\{\}{0pt}{}{n}{0} = \delta_{n,0}$ and $\genfrac\{\}{0pt}{}{n}{k} = 0$ for $k > n$.

We immediately have that $\genfrac\{\}{0pt}{}{n}{1} = \genfrac\{\}{0pt}{}{n}{n} = 1$ for $n \neq 0$. We enumerate some more values in table III.1.

| $k$<br>$n$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1 | | | |
| 2 | 1 | 1 | | |
| 3 | 1 | 3 | 1 | |
| 4 | 1 | 6 | 7 | 1 |

Table III.1: Stirling numbers of the second kind

**Proposition III.3.13.** *For* $1 \leq k \leq n$, $\left\{{n \atop k}\right\} = \left\{{n-1 \atop k-1}\right\} + k\left\{{n-1 \atop k}\right\}$

*Proof.* We split the partitions into two cases:

- The partition contains $\{n\}$ as a singleton. There are $\left\{{n-1 \atop k-1}\right\}$ such partitions.
- $n$ belongs to some other block. There are $\left\{{n-1 \atop k}\right\}$ ways to partition the remaining elements, and $k$ ways to choose which block $n$ belongs to. □

**Proposition III.3.14.** *The number of surjections from* $[n]$ *to* $[k]$ *is* $k!\left\{{n \atop k}\right\}$.

*Proof.* Any surjection is determined by a sequence $(p_1, p_2, \ldots, p_k)$ of preimages of $1, 2, \ldots, k$ respectively. These are simply permutations of $k$ blocks of $[n]$. □

**Corollary III.3.15.** *For all* $n \in \mathbb{N}$,

$$\sum_{j=0}^{n} \left\{{n \atop j}\right\} x^{\underline{j}} = x^n$$

*Proof.* For $x \in \mathbb{N}$, the RHS counts functions from $[n]$ to $[x]$. We split these functions by the size of the image.

For functions of image size $j$, there are $\binom{x}{j}$ ways to choose the image, and $j!\left\{{n \atop j}\right\}$ ways to choose the preimage. But $\binom{x}{j}j!$ is precisely $n^{\underline{j}}$.

Thus both sides agree at infinitely many points, and so they are equal. □

**Definition III.3.16** (Bell numbers)**.** The number of set partitions of $[n]$ is called the $n$th *Bell number*, denoted $B_n := \sum_{k=0}^{n} \left\{{n \atop k}\right\}$.

**Exercise III.3.17.** *Prove that $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$.*

*Solution.* Let $b_k$ be the number of partitions of $[n+1]$ with $n+1$ in a block of size $k+1$. Then $b_k = \binom{n}{k} B_{n-k}$. This gives the desired result by the re-indexing $k \mapsto n-k$. ■

Though this seems like a recurrence, it is not (for this course). By "recurrence" we will mean a recurrence relation dependent upon at most $M$ previous terms, for some fixed $M$.

## III.4 Permutations as Cycles

Let $S_n$ be the set of all permutations of $[n]$. Recall that any permutation $\pi \in S_n$ can be written as a product of cycles. A useful convention is to skip cycles of length 1. Thus we write $\sigma = 6754132$ as $(1635)(27)$. This allows us to consider $\pi$ as just a product (under composition) of permutations which are cyclic on some subset of $[n]$. *E.g.* $\pi = (1635) \circ (27)$, where $(27)$ for example is the permutation which swaps 2 and 7 and fixes everything else.

**Lemma III.4.1.** *Let $\sigma \in S_n$ and $j \in [n]$. Then there exists an $i \in \mathbb{N}^*$ such that $\sigma^i(j) = j$.*

*Proof.* Consider the sequence $(\sigma(j), \ldots, \sigma^n(j))$. If any of these are equal to $j$, we are done. Otherwise, by the pigeonhole principle, there exist $k < l$ such that $\sigma^k(j) = \sigma^l(j)$. Then $\sigma^{l-k}(j) = j$ (since $\sigma$ is a bijection). □

**Corollary III.4.2.** *Let $\sigma \in S_n$. Then $\sigma^{n!} = \mathrm{id}$.*

*Proof.* By the lemma, for each $j \in [n]$, there exists an $i_j \in [n]$ such that $\sigma^{i_j}(j) = j$. Since $i_j \mid n!$ for all $j$, we have $\sigma^{n!}(j) = j$ for all $j$. □

*Notation.* We will write cyclic decompositions of permutations as follows:

- Each cycle has its smallest element first.
- Cycles are written in increasing order of their smallest elements.

**Definition III.4.3.** The *cycle type* of a permutation $\sigma$, denoted type$(\sigma)$, is the partition formed by arranging its cycle lengths in weakly decreasing order.

**Lecture 12.**
Wed 31 Jan '24

*Notation.* We write a cycle type $\lambda$ in *frequency notation* as

$$\lambda = \langle 1^{a_1}, 2^{a_2}, \ldots, n^{a_n} \rangle$$

where $a_i$ is the number of times $i$ appears in $\lambda$.

**Theorem III.4.4.** *The number of permutations in $S_n$ with cycle type $\lambda = \langle 1^{a_1}, 2^{a_2}, \ldots, n^{a_n} \rangle$ is given by*

$$\frac{n!}{(1^{a_1} a_1!)(2^{a_2} a_2!) \ldots (n^{a_n} a_n!)}$$

*Proof.* For every permutation in $S_n$ in one-line notation, insert parentheses so that we first form $a_1$ cycles of length 1, $a_2$ cycles of length 2, and so on. This gives another permutation in cycle notation.

How many times does a given permutation $\sigma$ occur? For a cycle of length $j$, the same cycle occurs in $j$ ways. Since we have $a_j$ such cycles, we get a factor of $j^{a_j}$. Next, among all $j$-cycles, we get $\sigma$ if we permute these cycles, and this happends in $j!$ ways. Finally, note that for different values of $j$, these rearrangements are independent. □

*Example.* There are $\frac{n!}{n^1 1!} = (n-1)!$ permutations with cycle type $\lambda = (n) = \langle n^1 \rangle$.

*Notation.* We write $\left\langle {n \atop k} \right\rangle$ for the number of permutations in $S_n$ with exactly $k$ cycles.

**Proposition III.4.5.** *For $1 \leq k \leq n$, $\left\langle {n \atop k} \right\rangle = \left\langle {n-1 \atop k-1} \right\rangle + (n-1)\left\langle {n-1 \atop k} \right\rangle$.*

*Proof.* Consider a permutation $\sigma \in S_n$ with $k$ cycles. If $n$ is a fixed point, then $\sigma$ restricted to $[n-1]$ has $k-1$ cycles. This gives the first term.

If $n$ is not a fixed point, then removing $n$ gives a permutation in $S_{n-1}$ with $k$ cycles. There are $n-1$ ways to insert $n$ *after* any element of any cycle. □

**Proposition III.4.6.** *Let $n \in \mathbb{N}$. Then*

$$\sum_{k=0}^{n} \left\langle {n \atop k} \right\rangle x^k = x^{\bar{n}}$$

| $n \backslash k$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1 | | | |
| 2 | 1 | 1 | | |
| 3 | 2 | 3 | 1 | |
| 4 | 6 | 11 | 6 | 1 |

Table III.2: The first few values of $\left\langle {n \atop k} \right\rangle$.

*Proof.* Let $G_n(x) = x^{\overline{n}} = (x+n-1)G_{n-1}(x)$. Suppose the result holds for $n-1$. Then

$$\begin{aligned} G_n(x) &= (x+n-1)\sum_{k=0}^{n} \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle x^k \\ &= \sum_{k=0}^{n} \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle x^{k+1} + \sum_{k=0}^{n} (n-1) \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle x^k \\ &= \sum_{k=1}^{n} \left\langle \begin{matrix} n-1 \\ k-1 \end{matrix} \right\rangle x^k + (n-1)\sum_{k=0}^{n} \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle x^k \\ &= \sum_{k=1}^{n} \left( \left\langle \begin{matrix} n-1 \\ k-1 \end{matrix} \right\rangle + (n-1) \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle \right) x^k && \left(\text{because } \left\langle {n \atop 0} \right\rangle = 0\right) \\ &= \sum_{k=0}^{n} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle x^k \end{aligned}$$

The base case is $G_1(x) = x$. □

**Definition III.4.7.** The *(signed) Stirling numbers of the first kind* are $(-1)^{n-k}\left\langle {n \atop k} \right\rangle$.

**Recall:** $V = \mathbb{Q}[x]$ is the space of polynomials with rational coefficients. This is a vector space over $\mathbb{Q}$. The natural basis for $V$ is $\mathcal{B}_1 = \{1, x, x^2, \ldots\}$. But we also have another basis, $\mathcal{B}_2 = \{x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \ldots\} = \{1, x, x(x-1), \ldots\}$.

Let $S$ be the $\mathbb{N} \times \mathbb{N}$ matrix whose $(n,k)^{\text{th}}$ entry is $\left\{ {n \atop k} \right\}$. Then corollary III.3.15 shows that $S$ is the transition matrix from $\mathcal{B}_2$ to $\mathcal{B}_1$. Similarly, let $s$ be the $\mathbb{N} \times \mathbb{N}$ matrix whose $(n,k)^{\text{th}}$ entry is $(-1)^{n-k}\left\langle {n \atop k} \right\rangle$. Then $s$ is the transition matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$. Thus we have shown that:

**Proposition III.4.8.** $Ss = sS = \mathrm{id}$.

## III.5 Inclusion-Exclusion formula

Given sets $A_1, \ldots, A_n$ so that we know the size of the intersection of any $k$ of them, we can compute the size of the union of all of them.

**Theorem III.5.1** (Principle of Inclusion-Exclusion)**.**

$$\left| \bigcup_{i \in [n]} A_i \right| = \sum_{j=1}^{n} (-1)^{j-1} \sum_{\substack{S \subseteq [n] \\ |S| = j}} \left| \bigcap_{i \in S} A_i \right|$$

**Lecture 13.** Fri 02 Feb '24

*Example.* $n$ guests attend a concert and leave their coats. At the end, each leaves with a random coat. Find the probability that none of them get their own coat back.

**Definition III.5.2** (Derangement)**.** A permutation $\pi \in S_n$ is said to be a *derangement* if

$$\pi(i) \neq i \quad \text{for all } i \in [n].$$

The number of derangements of $n$ objects is denoted $!n$.

*Example.* The derangements of [3] are 312 and 231.

We find the number of derangements on $[n]$. Let $A_i = \{\pi \in S_n \mid \pi(i) = i\}$. Then,

$$A = \bigcup_i A_i = \{\pi \in S_n \mid \exists i(\pi(i) = i)\}.$$

Also note that $|A_i| = (n-1)!$, $|A_i \cap A_j| = (n-2)!$ and so on (where $i \neq j$). Thus by inclusion-exclusion,

$$\begin{aligned}
\left| \bigcup_i A_i \right| &= \sum_{j=1}^{n} (-1)^{j-1} \sum_{S \in \binom{[n]}{j}} \left| \bigcap_{i \in S} A_i \right| \\
&= \sum_{j=1}^{n} (-1)^{j-1} \binom{n}{j} (n-j)! \\
&= n! \sum_{j=1}^{n} (-1)^{j-1} \frac{1}{j!}
\end{aligned}$$

and so the number of derangements is

$$\begin{aligned} !n &= n! - \left|\bigcup_i A_i\right| \\ &= n! \sum_{j=0}^{n} (-1)^j \frac{1}{j!}. \end{aligned}$$

**Theorem III.5.3** (Derangement formula). *The number of derangements of $n$ objects is*

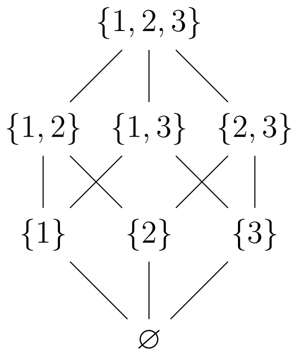
$$!n = n! \sum_{j=0}^{n} \frac{(-1)^j}{j!}.$$

This implies that $\frac{n!}{e}$ is a very good approximation for $!n$.

## A more general formula

Let $(P, \leq)$ be a poset. We say that $t$ *covers* $s$, or that $s$ is covered by $t$, if $s < t$ and there is no $u$ such that $s < u < t$. The *Hasse diagram* of $(P, \leq)$ is the graph whose vertex set is $P$ and whose edges are the cover relations. By convention, the smaller elements are drawn lower in the diagram.

*Example.* The Hasse diagram of the poset $(2^{[3]}, \subseteq)$ is



**Definition III.5.4** (Order ideal). An *order ideal* of a poset $(P, \leq)$ is a (non-empty) subset $I \subseteq P$ such that if $t \in I$ and $s \leq t$, then $s \in I$. The *principal order ideal* generated by $t \in P$ is the set

$$\downarrow t = \{s \in P \mid s \leq t\}.$$

The principal order ideal generated by $t$ is an order ideal because of transitivity, and is the minimal order ideal containing $t$.

**Definition III.5.5** (Möbius function)**.** Let $(P, \leq)$ be a poset for which every principal order ideal is finite. Then the *Möbius function* $\mu\colon P \times P \to \mathbb{Z}$ is given by

$$\begin{aligned} \mu(s,s) &= 1, \\ \mu(s,t) &= -\sum_{s \leq u < t} \mu(s,u). \end{aligned}$$

*Example.* In the previous example of the poset $(2^{[3]}, \subseteq)$, we have

$$\begin{aligned} \mu(\varnothing, \{1\}) &= -\sum_{\varnothing \leq t < \{1\}} \mu(\varnothing, t) = -\mu(\varnothing, \varnothing) = -1, \\ \mu(\varnothing, \{1,2\}) &= -\mu(\varnothing, \varnothing) - \mu(\varnothing, \{1\}) - \mu(\varnothing, \{2\}) \\ &= -1 + 1 + 1 = 1 \end{aligned}$$

**Theorem III.5.6** (Möbius inversion formula)**.** *Let $(P, \leq)$ be a poset for which every principal order ideal is finite. Let $f, g\colon P \to \mathbb{R}$ be functions satisfying*

$$g(t) = \sum_{s \leq t} f(s) \quad \text{for all } t \in P.$$

*Then*

$$f(t) = \sum_{s \leq t} \mu(s,t) g(s).$$

*Proof.* We have

$$\begin{aligned} \sum_{s \leq t} \mu(s,t) g(s) &= \sum_{s \leq t} \mu(s,t) \sum_{u \leq s} f(u) \\ &= \sum_{u \leq t} \sum_{u \leq s \leq t} \mu(s,t) f(u) \\ &= \sum_{u \leq t} f(u) \sum_{u \leq s \leq t} \mu(s,t) \\ &= \sum_{u \leq t} f(u) \delta_{u,t} \end{aligned}$$

where $\delta_{u,t}$ is the Kronecker delta. Why is the last step true? If $u = t$, then

$$\sum_{u \leq s \leq t} \mu(s,t) = \mu(t,t) = 1.$$

Otherwise,

$$\begin{aligned}\sum_{u \leq s \leq t} \mu(s,t) &= \sum_{u \leq s < t} \mu(s,t) + \mu(t,t) \\ &= 1 - \mu(u,t).\end{aligned}$$

□

## III.6 Generating Functions

**Lecture 14.** Fri 02 Feb '24

**Definition III.6.1** (Formal power series)**.** The set of *formal power series* $\mathbb{R}[[x]]$ consists of formal sums of the form $\sum_{n=0}^{\infty} a_n x^n$, where $a_n \in \mathbb{R}$ for all $n \in \mathbb{N}$.

$\mathbb{R}[[x]]$ is the completion of the ring of polynomials $\mathbb{R}[x]$. One can define $\mathbb{R}[[x]]$ as the set of all sequences $(a_n)_{n=0}^{\infty}$ of real numbers, with addition define componentwise and multiplication defined by convolution. That is,

$$\begin{aligned}(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} &= (a_n + b_n)_{n\in\mathbb{N}} \\ (a_n)_{n\in\mathbb{N}} \times (b_n)_{n\in\mathbb{N}} &= \left( \sum_{k=0}^{n} a_k b_{n-k} \right)_{n\in\mathbb{N}}\end{aligned}$$

**Definition III.6.2** (Generating function)**.** Given a sequence $(a_n)_{n\in\mathbb{N}}$, the FPS

$$A(x) = \sum_{n=0}^{\infty} a_n x^n$$

is called the *ordinary generating function* of $(a_n)_n$.

*Example.* Let $a_0 = 50$ and $a_n = 4a_{n-1} - 100$ for $n \geq 1$. Then

$$\begin{aligned} A(x) &= \sum_{n=0}^{\infty} a_n x^n \\ &= 50 + \sum_{n=1}^{\infty} (4a_{n-1} - 100)x^n \\ &= 50 + 4x \sum_{n=0}^{\infty} a_n x^n - 100x \sum_{n=0}^{\infty} x^n \\ &= 50 + 4xA(x) - \frac{100x}{1-x} \end{aligned}$$

so

$$A(x) = \frac{50}{1-4x} - \frac{100x}{(1-x)(1-4x)}.$$

We split the fraction into partial fractions as

$$\begin{aligned} \frac{x}{(1-x)(1-4x)} &= \frac{1}{3} \frac{(1-x)-(1-4x)}{(1-x)(1-4x)} \\ &= \frac{1}{3(1-4x)} - \frac{1}{3(1-x)}. \end{aligned}$$

This gives

$$\begin{aligned} A(x) &= \frac{50}{1-4x} - \frac{100}{3(1-4x)} + \frac{100}{3(1-x)} \\ &= \frac{50}{3(1-4x)} + \frac{100}{3(1-x)} \\ &= \frac{50}{3} \sum_{n=0}^{\infty} 4^n x^n + \frac{100}{3} \sum_{n=0}^{\infty} x^n. \end{aligned}$$

Thus $a_n = \frac{50}{3} 4^n + \frac{100}{3}$.

**Theorem III.6.3.** *Let $p_n$ be the number of partitions of $n$. Then*

$$\sum_{n=0}^{\infty} p_n q^n = \prod_{n=1}^{\infty} \frac{1}{1-q^n}.$$

*Proof.* Expand the RHS as

$$(1 + q + q^2 + \dots)(1 + q^2 + q^4 + \dots)\dots$$

Let any term in the product be $q^{a_1}q^{2a_2}\dots$. This can be identified with the partition $\lambda = \langle 1^{a_1}, 2^{a_2}, \dots\rangle$. Thus the coefficient of $q^n$ in the product is the number of partitions of $n$. □

We are now equipped to prove fact III.3.10.

*Proof.* Following the idea of the previous proof, we can see that the number of partitions of $n$ into odd parts has ogf

$$\sum_{n=0}^{\infty} p_{\text{odd}}(n)q^n = \prod_{n=1}^{\infty} \frac{1}{1-q^{2n-1}}$$

and the number of partitions into distinct parts has ogf

$$\sum_{n=0}^{\infty} p_{\text{distinct}}(n)q^n = \prod_{n=1}^{\infty} 1+q^n.$$

With some algebraic manipulation, we produce

$$\begin{aligned}\prod_{n=1}^{\infty} 1+q^n &= \prod_{n=1}^{\infty} \frac{1-q^{2n}}{1-q^n} \\ &= \prod_{n=1}^{\infty} \frac{1}{1-q^{2n-1}}\end{aligned}$$

since all the even terms cancel out. □

**Exercise III.6.4** (Product formula)**.** *Let $A(x)$ and $B(x)$ be the ogfs of sequences $(a_n)_n$ and $(b_n)_n$ respectively. Then $A(x) \cdot B(x)$ is the ogf of the sequence $(c_n)_n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.*

*Proof.* The coefficient of $x^n$ in $A(x) \cdot B(x)$ is obviously

$$\sum_{k=0}^{n} a_k b_{n-k}.$$

□

*Example.* Let $C_n$ be the number of ways of forming a valid word with $n$ pairs of parentheses.

Let $C(x)$ be the generating function of $(C_n)_n$. The key idea of the solution is to consider the matching closing parenthesis of the first opening parenthesis. Thus any valid word $w$ can be uniquely written as $w = (w_1)w_2$, where $w_1$ and $w_2$ are both valid. This gives

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

**Lecture 15.**
Mon 05 Feb '24

**Recall:** Let $C_n$ be the number of valid words with $n$ pairs of parentheses. We had derived the "recurrence"

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$$

Let $C(x)$ be the ogf of $(C_n)_{n\in\mathbb{N}}$. Then the RHS has ogf $C(x)^2$ and the LHS has the ogf $\frac{C(x)-1}{x}$ (exercise).

*Solution.* The RHS is by the product formula. For the LHS, let $C_{+1}(x)$ be the ogf of $(C_{n+1})_n$. That is,

$$\begin{aligned} C_{+1}(x) &= \sum_{n=0}^{\infty} C_{n+1}x^n \\ &= \frac{1}{x}\sum_{n=0}^{\infty} C_{n+1}x^{n+1} \\ &= \frac{1}{x}\sum_{n=1}^{\infty} C_n x^n \\ &= \frac{1}{x}(C(x)-1) \end{aligned}$$

since $C_0 = 1$ (the empty word). ■

So

$$\begin{aligned} C(x) - 1 &= xC(x)^2 \\ C(x) &= \frac{1 \pm \sqrt{1-4x}}{2x} \end{aligned}$$

The positive root is not a formal power series, so $C(x) = \frac{1-\sqrt{1-4x}}{2x}$.

**Exercise III.6.5.** *Use the binomial theorem to get*

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

*These are called the* Catalan numbers, *since they were first studied by Euler.*

*Solution.*

$$\begin{aligned} C(x) &= \frac{1 - (1 + \sum_{n=1}^{\infty} \binom{1/2}{n}(-4x)^n}{2x} \\ &= \frac{1}{2}\sum_{n=0}^{\infty}\binom{1/2}{n+1}(-1)^n 4^{n+1} x^n \end{aligned}$$

so by exercise III.2.14,

$$\begin{aligned} &= \frac{1}{2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(n+1)2^{2n+1}}\binom{2n}{n}(-1)^n 4^{n+1} x^n \\ &= \sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}x^n. \end{aligned}$$

which gives the desired result. ■

**Definition III.6.6** (Exponential generating function)**.** Given a sequence $(a_n)_{n\in\mathbb{N}}$, the formal power series $A(x) = \sum_{n\in\mathbb{N}} a_n \frac{x^n}{n!}$ is called the *exponential generating function* (egf) of $(a_n)_{n\in\mathbb{N}}$.

*Examples.*

- The egf of $(1)_{n\in\mathbb{N}}$ is $e^x$.
- Let $a_0 = 1$, $a_{n+1} = (n+1)(a_n - n + 1)$ for $n \geq 0$. Then its egf is

$$\begin{aligned} A(x) &= \frac{1}{1-x} + xe^x \\ \implies a_n &= n! + n \end{aligned}$$

**Exercise III.6.7** (Product formula)**.** *Let $A$ and $B$ be egfs for $(a_n)_{\mathbb{N}}$ and $(b_n)_{\mathbb{N}}$, respectively. Then the sequence $(c_n)_{\mathbb{N}}$ which has the egf $AB$ is given by*

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}.$$

**Theorem III.6.8** (Exponential formula)**.** *Let $a_n$ be the number of ways to build some structure on $[n]$ with $a_0 = 0$ and $h_n$ be the number of ways to form a set partition of $[n]$ and then build the same structure on its blocks, with $h_0 = 1$. If $A$ and $H$ are the egfs of $(a_n)_n$ and $(h_n)_n$ respectively, then*

$$H(x) = e^{A(x)}.$$

*Proof.* The number of ways of partitioning $[n]$ into $k$ blocks and building the same structure has egf $\frac{A(x)^k}{k!}$ by the product formula (since order doesn't matter). Thus

$$H(x) = 1 + \sum_{k=1}^{\infty} \frac{A(x)^k}{k!} = e^{A(x)}. \qquad \square$$

*Examples.*

- Let $a_n$ be the number of sets of $[n]$. That is, $a_n = 1$ for all $n$ except $a_0 = 0$. Then $A(x) = e^x - 1$. This gives $H(x) = e^{e^x - 1}$, which is the egf of $B_n$.

- Let $i_n$ be the number of involutions in $S_n$. That is,

  $$i_n = \#\{\sigma \in S_n \mid \sigma = \sigma^{-1}\}.$$

  We had shown earlier that an involution has cycle type $\langle 1^{a_1}, 2^{a_2} \rangle$.

  The structure here is to have only one- and two-cycles. So

  $$a_n = \begin{cases} 1 & \text{if } n = 1, \\ 1 & \text{if } n = 2, \\ 0 & \text{otherwise.} \end{cases}$$

  This gives $A(x) = x + \frac{x^2}{2}$. So by the exponential formula, the number of involutions has egf

  $$H(x) = \sum_{n \in \mathbb{N}} i_n \frac{x^n}{n!} = e^{x + \frac{x^2}{2}}.$$

# Chapter IV

# Graph Theory

⟨Insert oft-repeated story about Königsberg bridges.⟩

## IV.1 Graphs & Trees

**Lecture 16.**
Wed 07 Feb '24

**Definition IV.1.1** (Simple graph)**.** A *simple graph* is an ordered pair $G = (V, E)$, where $V$ is a set of *vertices* and $E$ is a collection of 2-element subsets of $V$, called *edges*.

A vertex $v \in V$ is said to be *incident* to an edge $e \in E$ if $v \in e$. Vertices $v_1, v_2 \in V$ are said to be *adjacent* if $\{v_1, v_2\} \in E$.

In general, we could also have *loops*, which are edges from one vertex to itself, and *multiple edges*, which are two or more edges between the same pair of vertices. One could also have *weighted* edges. We will not define these formally, since it is too many definitions.

We will also only focus on finite graphs, *i.e.*, graphs with a finite number of vertices (which implies a finite number of edges).

**Definition IV.1.2** (Walks, paths and trails)**.**

- A *walk* is a sequence of vertices $(v_1, v_2, \ldots, v_n)$, where $v_i$ is adjacent to $v_{i+1}$ for $1 \leq i < n$.
- A walk is a *trail* if all the edges are distinct.
  - A trail is said to be *Eulerian* if all the edges of the graph are used in it.
- A walk is a *path* if all the vertices are distinct.

- A walk, trail, or path is *closed* if $v_1 = v_n$.
- A closed path[1] is a *cycle*.

**Definition IV.1.3** (Connectedness)**.** A graph is *connected* if there is a path between any two vertices.

**Definition IV.1.4** (Degree)**.** The *degree* of a vertex $v$ in a graph $G = (V, E)$ (possibly with loops or parallel edges) is the number of edges incident to $v$. An edge from $v$ to itself counts twice.

**Theorem IV.1.5** (Königsberg)**.** *A connected graph $G$ (possibly with loops or parallel edges) has a closed Eulerian trail iff all vertices of $G$ have even degree.*

*Proof.* Suppose $G$ has a closed Eulerian trail. Then, every time a vertex is visited, it is also left by a different edge. Thus each vertex has even degree.

Conversely, suppose all vertices have even degree. Choose any vertex $v$ and an edge $e_1$ incident to $v$ (this exists by connectedness, except in the case of the singleton graph) and go to $v_1$. Pick another edge $e_2$ incident to $v_1$ and go to $v_2$. Continue in this way until we return to $v$, forming a closed trail $C_1$. This must happen, since $G$ is finite, and each vertex has even degree. If $C_1$ contains every edge, we are done.

Otherwise, there must exist a vertex $w$ in $C_1$, which is incident to an edge not yet visited, since the graph is connected. Continue from $w$ as before to form a disjoint closed trail $C_2$. We can merge this with $C_1$ to form a longer closed trail.

Continue in this way until we have used all edges. □

**Corollary IV.1.6.** *The Königsberg bridge problem has no solution.*

**Corollary IV.1.7.** *A connected graph $G$ (possibly with loops or parallel edges) has an Eulerian trail iff it has at most two vertices of odd degree.*

*Proof.* First note that it is not possible for exactly one vertex to have odd degree.

Suppose $G$ has an Eulerian trail. If it is closed, then all vertices have even degree. If not, add an edge between the first and last vertices to form a cycle, and apply the theorem.

Conversely, suppose $G$ has at most two vertices of odd degree. If there are none, then $G$ has a closed Eulerian trail. If there are two, then add an edge between them to form a cycle, and apply the theorem. □

[1] How can a path be closed? We mean that the first and last vertices are the same, but no other vertex repeats.

**Definition IV.1.8** (Hamiltonian path)**.** A *Hamiltonian path* is a cycle that visits every vertex.

**Fact IV.1.9** (Dirac's theorem)**.** *If every vertex of $G$ has degree at least $n/2$, where $|v| = n$, then $G$ has a Hamiltonian path.*

**Definition IV.1.10** (Graph isomorphism)**.** Two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are said to be *isomorphic* if there exists a bijection $\Phi \colon V_1 \to V_2$ such that $\{v_1, v_2\} \in E_1$ iff $\{\Phi(v_1), \Phi(v_2)\} \in E_2$.

**Definition IV.1.11** (Tree)**.** A connected graph with no cycles is called a *tree.*

**Lecture 17.** Fri 09 Feb '24

**Proposition IV.1.12.** *Let $T$ be a tree.*

*(1) Deleting any edge in $T$ disconnects it.*

*(2) Adding a new edge to $T$ creates a cycle.*

*(3) For any $v, w \in V(T)$, there is a* unique *path from $u$ to $v$.*

*Proof.* (3) holds because if there is more than one path from $u$ to $v$, then we must create a cycle.

For (2), suppose $\{v, w\} \notin E(T)$ and we add it. By (3), there is a unique path from $v$ to $w$ in $E(T)$, and so adding this edge creates a cycle.

For (1), suppose that removing an edge $\{v, w\}$ from $T$ still left it connected. Then we would have two paths from $v$ to $w$, contradicting (3). □

**Definition IV.1.13.** A vertex with degree 1 is called a *leaf* or *pendant vertex.*

**Lemma IV.1.14.** *Every tree on $n \geq 2$ vertices has at least 2 leaves.*

*Proof.* Let the longest path in the tree be $(v_1, \dots, v_k)$. Then $v_1$ and $v_k$ must be leaves, for otherwise the path could be made longer. □

**Theorem IV.1.15.** *All trees on $n$ vertices have $n - 1$ edges.*

*Proof.* This is clearly true for the singleton tree. Let $T$ be a tree with $n + 1$ vertices, and let $l$ be a leaf (by the previous lemma). Removing $l$ and its incident edge gives a tree of $n$ vertices with $n - 1$ edges. Thus, $T$ has $n$ edges. Winduction. □

**Theorem IV.1.16.** *Any connected graph on $n$ vertices with $n-1$ edges is a tree.*

*Proof.* True for $n = 1$. □

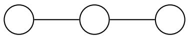
**Definition IV.1.17.** A *forest* is a graph with no cycles.

A tree is a connected forest.

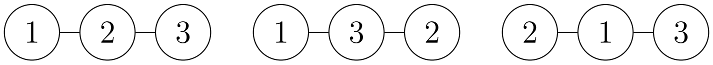
We wish to count the number of trees on vertices labelled $[n]$.

*Examples.*

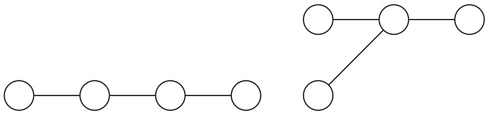
- For $n = 3$, first note that there is exactly one unlabelled tree,



This gives rise to 3 labelled trees

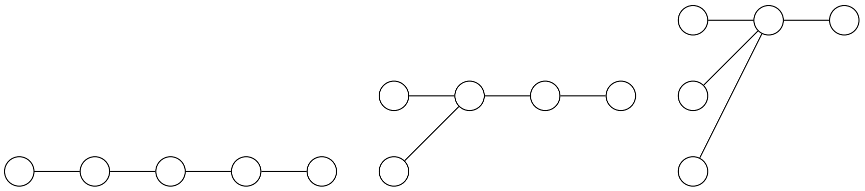


- For $n = 4$, there are 2 unlabelled trees.



The first tree gives rise to $4!/2 = 12$ labelled trees, and the second gives rise to 4 labelled trees. Total: 16 labelled trees.

- For $n = 5$, there are 3 unlabelled trees.



The first tree gives rise to $5!/2 = 60$ labelled trees, the second to $5!/2! = 60$, and the last to 5. Total: 125 labelled trees.

We observe the pattern that the number of labelled trees on $n$ vertices is $n^{n-2}$.

**Theorem IV.1.18** (Cayley's formula)**.** *The number of trees labelled* $[n]$ *is* $n^{n-2}$.

*Remark.* The book presents a bijective proof.

**Definition IV.1.19.** A *rooted tree* $T(v)$ is a tree with a marked vertex $v$, called the *root*.

A *branching* of a rooted tree $T(v)$ is an orientation of $T$, *i.e.*, an assignment of directions to the edges of $T$, in which every edge is directed away from $v$.

A *rooted forest* is one where every component has a root, and we can constuct branchings in the same way.

We will show that the number of branchings which is equal to the number of rooted trees is $n^{n-1}$.

*Proof of Cayley's formula.* We start with the empty graph over $n$ vertices and add edges one at a time to form a branching. Initially, there are $n$ components. At the $k^{\text{th}}$ stage, we will have $n-k$ components. Consider the following algorithm:

> For $1 \leq k \leq n-1$, at the $k^{\text{th}}$ stage, add an oriented edge $(u, v)$ from any vertex to the root of one of the components to which it does not belong.

At the first stage, we have $n$ choices for $u$ and $n-1$ choices for $v$. At the second stage, we have $n$ choices for $u$ and $n-2$ choices for $v$, and so on. Thus at the $k^{\text{th}}$ stage, we have $n(n-k)$ ways of forming a rooted forest. Continuing this way, we get that the number of branchings is $n^{n-1}(n-1)!$.

But note that every branching occurs $(n-1)!$ times, because of different orderings of the edges. Thus the total number of rooted trees is $n^{n-1}$. □

Cayley's formula follows as a corollary. (A factor of $n$ comes from the choice of root.)

**Exercise IV.1.20.** *The number of rooted forests on* $n$ *vertices is* $(n+1)^{n-1}$.

*Proof.* Introduce a special vertex $v_{-1}$, and consider all rooted trees on $n+1$ vertices with root $v_{-1}$. Removing $v_{-1}$ gives a rooted forest on $n$ vertices, and every rooted forest on $n$ vertices arises in this way. Thus by Cayley's formula, the number of rooted forests on $n$ vertices is $(n+1)^{n-1}$. □

**Definition IV.1.21.** Let $G = (V, E)$ and $|V| = n$. The *adjacency matrix* $A$ is the $n \times n$ matrix indexed by $V$ whose entries are

$$A_{v,w} = \mathbf{1}_{\{v,w\} \in E}.$$

**Proposition IV.1.22.** *Let $G$ be a graph and $A$ be its adjacency matrix. Then $(A^k)_{v,w}$ counts the number of walks from $v$ to $w$ of length $k$.*

**Lecture 18.**
Mon 12 Feb '24

We aim to generalise Cayley's formula.

**Definition IV.1.23** (Subgraph)**.** Let $G = (V, E)$. A *subgraph* of $G$ is a graph $G' = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E \cap 2^{V'}$.

**Definition IV.1.24** (Spanning tree)**.** A *spanning tree* $T$ of a graph $G = (V, E)$ is a subgraph with vertex set $V$ such that $T$ is a tree.

*Example.* A spanning tree of the complete graph $K_5$ with vertex set $[5]$ has the edges $\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}$.

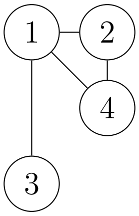
**Definition IV.1.25** (Complete graph)**.** The *complete graph* $K_n$ is the graph on $n$ vertices with an edge between every pair of vertices.

In the language of spanning trees, Cayley's formula states that the number of spanning trees of $K_n$ is $n^{n-2}$.

**Definition IV.1.26** (Laplacian)**.** The *Laplacian* of a (simple) graph $G = (V, E)$ is the matrix given by $L = D - A$, where $A$ is the adjacency matrix of $G$ and $D = \operatorname{diag}(\deg(v_1), \ldots, \deg(v_n))$.

The *reduced Laplacian* $L_0$ is obtained by deleting the last row and column of $L$.

*Example.* Let $G$ be given by



**Theorem IV.1.27** (Kirchoff's matrix tree theorem)**.** *Let $G$ be a graph and $L_0$ its reduced Laplacian. Then the number of spanning trees of $G$ is given by* $\det(L_0)$.

**Definition IV.1.28.** Let $G = (V, E)$ be a graph, $V = [n]$, and $m = |E|$, with the edges labelled by $[m]$. Suppose the edges of $G$ are oriented in some way. Then the *incidence matrix* $\mathcal{I}(G) = \mathcal{I}$ is the $n \times m$ matrix given by

$$\mathcal{I}_{v,e} = \begin{cases} 1 & \text{if } v \text{ is the head of } e, \\ -1 & \text{if } v \text{ is the tail of } e, \\ 0 & \text{otherwise.} \end{cases}$$

**Exercise IV.1.29.** *Let $G = (V, E)$ be a graph with Laplacian $L$. Let $\mathcal{I}_0$ be the incidence matrix with the last row removed, for any orientation of the edges of $G$. Then, independent of the choice of orientation,*

- $L = \mathcal{I}\mathcal{I}^\top$,
- $L_0 = \mathcal{I}_0\mathcal{I}_0^\top$.

**Theorem IV.1.30** (Cauchy-Binet formula)**.** *Let $A$ be an $n \times m$ matrix and $B$ an $m \times n$ matrix with $n < m$. For an $n$-sized subset $S$ of $[m]$, let $A_{[n],S}$ (resp. $B_{S,[n]}$) be the $n \times n$ submatrix of $A$ (resp. $B$) formed by choosing the columns of $A$ (resp. rows of $B$) with indices in $S$. Then*

$$\det AB = \sum_{S \in \binom{[m]}{n}} \det A_{[n],S} \det B_{S,[n]}.$$

*Proof of theorem IV.1.27.* We will use the face that $L_0 = \mathcal{I}_0\mathcal{I}_0^\top$ and Cauchy-Binet. Fix a subset $S$ of $[m]$, *i.e.*, edges in $G$, of size $n-1$. Let $X = (\mathcal{I}_0)_{[n-1],S}$ Then the summand on the right-hand side of Cauchy-Binet for the determinant of $L_0 = \mathcal{I}_0\mathcal{I}_0^\top$ is $\det X \det X^\top = (\det X)^2$.

We claim that $(\det X)^2 = [(V, S) \text{ is a tree}]$.

Suppose there exists a vertex $i$ of degree 1 in $G' = (V, S)$. Then the $i^{\text{th}}$ row in $X$ has only one non-zero entry, either 1 or $-1$. Expand $\det X$ using that row and use the induction hypothesis. The remaining graph is a tree iff $G'$ is a tree.

If there are no vertices of degree 1 in $G'$, then $G'$ cannot be a tree. Since $G'$ has $n-1$ edges, it is disconnected and must contain a cycle. The columns of $X$ corresponding to the cycle must be linearly dependent, so $\det X = 0$.

Thus the claim is proved, and therefore $\det L_0 = \det \mathcal{I}_0 \mathcal{I}_0^\top$ gets a contribution of 1 from each spanning tree of $G$. □

**Corollary IV.1.31** (Cayley's formula)**.** *The number of spanning trees of $K_n$ is $n^{n-2}$.*

*Proof.* Let $G = K_n$. Then

$$\begin{aligned}\det L_0 &= \det \begin{pmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{pmatrix}_{(n-1)\times(n-1)} \\ &= \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{pmatrix} \\ &= n^{n-2}.\end{aligned}$$

□

**Lecture 19.**
Wed 14 Feb '24

We will now prove the Cauchy-Binet formula.

**Lemma IV.1.32** (Sylvester's determinant identity)**.** *Let $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times n}$. Then*

$$\lambda^m \det(\lambda I_n + AB) = \lambda^n \det(\lambda I_m + BA).$$

*Proof.* Use $2 \times 2$ block matrices. Note that

$$\begin{aligned}\begin{pmatrix}\lambda I_n & A \\ B & \lambda I_m\end{pmatrix} &= \begin{pmatrix}I_n & 0 \\ B & I_m\end{pmatrix}\begin{pmatrix}\lambda I_n & 0 \\ 0 & \lambda I_m - BA\end{pmatrix}\begin{pmatrix}I_n & A \\ 0 & I_m\end{pmatrix} \\ &= \begin{pmatrix}I_n & A \\ 0 & I_m\end{pmatrix}\begin{pmatrix}\lambda I_n - AB & 0 \\ B & \lambda I_m\end{pmatrix}\begin{pmatrix}I_n & 0 \\ B & I_m\end{pmatrix}.\end{aligned}$$

**Fact IV.1.33.** *For such block matrices, the determinant of an upper or lower triangular block matrix is the product of the determinants of the diagonal blocks.*

Then we have

$$\det\begin{pmatrix}\lambda I_n & A \\ B & \lambda I_m\end{pmatrix} = \lambda^n$$

□

*Proof of Cauchy-Binet.* Compare the coefficient of $\lambda^{m-n}$ in the two sides of

$$\lambda^{m-n}\det(\lambda I_n + AB) = \det(\lambda I_m + BA).$$

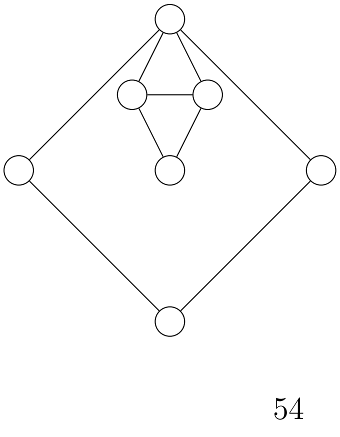
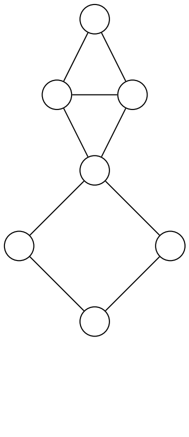
□

## IV.2 Planar Graphs

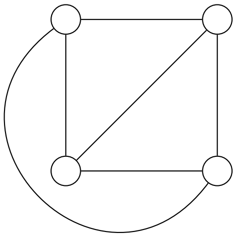
**Definition IV.2.1** (Planar graph)**.** A graph which can be drawn in the plane without edges intersecting in non-vertices is called a *planar graph.* We can allow loops and parallel edges.

A planar graph together with its planar embedding is called a *plane graph.*

*Examples.*

-  and  are isomorphic as planar graphs, but not as plane graphs.

- Edges need not be straight lines. Thus $K_4$ is planar.



**Fact IV.2.2** (Jordan curve theorem)**.** *A plane graph partitions the plane into disjoint regions, which we call* faces. *We also include the unbounded face.*

**Theorem IV.2.3** (Euler's theorem)**.** *Let $G$ be a connected planar graph with $V$ vertices, $E$ edges, and $F$ faces. Then*

$$V - E + F = 2.$$

*Proof.* We induct on $E$. If $E = 0$, then $V = 1$ and $F = 1$, so the result holds. Suppose the result holds for all connected planar graphs with $E - 1$ edges.

We find an edge $e$ such that removing $e$ from $G$ gives a connected graph $G'$. Removing $e$ will merge the two faces on either side of $e$. Then $G'$ has $V$ vertices, $E - 1$ edges and $F - 1$ faces. Then $V - E + F = V - (E - 1) + (F - 1) = 2$.

If such an edge does not exist, *i.e.*, removing any edge disconnects the graph, then $G$ is a tree. So $V - E + F = V - (V - 1) + 1 = 2$. □

*Remark.* Planar graphs can also be embedded on a sphere.

**Definition IV.2.4** (Bipartite graph)**.** A *bipartite graph* $G = (V, E)$ is one where $V = V_1 \sqcup V_2$ such that no edge connects two vertices in the same set. The *complete bipartite graph* $K_{m,n}$ is the bipartite graph where $|V_1| = m$, $|V_2| = n$, and $\{v_1, v_2\} \in E$ for all $v_1 \in V_1$ and $v_2 \in V_2$.

**Corollary IV.2.5.** *$K_{3,3}$ is not planar.*

*Proof.* Suppose it were. $V = 6$, $E = 9$, so by Euler's theorem, $F = 5$. But each face must have at least 4 edges since $K_{3,3}$ is bipartite. Summing over all faces, we get $2E \geq 4F$, a contradiction. □

**Definition IV.2.6** (Minor)**.** A *minor* of a graph $G$ is one obtained by deleting vertices or edges, or *contracting* edges. An edge is contracted by removing it and merging its two endpoints.

**Fact IV.2.7** (Kuratowski's theorem)**.** *A graph is planar iff it has no minor isomorphic to* $K_5$ *or* $K_{3,3}$.

**Definition IV.2.8** (Colouring)**.** A *(vertex) colouring* of a graph $G$ is an assignment of colours to the vertices of $G$ so that adjacent vertices have different colours.

**Fact IV.2.9** (Four colour theorm)**.** *Any planar graph can be coloured using at most 4 colours.*

# Chapter V

# Number Theory

**Lecture 20.**
Mon 26 Feb '24

## V.1 Algebraic Structures

**Definition V.1.1** (Group)**.** A *group* is a set $G$ together with a binary operation $*\colon G \times G \to G$ which satisfies the following conditions:

(G1) **associativity:** $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$.

(G2) **identity:** there exists an element $e \in G$, called the *identity*, such that $a * e = e * a = a$ for all $a \in G$.

(G3) **inverses:** for all $a \in G$, there exists an element $a^{-1} \in G$ called the *inverse* of $a$, such that $a * a^{-1} = a^{-1} * a = e$.

*Notation.* We write $ab$ for $a * b$.

*Examples.*

- $(\mathbb{Z}, +)$, which has $e = 0$ and $a^{-1} = -a$. Similarly, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ or $(\mathbb{C}, +)$ are groups.

- $(\mathbb{Q}^*, *)$, which has $e = 1$ and $a^{-1} = 1/a$. Similarly, $(\mathbb{R}^*, *)$ and $(\mathbb{C}^*, *)$ are groups. But $(\mathbb{Z}^*, *)$ is not a group as it lacks inverses.

- Let $(V, +)$ be a vector space over any field. Then it is also a group.

- Let $S_n$ be the set of all permutations on $[n]$. Then $(S_n, \circ)$ is a group. In fact, it is the only non-abelian (for $n \geq 3$) group in this list.

**Definition V.1.2** (Abelian group)**.** A group is said to be *abelian* if the group operation is commutative.

**Definition V.1.3** (Subgroup)**.** Let $(G, *)$ be a group. A *subgroup* of $G$ is a non-empty subset $H \subseteq G$ which is closed under products and inverses.

*Example.* $(2\,\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Z}, +)$.

**Definition V.1.4** (Ring)**.** A *ring* is a set $R$ together with two binary operations, $+\colon R \to R$ and $*\colon R \times R \to R$ called addition and multiplication, respectively, which satisfy the following conditions:

(R1) $(R, +)$ is an abelian group.

(R2) **associativity:** $(a * b) * c = a * (b * c)$ for all $a, b, c \in R$.

(R3) **distributivity:** $(a + b) * c = a * c + b * c$ and $a * (b + c) = a * b + a * c$ for all $a, b, c \in R$.

$R$ is said to be *commutative* if it saisfies

(R4) **commutativity:** $a * b = b * a$ for all $a, b \in R$,

and *with identity* if

(R5) **identity:** there exists an element $1 \in R$ such that $1 * a = a * 1 = a$ for all $a \in R$.

A subring of $R$ is a subgroup of $R$ (under addition) which is closed under multiplication.

*Examples.*

- $(\mathbb{Z}, +, *)$ is a commutative ring with identity.
- $(2\,\mathbb{Z}, +, *)$ is a commutative ring without identity. It sits inside $(\mathbb{Z}, +, *)$ as a subring.
- Let $M_n(F)$ be the set of all $n \times n$ matrices with entries in a field $F$. Then $(M_n(F), +, *)$ is a non-commutative ring with identity.
- $(\mathbb{R}^{\mathbb{R}}, +, *)$ is a commutative ring with identity, where addition and multiplication are defined pointwise.

## V.2 Primes

**Definition V.2.1** (Prime)**.** An integer $p$ is said to be *prime* if it has no non-trivial divisors (other than $\pm 1$ and $\pm p$). We additionally define 1 and $-1$ to not be primes.

*Notation.* We write $a \mid b$ if $b$ is divisible by $a$ and $a \nmid b$ if not. That is, $a \mid b \iff \exists k(ka = b)$. However, we leave $0 \mid 0$ undefined.

We assume some standard facts about divisibility:

- for all $a \neq 0$, $a \mid a$.
- if $a \mid b$ and $b \mid a$, then $a = \pm b$.
- if $a \mid b$ and $b \mid c$, then $a \mid c$.
- if $a \mid b$ and $a \mid c$, then $a \mid (b + c)$.

**Definition V.2.2** (Order and valuation)**.** Suppose $n \in \mathbb{Z}$ and $p$ is a prime such that $p^\alpha \mid n$ but $p^{\alpha+1} \nmid n$. Then we call $\alpha$ the *order of* $n$ *at* $p$ or the *p-adic valuation of* $n$, written $\operatorname{ord}_p(n) = \alpha$.

### V.2.1 Unique prime factorization

**Lemma V.2.3.** *If $a, b \in \mathbb{Z}$ and $b > 0$, then there exist $q, r \in \mathbb{Z}$ such that $a = qb + r$ and $0 \leq r < b$.*

*Proof.* Take the supremum of the set $X = \{k \in \mathbb{Z}, kb \leq a\}$ to be $q$ and $r = a - qb$. Since $q \in X$, we have $0 \leq r$. Since $q + 1 \notin X$, we have $r < b$. □

**Definition V.2.4.** Let $a_1, \ldots, a_n \in \mathbb{Z}$. Define $A = (a_1, \ldots, a_n)$ to be the set of all linear combinations of $a_1, \ldots, a_n$ over $\mathbb{Z}$.

Note that if $a, b \in A$, then $a \pm b \in A$ and $ka \in A$ for all $k \in \mathbb{Z}$. This motivates the following generalisation:

**Definition V.2.5** (Ideal)**.** A *left* (resp. *right*) *ideal* of a ring $R$ is a subring $I \subseteq R$ which is closed under multiplication by elements of $R$ on the left (resp. right). That is, for all $a \in I$ and $r \in R$, we have $ra \in I$ (resp. $ar \in I$).

For a commutative ring, we simply call it an *ideal.*

**Lemma V.2.6.** *If $a, b \in \mathbb{Z}$, then there exists a $d \in \mathbb{Z}$ such that $(a, b) = (d)$.*

*Proof.* Assume that both $a$ and $b$ are non-zero. Then there exist positive elements in $(a, b)$. Let $d$ be the smallest positive element in $(a, b)$. Then $(d) \subseteq (a, b)$.

For the reverse inclusion, suppose $c \in (a, b)$. Apply the previous lemma to conclude the existence of $q$ and $r$ such that $c = qd + r$ with $0 \leq r < d$. Since $c, d \in (a, b)$, $r = c - qd \in (a, b)$. But $d$ is the smallest positive element in $(a, b)$, so $r$ cannot be positive. Thus $r = 0$ and $c = qd \in (d)$. □

**Definition V.2.7** (GCD)**.** Let $(a, b) \in \mathbb{Z}$. An integer $d$ is a *greatest common divisor* (gcd) of $a$ and $b$ if it divides both $a$ and $b$, and any other common divisor also divides $d$.

**Lecture 21.**
Wed 28 Feb '24

**Corollary V.2.8.** *Let $a, b \in \mathbb{Z}$. If $(a, b) = (d)$, then $d = \gcd(a, b)$.*

*Proof.* Since $a, b \in (d)$, $d$ is a common divisor of both $a$ and $b$. Let $c$ be another common divisor. Then $c \mid ax + by$, so $c \mid d$. Thus $d$ is the greatest common divisor. □

*Notation.* We will write $(a, b)$ for the gcd of $a$ and $b$. Whether this refers to the gcd or the ideal will (should) be clear from the context.

**Definition V.2.9** (Coprime)**.** Two integers are said to be *coprime* if their only common divisors are ±1.

Thus $a$ and $b$ are coprime iff $(a, b) = 1$. There is a generalization of this to other rings, where instead of ±1 we say that two elements are coprime if their only common divisors are *units.*

**Proposition V.2.10.** *Suppose $(a, b) = 1$ and $a \mid bc$. Then $a \mid c$.*

*Proof.* There exist $x$, $y$ such that $ax + by = 1$. Then $c = cax + cby$. But $a \mid cb$, so $a \mid c$. □

**Corollary V.2.11.** *If $p$ is a prime and $p \mid bc$, then $p \mid b$ or $p \mid c$. Equivalently, if $p \nmid b$ and $p \nmid c$, then $p \nmid bc$.*

*Proof.* Since $p$ is a prime, its only divisors are $\pm 1$ and $\pm p$. Thus, either $(p, b) = 1$ or $p \mid b$. If $p \mid b$, then we are done. Otherwise, by the previous proposition, $p \mid c$. □

**Corollary V.2.12.** *Suppose $p$ is a prime and $a, b \in \mathbb{Z}$. Then $\operatorname{ord}_p(ab) = \operatorname{ord}_p(a) + \operatorname{ord}_p(b)$.*

*Proof.* Let $\alpha = \operatorname{ord}_p(a)$, $\beta = \operatorname{ord}_p(b)$ so that $a = p^\alpha a'$ and $b = p^\beta b'$ where $p \nmid a', b'$. Then $ab = p^{\alpha+\beta} a'b'$. By the previous corollary, $p \nmid a'b'$. Thus $\operatorname{ord}_p(ab) = \alpha + \beta$. □

**Lemma V.2.13** (Existence of prime factorization)**.** *Every integer $n \neq 0, \pm 1$ has a prime factorization.*

*Proof.* Let $n$ be the smallest positive integer without a prime factorization. Then $n$ is not prime, so $n = ab$ for some $a, b \in \mathbb{Z}$. But $a, b < n$ have prime factorizations, so $n$ has a prime factorization.

If every positive integer has a prime factorization, then so will the negative of any such integer, by taking an additional factor of $-1$. □

**Theorem V.2.14** (Fundamental theorem of arithmetic)**.** *Every integer $n \neq 0$ has a unique prime factorization.*

*Proof.* Write $n$ as

$$n = (-1)^{\epsilon(n)} \prod_{\substack{p \text{ prime} \\ p>0}} p^{a(p)}.$$

For any prime $q$, apply $\operatorname{ord}_q$ to both sides. Then

$$\operatorname{ord}_q(n) = \epsilon(n) \operatorname{ord}_q(-1) + \sum_p^q a(p) \operatorname{ord}_q(p)$$

by corollary V.2.12. But by the definition of $\operatorname{ord}_q$, $\operatorname{ord}_q(-1) = 0$ and $\operatorname{ord}_q(p) = \delta_{pq}$. Thus $a(q) = \operatorname{ord}_q(n)$ is uniquely determined. □

## V.3 In Other Rings

**Definition V.3.1** (Field)**.** A *field* is a commutative ring with identity $1 \neq 0$, where all non-zero elements have multiplicative inverses.

*Example.* $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, finite fields $\mathbb{F}_q$, where $q$ is a prime power.

**Definition V.3.2** (Ring of polynomials)**.** For a field $k$, $k[x]$ is the *ring of polynomials* in $x$ with coefficients from $k$. There is a notion of divisibility in $k[x]$. We thus write $f \mid g$ if $g = fp$ for some $p \in k[x]$.

A non-constant polynomial $p$ is *irreducible* if $q \mid p$ only when $q$ is constant or a multiple of $p$.

*Examples.*

- $3 \mid 1 + x$.
- Linear polynomials are always irreducible.
- $x^2 + 1$ is irreducible in $\mathbb{Q}[x]$ but not in $\mathbb{C}[x]$.

**Lemma V.3.3.** *Every non-constant polynomial is a product of irreducible polynomials.*

*Proof idea.* Same as for $\mathbb{Z}$, but we use induction on the degree of the polynomial. □

**Definition V.3.4** (Monic polynomial)**.** A polynomial is *monic* if its leading coefficient is 1.

**Definition V.3.5** (Order)**.** Let $f, p \in k[x]$, $p$ irreducible. Then $\operatorname{ord}_p(f) = a$ if $f = p^a q$ for some $q \in k[x]$ and $p \nmid q$.

**Theorem V.3.6** (Unique factorization of polynomials)**.** *Let $f \in k[x]$. Then we can write*

$$f = c \prod_p p^{a(p)}$$

*where the product runs over all monic irreducible polynomials, $a(p) = \operatorname{ord}_p(f)$ and $c \in k$.*

**Lecture 22.** Fri 01 Mar '24

**Definition V.3.7** (Integral domain)**.** An *integral domain* is a commutative ring with no zero divisors.

For integral domains, the cancellation law holds. $ac = bc \wedge c \neq 0 \implies a = b$.

*Example.* $\mathbb{Z}$, $k[x]$.

**Definition V.3.8** (Euclidean domain)**.** A *Euclidean domain* is an integral domain $R$ together with a function $\lambda\colon R^* \to \mathbb{N}$ such that if $a, b \in R$ with $b \neq 0$, there exist $c, d \in R$ with $a = cb + d$, then either $d = 0$ or $\lambda(d) < \lambda(b)$.

Recall that for $a_1, \ldots, a_n \in R$,

$$(a_1, \ldots a_n) = \{x_1 a_1 + \ldots x_n a_n \mid x_1, \ldots, x_n \in R\}$$

is the ideal generated by $a_1, \ldots, a_n$.

**Definition V.3.9** (Principal ideals)**.** If an ideal $I$ can be written as $I = (a_1, \ldots, a_n)$, we say $I$ is *finitely generated.* If $I = (a)$, we say that $I$ is a *principal ideal.* An integral domain is called a *principal ideal domain* (PID) if all finitely generated ideals are principal.

*Example.* $\mathbb{Z}$ is a PID.

**Proposition V.3.10.** *Every Euclidean domain is a principal ideal domain.*

*Proof.* Let $I$ be an ideal in a Euclidean domain $R$. Consider the set $\{\lambda(b) \mid b \in I^*\} \subseteq \mathbb{N}$. So there exists a minimal element $a \in I^*$ such that $\lambda(a) \leq \lambda(b)$ for all $b \in I^*$.

We claim that $I = (a) = Ra = \{ra \mid r \in R\}$. Since $a \in I$ and $I$ is an ideal, $Ra \subseteq I$. Let $b \in I$. Then there exist $q, r \in R$ such that $b = qa + r$ with $r = 0$ or $\lambda(r) < \lambda(a)$. But $r = b - qa \in I$. Since $\lambda(a)$ is minimal, $r = 0$, which gives $b = qa \in Ra$ and $I \subseteq Ra$. □

The converse is false, but it is hard to find a counterexample.

**Definition V.3.11.** Let $R$ be a principal ideal domain.

- For $a \in R$, $b \in R^*$, we say that $a$ *divides* $b$ (denoted $a \mid b$) if $b = ac$ for some $c \in R$. In other words, $(b) \subseteq (a)$.
- An element $u \in R$ is called a *unit* if $u \mid 1$. In other words, $(u) = R$.
- Two elements $a, b \in R$ are called *associates* if $a = bu$ for some unit $u \in R$. In other words, $(a) = (b)$.
- A non-unit $p \in R$ is called a *prime* if $p \neq 0$ and for all $a, b \in R$, $p \mid ab$ only if $p \mid a$ or $p \mid b$. In other words, if $ab \in (p)$, then $a \in (p)$ or $b \in (p)$.

**Exercise V.3.12.** *Prove the "in other words" above.*

### V.3.1 Unique factorization for PIDs

- Show that the greatest common divisor of $a, b \in R$ exists and is unique up to associates, and $(a, b) = (d)$.
- We can find for every $a$ and $p$ prime, the *order* $\operatorname{ord}_p(a)$, which satisfies $\operatorname{ord}_p(ab) = \operatorname{ord}_p(a) + \operatorname{ord}_p(b)$.

Let $S$ be a set of primes in $R$ satisfying

(i) every prime in $R$ is associate to some prime in $S$, and

(ii) no two primes in $S$ are associates.

**Theorem V.3.13** (Unique factorization theorem)**.** *Let $R$ be a principal ideal domain and $S$ be as above. Then for all $a \in R^*$, we can write*

$$a = u \prod_{p \in S} p^{e(p)}$$

*where $e(p) = \operatorname{ord}_p(a)$ and $u$ is a unit. Further, this is unique.*

**Definition V.3.14** (Unique factorization domain)**.** A domain $R$ for which unique factorization holds is called a *unique factorization domain* (UFD).

*Examples.*

- $\mathbb{Z}$ is a UFD.
- $k[x_1, \ldots, x_n]$ is a UFD but not a PID.
- $\mathbb{Z}[\sqrt{3}i]$ is a ring. It is also an integral domain by virtue of being a subring of $\mathbb{C}$. $2, 1 \pm \sqrt{3}i$ are primes (absolute value 2 is minimal). The only units are $\pm 1$, so no two are associates of each other. But $4 = 2 * 2 = (1 + \sqrt{3}i)(1 - \sqrt{3}i)$. Thus $\mathbb{Z}[\sqrt{3}i]$ is not a UFD.
- $\mathbb{Z}[\sqrt{7}]$ has $6 = 2 * 3 = (\sqrt{7} + 1)(\sqrt{7} - 1)$. But 2 and 3 are not prime! (exercise) $\mathbb{Z}[\sqrt{7}]$ does turn out to be a UFD.

**Fact V.3.15** (Gauss' conjecture)**.** *Let $d$ be a square-free positive integer. Consider $\mathbb{Q}[i\sqrt{d}]$. The subring of algebraic integers in it is a UFD iff $d$ is a* Heegner *number. That is,*

$$d \in \{1, 2, 3, 7, 11, 19, 43, 67, 163\}.$$

**Lecture 23.**
Mon 04 Mar '24

If $d = 1$, this subring is $\mathbb{Z}[i]$. But if $d = 3$, it is $\mathbb{Z}[e^{i\pi/3}]$, not $\mathbb{Z}[i\sqrt{3}]$.

*Examples* (UFD).

- $\mathbb{Z}[i]$, the *Gaussian integers.*
- $\mathbb{Z}[\omega]$, the *Eisenstein integers*, where $\omega = e^{\frac{2\pi i}{3}}$.

**Proposition V.3.16.** *$\mathbb{Z}[i]$ is a Euclidean domain.*

*Proof.* Define $\lambda\colon \mathbb{Q}[i] \to \mathbb{N}$ as $\lambda(a + ib) = a^2 + b^2$. Let $\alpha = a + ib$, $\gamma = c + id \neq 0$. Write $\frac{\alpha}{\gamma} = r + is$, where $r, s \in \mathbb{Q}$. Choose $m, n \in \mathbb{Z}$ such that $|r - m| \leq \frac{1}{2}$ and $|s - n| \leq \frac{1}{2}$. Let $\delta = m + in$. Then $\lambda(\frac{\alpha}{\gamma} - \delta) = (r - m)^2 + (s - n)^2 \leq \frac{1}{2}$. Define $\rho = \alpha - \gamma\delta$, Either $\rho = 0$, or

$$\begin{aligned} \lambda(\rho) &= \lambda(\gamma)\lambda\left(\frac{\alpha}{\gamma} - \delta\right) \\ &\leq \frac{1}{2}\lambda(\gamma) \\ &< \lambda(\gamma). \end{aligned}$$

So $\alpha = \delta\gamma + \rho$ with $\rho = 0$ or $\lambda(\rho) < \lambda(\gamma)$. □

**Corollary V.3.17.** *$\mathbb{Z}[i]$ is a PID and hence a UFD.*

**Exercise V.3.18.** *Prove that $\mathbb{Z}[\omega]$ is a Euclidean domain.*

# Chapter VI

# The Study of Primes

**Theorem VI.0.1** (Euclid)**.** *There are infinitely many primes in $\mathbb{Z}$.*

*Proof.* Suppose not. Label the positive primes $p_1, p_2, \ldots, p_n$. Define $N = p_1 p_2 \ldots p_n + 1$. Clearly, $N$ is not divisible by any $p_i$. But $N$ must be a product of primes. This is a contradiction. □

*Remark.* Check out the proofs of this theorem in *Proofs from THE BOOK*.

**Exercise VI.0.2.** *There are infinitely many monic irreducible polynomials in $k[x]$, assuming $k$ is infinite.*

*Proof.* $x + a$ for each $a \in k$. □

## VI.1 Arithmetic Functions

- $\nu(n)$ = number of positive divisors of $n$.
- $\sigma(n)$ = sum of positive divisors of $n$.
- The *Möbius function*

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \text{ is not square-free} \\ (-1)^{\#\text{ prime factors of } n} & \text{otherwise} \end{cases}$$

- The *Euler totient function*

$$\phi(n) = \#\{1 \leq m \leq n \mid \gcd(m, n) = 1\}$$

*Examples.*

- $\nu(3) = 2$, $\sigma(3) = 4$, $\mu(3) = -1$, $\phi(3) = 2$.
- $\nu(6) = 4$, $\sigma(6) = 12$, $\mu(6) = 1$, $\phi(6) = 2$.
- $\sigma(28) = 56$, since it is a perfect number.

**Proposition VI.1.1.** *Write $n$ as $n = p_1^{a_1} p_2^{a_2} \dots p_l^{a_l}$ in terms of its prime factors. Then*

*(i)* $\nu(n) = (a_1 + 1)(a_2 + 1) \dots (a_l + 1)$.

*(ii)* $\sigma(n) = (1 + p_1 + \dots p^{a_l}) \dots (1 + p_l + \dots + p_l^{a_l})$.

*Proof.* For the first part, every $l$-tuple $(b_1, \dots, b_l)$ can be transformed bijectively to a divisor of $n$.

For the second, write

$$\begin{aligned}\sigma(n) &= \sum_{d|n} d \\ &= \sum_{\substack{0 \le b_i \le a_i \\ 1 \le i \le l}} p_1^{b_1} \dots p_l^{b_l} \\ &= \prod_{i=1}^{l} \sum_{0 \le b_i \le a_i} p_i^{b_i} \\ &= \prod_{i=1}^{l} \frac{p_i^{a_i+1} - 1}{p_i - 1}.\end{aligned}$$

□

**Proposition VI.1.2.** $\sum_{d|n} \mu(d) = \delta_{n,1}$.

*Proof.* True for $n = 1$. For $n > 1$, write $n$ as $p_1^{a_1} \dots p_l^{a_l}$. Since $\mu(d) = 0$ whenever $d$ is not square-free, we have

$$\begin{aligned}\sum_{d|n} \mu(d) &= \sum_{\substack{b_i \in \{0,1\} \\ 1 \le i \le l}} \mu(p_1^{b_1} \dots p_l^{b_l}) \\ &= \sum_{k=0}^{l} \binom{l}{k} (-1)^k \\ &= (1 - 1)^k \\ &= 0\end{aligned}$$

□

**Definition VI.1.3** (Dirichlet convolution)**.** Let $f, g\colon \mathbb{N}^* \to \mathbb{C}$. Then the *Dirichlet convolution* of $f$ and $g$ is

$$(f \circ g)(n) = \sum_{d|n} f(d) g(\frac{n}{d}) = \sum_{d_1 d_2 = n} f(d_1) g(d_2)$$

**Exercise VI.1.4.** $(f \circ g) \circ h = f \circ (g \circ h)$.

Let $\varepsilon(n)$ be the multiplicative identity. That is, $\varepsilon(n) = \delta_{n,1}$. Check that $f \circ \varepsilon = \varepsilon \circ f = f$. Let $\mathbb{1}$ be the constant function $\mathbb{1}(n) = 1$ for all $n$. Then $f \circ \mathbb{1} = \mathbb{1} \circ f = \sum_{d|\cdot} f(d)$.

**Lemma VI.1.5.** *$\mathbb{1} \circ \mu = \mu \circ \mathbb{1} = \varepsilon$.*

*Proof.* First,

$$\begin{aligned}(\mathbb{1} \circ \mu)(1) = (\mu \circ \mathbb{1})(1) &= \sum_{d|1} \mu(d) \\ &= \mu(1) \\ &= 1.\end{aligned}$$

For $n > 1$,

$$\begin{aligned}(\mathbb{1} \circ \mu)(n) = (\mu \circ \mathbb{1})(n) &= \sum_{d|n} \mu(d) \\ &= 0\end{aligned}$$

by proposition VI.1.2. □

**Theorem VI.1.6** (Möbius inversion formula)**.** *Let $F(n) = \sum_{d|n} f(d)$. Then $f(n) = \sum_{d|n} \mu(d) F(\frac{n}{d})$.*

*Proof.* Note that $F = f \circ \mathbb{1}$. So

$$\begin{aligned}\sum_{d|\cdot} \mu(d) F\left(\frac{\cdot}{d}\right) &= F \circ \mu \\ &= (f \circ \mathbb{1}) \circ \mu \\ &= f \circ (\mathbb{1} \circ \mu) \\ &= f \circ \varepsilon \\ &= f\end{aligned}$$

□

**Lecture 24.**
Wed 06 Mar '24

**Definition VI.1.7** (Dirichlet series)**.** Given a function $f\colon \mathbb{N}^* \to \mathbb{C}$, the *Dirichlet series* (or *generating function*) associated to $f$ is

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

for $s$ in some subset of $\mathbb{C}$.

Recall the Dirichlet convolution,

$$(f \circ g)(n) = \sum_{d|n} f(d) g\Big(\frac{n}{d}\Big)$$

**Exercise VI.1.8.** *The sequence corresponding to a product of two Dirichlet series is the corresponding Dirichlet convolution.*

**Definition VI.1.9** (Multiplicative)**.** A function $f\colon \mathbb{N}^* \to \mathbb{C}$ is called *multiplicative* if $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$.

**Fact VI.1.10.** *Let $f$ be a multiplicative sunction. Then*

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p \ prime} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right)$$

*In particular, for $\mathbb{1}$,*

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \ prime} \frac{1}{1 - p^{-s}}$$

*is the* Riemann zeta function*.*

If $F(s)$ is such that $F(1) \neq 0$, then $F(s)^{-1}$ is also a Dirichlet series. In particular, $\frac{1}{\zeta}$ corresponds to the Möbius function $\mu$.

Recall Euler's totient function $\phi(n)$, which counts the number of positive integers less than $n$ that are coprime to it.

**Proposition VI.1.11.**

$$\sum_{d|n} \phi(d) = n$$

*Proof.* Write the rationals $\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, \frac{n}{n}$ in lowest terms. Now if $d \mid n$, exactly $\phi(d)$ of these rationals have denominator $d$. Why? Suppose $n = d_1 d$. Then $\frac{d_1 k}{n} = \frac{k}{d} \leq 1$ will be in lowest terms iff $(k, d) = 1$ and $k < d$. Conversely, all denominators of the rationals are divisors of $n$. Thus, $\sum_{d|n} \phi(d)$ simply counts the number of these rationals, which is $n$. □

*Example.*

$$\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}, \frac{6}{6} = \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, \frac{1}{1}$$

has $\phi(1) = 1$, $\phi(2) = 1$, $\phi(3) = 2$ and $\phi(6) = 2$.

**Proposition VI.1.12.** *If* $n = p_1^{a_1} \dots p_l^{a_l}$, *then*

$$\phi(n) = n \prod_{i=1}^{l} \left(1 - \frac{1}{p_i}\right)$$

*Proof.* By proposition VI.1.11 and Möbius inversion formula,

$$\begin{aligned} \phi(n) &= \sum_{d|n} \mu(d) \frac{n}{d} \\ &= \frac{n}{1} - \sum_{i=1}^{l} \frac{n}{p_i} + \sum_{i<j} \frac{n}{p_i p_j} - \cdots \\ &= n \prod_{i=1}^{l} \left(1 - \frac{1}{p_i}\right) \end{aligned}$$

(the last step is similar to Viete's formulas). □

**Definition VI.1.13** (Prime counting function)**.** The *prime counting function*

$$\pi(x) = \#\{1 \leq p \leq x \mid p \text{ is prime}\}.$$

**Proposition VI.1.14.**

$$\pi(x) \geq \log \log x$$

*Proof.* Let $p_n$ be the $n$th prime. Since $p_1, \dots, p_n$ cannot divide $p_1 \dots p_n + 1$, we have that $p_{n+1} \leq p_1 p_2 \dots p_n + 1$.

**Claim:** $p_n < 2^{2^n}$. True for $n = 1$. Suppose true for $n$. Then

$$\begin{aligned} p_{n+1} &\leq p_1 p_2 \dots p_n + 1 \\ &< 2^{2^1} \dots 2^{2^n} + 1 \\ &= 2^{2^1 + \dots + 2^n} + 1 \\ &= 2^{2^{n+1} - 2} + 1 \\ &< 2^{2^{n+1}} \end{aligned}$$

This proves the claim.

So $\pi(2^{2^n}) \geq n$. Let $x > e$ and choose $n$ such that $e^{e^{n-1}} < x \leq e^{e^n}$. So $n \geq \log \log x$. Since $2/3 < \log 2$, if $n > 3$ we get $e^{n-1} > 2^n$. So

$$\begin{aligned} \pi(x) &\geq \pi(e^{e^{n-1}}) \\ &\geq \pi(e^{2^n}) \\ &\geq \pi(2^{2^n}) \\ &\geq n \\ &\geq \log \log x \end{aligned} \qquad \square$$

**Fact VI.1.15.**

$$\frac{\pi(x)}{x} \to 0 \quad \textit{as} \quad x \to \infty$$

**Fact VI.1.16** (Prime number theorem)**.**

$$\pi(x) \sim \frac{x}{\log x} \quad \textit{as} \quad x \to \infty$$

Proved by Hadamard and de la Vallée Poussin in 1896.

**Theorem VI.1.17** (Dirichlet)**.** *Let $a, d \in \mathbb{Z}$ with $(a, d) = 1$. Then there are infinitely many primes of the form $a + nd$.*

# Chapter VII

# Congruences

Let $e/o$ denote an even/odd number. Then we can form the addition and multiplication tables given in table VII.1.

To find solutions of polynomial equations, congruences are useful. We want to find roots of polynomials in $\mathbb{Z}[x]$ which are themselves integers. These are called *Diophantine equations.*

*Example.* Can $x^2 - 117x + 31 = 0$ have integer solutions?

No. Suppose $x$ is even/odd. Then the first two terms are both even/odd, so their sum is even. But 31 is odd, so the total sum is odd and hence never 0.

**Lecture 25.**
Wed 13 Mar '24
**Lecture 26.**
Wed 13 Mar '24

**Proposition VII.0.1.** *There are exactly $\phi(m)$ units in $\mathbb{Z}/m\mathbb{Z}$.*

*Proof.* $a \in \mathbb{Z}/m\mathbb{Z}$ is a unit iff $ax \equiv 1 \pmod{m}$ for some $x \in \mathbb{Z}/m\mathbb{Z}$. This is equivalent to $(a, m) = 1$, and there are $\phi(m)$ such $a$ in $\{0, 1, \ldots, m-1\}$. □

**Corollary VII.0.2.** *$\mathbb{Z}/p\mathbb{Z}$ is a field iff $p$ is prime.*

*Proof.* If $p$ is prime, then every element is a unit.

Conversely, if $p = p_1p_2$, then $\overline{p_1}, \overline{p_2} \neq \bar{0}$, but $\overline{p_1p_2} = \bar{0}$. So $\mathbb{Z}/p\mathbb{Z}$ is not a field. □

Table VII.1: Addition and multiplication tables modulo 2

| $+$ | $e$ | $o$ |
|---|---|---|
| $e$ | $e$ | $o$ |
| $o$ | $o$ | $e$ |

(a) Addition

| $\times$ | $e$ | $o$ |
|---|---|---|
| $e$ | $e$ | $e$ |
| $o$ | $e$ | $o$ |

(b) Multiplication

*Notation.* We will denote by $U(\mathbb{Z}/m\mathbb{Z})$ the set of units in $\mathbb{Z}/m\mathbb{Z}$.

**Lemma VII.0.3.** *$U(\mathbb{Z}/m\mathbb{Z})$ forms a group under multiplication.*

*Proof.* If $a$ and $b$ are units, then so is $ab$.

1 is a unit and an identity.

If $a$ is a unit, there exists a unique $x$ such that $ax \equiv 1 \pmod{m}$. Then $x$ is a unit and the unique inverse of $a$. □

**Theorem VII.0.4** (Euler)**.** *If $(a, m) = 1$, then $a^{\phi(m)} \equiv 1 \pmod{m}$.*

*Proof.* By the previous lemma, $a \in G = U(\mathbb{Z}/m\mathbb{Z})$ and $|G| = \phi(m)$. Consider the map $\psi \colon G \to G$ given by $\psi(x) = ax$.

**Claim:** $\psi$ is a bijection.

**Proof of claim:** Since $G$ is a group, the inverse of $a$ exists. Suffices to show that $\psi$ is injective (finite set). $\psi(x) = \psi(y) \iff ax = ay \iff x = y$.

Using this claim, we can write

$$\begin{aligned} \prod_{x \in G} ax &= \prod_{x \in G} x \\ a^{\phi(m)} \prod_{x \in G} x &= \prod_{x \in G} x \\ a^{\phi(m)} &= 1 \end{aligned}$$

□

**Corollary VII.0.5** (Fermat's little theorem)**.** *If $p$ is prime and $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.*

*Proof.* $\phi(p) = p - 1$. □

**Lemma VII.0.6.** *If $a_1, a_2, \ldots, a_j$ are coprime to $m$, then so is $a_1 a_2 \ldots a_j$.*

*Proof.* They are all units, so their product is a unit. □

**Lemma VII.0.7.** *If $a_1, a_2, \ldots, a_j$ divide $m$ and $(a_i, a_j) = 1$ for all $i \neq j$, then $a_1 a_2 \ldots a_j$ divides $m$.*

*Proof.* Induction. The base case $j = 1$ is obvious.

Suppose the statement is true for $a_1, a_2, \ldots, a_{j-1}$. Then by the previous lemma, $a_1 a_2 \ldots a_{j-1}$ is coprime to $a_j$. So we can write $r \cdot a_1 \ldots a_{j-1} + s \cdot a_j = 1$. Multiplying by $m$, we get

$$r \cdot a_1 \ldots a_{j-1} m + s \cdot a_j m = m.$$

But $a_j$ divides the $m$ in the first term, and by the induction hypothesis, $a_1 \ldots a_{j-1}$ divides the $m$ in the second term. So $a_1 \ldots a_j$ divides $m$. □

**Theorem VII.0.8** (Chinese remainder theorem)**.** *Write $m = m_1 \dots m_k$ with $(m_i, m_j) = 1$ for all $i \neq j$. Let $b_1, \dots, b_j \in \mathbb{Z}$ and consider the system of congruences*

$$\begin{aligned} x &\equiv b_1 \pmod{m_1} \\ x &\equiv b_2 \pmod{m_2} \\ &\vdots \\ x &\equiv b_k \pmod{m_k}. \end{aligned}$$

*Then the system always has solutions and any two solutions differ by a multiple of $m$.*

*Proof.* Let $n_i = \frac{m}{m_i} = m_1 \dots m_{i-1} m_{i+1} \dots m_k$. Each $m_j$, $j \neq i$, is coprime to $m_i$, so by lemma VII.0.6, $(m_i, n_i) = 1$. Thus we have $r_i$ and $s_i$ such that $r_i m_i + s_i n_i = 1$. Let $e_i = s_i n_i$. Then $e_i \equiv 1 \pmod{m_i}$. Since each $m_j \neq n_j$ divides $m$, $e_i \equiv 0 \pmod{m_j}$ for all $j \neq i$.

This gives a solution

$$x_0 = b_1 e_1 + b_2 e_2 + \dots + b_k e_k.$$

Suppose $x_1$ is another solution. Then $x_1 - x_0 \equiv 0 \pmod{m_i}$ for all $i$. So each of $m_1$, $m_2$, …, $m_k$ divides $x_1 - x_0$. By lemma VII.0.7, $m$ divides $x_1 - x_0$. □

*Example* (Original example of Sunzi). A certain number leaves a remainder of 2 when divided by 3, a remainder of 3 when divided by 5, and a remainder of 2 when divided by 7. What is the number?

We have

$$\begin{aligned} m_1 &= 3 & \qquad m_2 &= 5 & \qquad m_3 &= 7 \\ b_1 &= 2 & \qquad b_2 &= 3 & \qquad b_3 &= 2 \end{aligned}$$

and compute

$$n_1 = 35 \qquad\qquad n_2 = 21 \qquad\qquad n_3 = 15.$$

We want

$$3r_1 + 35s_1 = 1 \qquad\qquad 5r_2 + 21s_2 = 1 \qquad\qquad 7r_3 + 15s_3 = 1.$$

One solution is

$$r_1, s_1 = 12, -1 \qquad\qquad r_2, s_2 = -4, 1 \qquad\qquad r_3, s_3 = -2, 1.$$

This gives

$$e_1 = -35 \qquad\qquad e_2 = 21 \qquad\qquad e_3 = 15,$$

and finally the solution

$$\begin{aligned} x &= 2(-35) + 3(21) + 2(15) \\ &= -70 + 63 + 30 \\ &= 23. \end{aligned}$$

**Lecture 27.** Fri 15 Mar '24

**Proposition VII.0.9.** *If $R_1, \dots, R_n$ are rings, then $S = R_1 \times \cdots \times R_n$ is also a ring under componentwise addition and multiplication.*

*Proof.* Zero is $(0, \dots, 0)$ and one is $(1, \dots, 1)$. Inverses are also componentwise. Everything else works componentwise. □

**Exercise VII.0.10.** *$u = (u_1, \dots, u_n)$ is a unit in $S$ iff each $u_i$ is a unit in $R_i$.*

**Theorem VII.0.11.** *If $m = m_1 \dots m_k$ and $(m_i, m_j) = 1$ for all $i < j$, then*

$$\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}.$$

*That is, they are isomorphic as rings.*

*Proof.* Define $\psi_i \colon \mathbb{Z} \to \mathbb{Z}/m_i\mathbb{Z}$ as $\psi_i(a) = a \bmod m_i$. Define $\psi = (\psi_1, \dots, \psi_k)$.

By the Chinese Remainder Theorem, $\psi(n) = (b_1, \dots, b_k)$ always has a solution, so $\psi$ is surjective.

If $\psi(n) = 0$, then $n \equiv 0 \pmod{m_i}$ for all $i$, so $n \equiv 0 \pmod{m}$. Thus $\psi$ can be restricted to $\mathbb{Z}/m\mathbb{Z}$ in a natural way, and is then a bijection since its domain and codomain have the same size.

It is easy to check that $\psi$ respects addition and multiplication. □

**Corollary VII.0.12.**

$$U(\mathbb{Z}/m\mathbb{Z}) \cong U(\mathbb{Z}/m_1\mathbb{Z}) \times \cdots \times U(\mathbb{Z}/m_k\mathbb{Z}).$$

Thus we can restrict our attention to the study of $U(\mathbb{Z}/p^n\mathbb{Z})$ for $p$ prime.

**Lemma VII.0.13.** *Let $k$ be a field and $f \in k[x]$ with $\deg f = n$. Then $f$ has at most $n$ distinct roots in $k$.*

*Proof.* Induction. Trivial for $n = 1$.

If $f$ has no roots in $k$, we are done. Otherwise, let $\alpha$ be a root of $f$. Divide $f$ by $(x - \alpha)$ to get $f(x) = (x - \alpha)q(x) + r$. $r$ has degree less than $(x - a)$, so $r$ is a constant and hence 0.

Thus $f(x) = (x - \alpha)q(x)$ where $q$ has degree $n - 1$. Suppose $\beta \neq \alpha$ is a root of $f$. Then $0 = f(\beta) = (\beta - \alpha)q(\beta)$, so $\beta$ is a root of $q$.

But by the induction hypothesis, $q$ has at most $n - 1$ roots, so $f$ has at most $n$ roots. Winduction. □

*Remark.* If $k$ is not a field, this need not hold. For example, let $k = \mathbb{Z}/4\mathbb{Z}$ and let $f(x) = 2x(x+1)$. Then $0, 1, 2, 3$ are all roots of $f$.

What's wrong? $\mathbb{Z}/4\mathbb{Z}$ has zero divisors. In fact, the above lemma can be generalized to any integral domain.

**Corollary VII.0.14.** *Let $f, g \in k[x]$ with $\deg f = \deg g = n$. If $f$ and $g$ agree at $n+1$ points, then $f = g$.*

*Proof.* Take the difference. This has degree at most $n$ but has $n+1$ roots, so it is the zero polynomial. □

**Proposition VII.0.15.** *For any prime $p$,*

$$x^{p-1} - 1 \equiv (x-1)(x-2)\dots(x-(p-1)) \pmod{p}$$

*for all $x$.*

*Proof.* View this polynomial over the field $\mathbb{Z}/p\mathbb{Z}$. Let $f$ be the difference of the two sides,

$$f(x) = x^{p-1} - 1 - (x-1)(x-2)\dots(x-(p-1)).$$

Note that the $x^{p-1}$ term cancels out, so $\deg f \le p-2$.

By Fermat's little theorem, $x^{p-1} = 1$ for all $x \neq 0$. Thus $f(x) = 0$ for all $x \neq 0$. Thus $f$ has at least $p-1$ roots, so it must be the zero polynomial. □

**Corollary VII.0.16** (Wilson's theorem)**.** *If $p$ is prime, then*

$$(p-1)! \equiv -1 \pmod{p}.$$

*Proof.* Set $x = 0$ in the above proposition. $p = 2$ is verified by hand. Every other prime is odd, so the powers of $-1$ on the RHS cancel out. □

**Proposition VII.0.17.** *If $p$ is prime and $d \mid p-1$, then $x^d \equiv 1 \pmod{p}$ has $d$ solutions.*

*Proof.* Let $d' = (p-1)/d$. Then

$$\begin{aligned} \frac{x^{p-1}-1}{x^d-1} &= \frac{(x^d)^{d'}-1}{x^d-1} \\ &= 1 + x^d + \cdots + (x^d)^{d'-1} \\ \implies x^{p-1} - 1 &= (x^d - 1)g(x) \end{aligned}$$

where $g(x)$ has degree $dd' - d = p - 1 - d$. By the previous proposition, $x^{p-1} - 1$ has $p-1$ roots, so $x^d - 1$ has at least $d$ roots. Since $x^d - 1$ has degree $d$, it has exactly $d$ roots. □

**Definition VII.0.18** (Cyclic group)**.** A group $H$ is said to be *cyclic* if it is generated by a single element $x$, *i.e.*,

$$H = \{x^n \mid n \in \mathbb{Z}\}.$$

*Examples.*

- $(\mathbb{Z}, +)$ is cyclic, generated by 1.
- $(\mathbb{Z}/n\mathbb{Z}, +)$ is cyclic, generated by $\bar{1}$.
- $(\mathbb{Z}/4\mathbb{Z}, +)$ is generated by $\bar{1}$ and $\bar{3}$, but not by $\bar{2}$, which only generates a subgroup.

**Definition VII.0.19** (Order of an element)**.** The *order* of an element $x \in H$ is the smallest positive integer $n$ such that $x^n = 1$. If no such $n$ exists, we say that $x$ has *infinite order.*

*Examples.*

- In $(\mathbb{Z}, +)$, 1 has infinite order.
- In $(\mathbb{Z}/4\mathbb{Z}, +)$, $\bar{1}$ has order 4 but $\bar{2}$ has order 2.

**Theorem VII.0.20** (Gauss)**.** *If $p$ is prime, then $G = U(\mathbb{Z}/p\mathbb{Z})$ is cyclic.*

*Proof.* For a divisor $d \mid p-1$, define $\psi(d)$ to be the number of elements of orger $d$ in $G$.

By proposition VII.0.17, $x^d - 1$ has $d$ solutions in $\mathbb{Z}/p\mathbb{Z}[x]$. Thus there are $d$ elements whose $d$th power is 1. Thus

$$\sum_{c|d} \psi(c) = d.$$

By Möbius inversion,

$$\sum_{c|d} \mu(c)\frac{d}{c} = \psi(d).$$

By **??**,

$$\psi(d) = \phi(d).$$

In particular, $\psi(p-1) = \phi(p-1)$.

If $p = 2$, then $|G| = 1$ makes the result trivial. If $p > 2$, then $\phi(p-1) > 1$, so there exists an element with order $p-1$. That element generates $G$. □

*Example.* For $p = 5$, $U(\mathbb{Z}/p\mathbb{Z}) = \{1, 2, 3, 4\}$. Then

$$\begin{array}{llll} 2^1 \equiv 2 & 2^2 \equiv 4 & 2^3 \equiv 3 & 2^4 \equiv 1 \\ 3^1 \equiv 3 & 3^2 \equiv 4 & 3^3 \equiv 2 & 3^4 \equiv 1 \\ 4^1 \equiv 4 & 4^2 \equiv 1. & & \end{array}$$

So the group is cyclic, with $\phi(5) = 2$ choices for the generator.

**Lecture 28.**
Mon 18 Mar '24

**Definition VII.0.21** (Primitive root)**.** An integer $a \in \mathbb{Z}/n\mathbb{Z}$ is called a *primitive root modulo* $n$ if it generates $U(\mathbb{Z}/n\mathbb{Z})$.

*Examples.*

- 2 is a primitive root modulo 5, but not modulo 7.
- There are no primitive roots modulo 12.

**Fact VII.0.22.**

*(i) If $p$ is an odd prime, then $U(\mathbb{Z}/p^l\mathbb{Z})$ is cyclic.*

*(ii) $U(\mathbb{Z}/2\mathbb{Z})$ and $U(\mathbb{Z}/4\mathbb{Z})$ are cyclic, but for $l \geq 3$, $U(\mathbb{Z}/2^l\mathbb{Z})$ is a product of cyclic groups of order 2 and $2^{l-2}$.*

**Theorem VII.0.23.** *Let $n = 2^a p_1^{a_1} \dots p_l^{a_l}$ be the prime decomposition of $n$. Then*

$$U(\mathbb{Z}/n\mathbb{Z}) \cong U(\mathbb{Z}/2^a\mathbb{Z}) \times \underbrace{U(\mathbb{Z}/p_1^{a_1}\mathbb{Z}) \times \cdots \times U(\mathbb{Z}/p_l^{a_l}\mathbb{Z})}_{cyclic}$$

**Corollary VII.0.24.** *An integer possesses primitive roots iff $n$ is of the form 2, 4, $p^a$, or $2p^a$ for some odd prime $p$.*

# Chapter VIII

# Quadratic Residues

**Definition VIII.0.1** (Quadratic residue)**.** Let $(a, m) = 1$. Then $a$ is called a *(quadratic) residue modulo $m$* if there is an $x \in \mathbb{Z}/m\mathbb{Z}$ such that $x^2 \equiv a \pmod{m}$.

If not, we say that $a$ is a *(quadratic) non-residue modulo $m$.*

*Example.* For $m = 7$,

$$1^2, 2^2, \dots, 6^2 \equiv 1, 4, 2, 2, 4, 1 \pmod{7}.$$

So the quadratic residues modulo 7 are precisely 1, 2, and 4.

**Proposition VIII.0.2.** *Let $m = 2^e p_1^{e_1} \dots p_l^{e_l}$ and let $(a, m) = 1$. Then $x^2 \equiv a \pmod{m}$ has a solution iff*

- *if $e = 2$, then $a \equiv 1 \pmod{4}$ and if $e \geq 3$, then $a \equiv 1 \pmod{8}$; and*
- *for each $i$, $a^{(p_i-1)/2} \equiv 1 \pmod{p_i}$.*

*Proof.* By the proof of the Chinese Remainder Theorem, $x^2 \equiv a \pmod{m}$ has a solution iff

$$x^2 \equiv a(2^e), \quad x^2 \equiv a(p_1^{e_1}), \quad \dots, \quad x^2 \equiv a(p_l^{e_l})$$

We restrict our attention to $x^2 \equiv a \pmod{2^e}$ to get the first condition. First note that 1 is the only quadratic residue modulo 4 and modulo 8.

**Lemma VIII.0.3.** *If $x^2 \equiv a \pmod{8}$, where $a$ is odd, has a solution, then so does $x^2 \equiv a \pmod{2^e}$ for all $e \geq 3$.*

*Proof.* Induction. Suppose $x_0^2 \equiv a \pmod{2^e}$. Note that since $a$ is odd, so is $x_0$. Take $x_0^2 = s2^e + a$.

Let $x_1 = x_0 - s2^{e-1}$. Then modulo $2^{e+1}$,

$$\begin{aligned} x_1^2 &\equiv (x_0 - s2^{e-1})^2 \\ &\equiv x_0^2 + sx_0 2^e + s^2 2^{2e-2} \\ &\equiv a + s2^e + sx_0 2^e + s^2 2^{e-3} 2^{e+1} \\ &\equiv a + s(1 + x_0) 2^e + 0 \\ &\equiv a \end{aligned}$$

since $x_0$ is odd. Winduction. □

**Lemma VIII.0.4.** *Let $p$ be an odd prime, $p \nmid a$ and $p \nmid n$. Then if $x^n \equiv a \pmod{p}$ has a solution, so does $x^n \equiv a \pmod{p^e}$ for all $e \geq 1$.*

*Proof.* Induction. If $x_0$ is a solution of $x^n \equiv a \pmod{p^e}$, then let $x_0^n = sp + a$ and choose $x_1 = x_0 + bp^{e-1}$, where $b$ is chosen so that $b \equiv -sx_0 \pmod{p}$. (This exists since $b$ is non-zero modulo $p$.) This gives a solution modulo $p^{e+1}$. Winduction. □

Thus, it suffices to look at $x^2 \equiv a \pmod{p_i}$.

Let $g$ be a primitive root modulo $p = p_i$ and let $x = g^y$. We wish to solve $g^{2y} = g^b \pmod{p}$. It is enough to solve for $2y \equiv b \pmod{p-1}$. This has a solution if $(2, p-1) \mid b$.

If $2 \mid b$, then $a^{\frac{p-1}{2}} = g^{\frac{b}{2}(p-1)} \equiv 1 \pmod{p}$ by Fermat's little theorem.

Conversely, if $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$, then $g^{\frac{b}{2}(p-1)} \equiv 1 \pmod{p}$ and so $p - 1 \mid \frac{b}{2}(p-1)$ which implies $2 \mid b$. □

**Definition VIII.0.5** (Legendre symbol)**.** Let $p$ be an odd prime. The *Legendre symbol* is defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a residue modulo } p \\ -1 & \text{if } a \text{ is a non-residue modulo } p \\ 0 & \text{if } p \mid a \end{cases}$$

**Proposition VIII.0.6.**

*(i)* $a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) \pmod{p}$.

*(ii)* $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

*(iii) If* $a \equiv b \pmod{p}$*, then* $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.

*Proof.* If $p \mid a$ or $p \mid b$, then all of these are trivial. Thus assume that $p \nmid ab$. For (i), we have by Fermat's little theorem that

$$\begin{aligned} a^{p-1} &\equiv 1 \pmod p \\ (a^{(p-1)/2} - 1)(a^{(p-1)/2} + 1) &\equiv 0 \pmod p \end{aligned}$$

but since $\mathbb{Z}/p\mathbb{Z}$ is a field, one of these is zero. So $a^{(p-1)/2} \equiv \pm 1 \pmod p$. By proposition VIII.0.2, $a^{(p-1)/2} \equiv 1 \pmod p$ iff $a$ is a residue modulo $p$.

For (ii), we have using (i) that

$$(ab)^{(p-1)/2} \equiv \left(\frac{ab}{p}\right) \pmod p$$

and

$$a^{(p-1)/2} b^{(p-1)/2} \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod p$$

so $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

(iii) is obvious from the definition. □

**Lecture 29.**
Wed 20 Mar '24

*Remark.* By (i), $\left(\frac{1}{p}\right) = 1$.

**Corollary VIII.0.7.** *The number of quadratic residues modulo $p$ is equal to the number of quadratic non-residues modulo $p$.*

*Proof.* □

**Corollary VIII.0.8.** *The multiplication table of residues modulo $p$ is as follows:*

| $\cdot$ | *Residue* | *Non-residue* |
|---|---|---|
| *Residue* | *Residue* | *Non-residue* |
| *Non-residue* | *Non-residue* | *Residue* |

*Proof.* Follows from definition and (ii). □

**Corollary VIII.0.9.** $(-1)^{\frac{p-1}{2}} = \left(\frac{-1}{p}\right)$.

*Proof.* Set $a = -1$ in (i). □

**Exercise VIII.0.10.** *Use this result to show that there are infinitely many primes congruent to* $1 \pmod 4$.

**Definition VIII.0.11** (Least residues)**.** Let $p$ be a prime. The set

$$S = \left\{ -\frac{p-1}{2}, -\frac{p-3}{2}, \ldots, -1, 1, \ldots, \frac{p-3}{2}, \frac{p-1}{2} \right\}$$

is called the set of *least residues modulo* $p$. These together with 0 form a complete set of residues modulo $p$.

*Example.* For $p = 7$, $S = \{-3, -2, -1, 1, 2, 3\}$.

**Definition VIII.0.12.** Let $p$ be a prime that does not divide $a$. We define

$$\mu(p, a) = \text{number of negative least residues modulo } p \text{ of the integers} \left\{ a, 2a, \ldots, \frac{p-1}{2} a \right\}.$$

**Theorem VIII.0.13** (Gauss' lemma)**.**

$$\left( \frac{a}{p} \right) = (-1)^{\mu(p,a)}$$

*for $p$ prime and $p \nmid a$.*

*Proof.* For each $1 \leq l \leq \frac{p-1}{2}$, let the least residue of $la$ be $\pm m_l$, $m_l > 0$.

**Claim:** $m_l = m_k$ iff $l = k$.

*Proof.* If $l \neq k$ and $m_l = m_k$, then $la \equiv \pm ka \pmod{p}$.

But $p \nmid a$, so $l \pm k \equiv 0 \pmod{p}$, But $|l \pm k| \leq p - 1$, so $l \not\equiv \pm k \pmod{p}$. Contradiction. □

Thus $\left\{ m_1, \ldots, m_{\frac{p-1}{2}} \right\}$ is the complete set $\left\{ 1, 2, \ldots, \frac{p-1}{2} \right\}$.

Then

$$\prod_{l=1}^{(p-1)/2} la \equiv (-1)^{\mu(p,a)} \prod_{l=1}^{(p-1)/2} m_l \pmod{p}$$

$$a^{\frac{p-1}{2}} \left( \frac{p-1}{2} \right)! \equiv (-1)^{\mu(p,a)} \left( \frac{p-1}{2} \right)!$$

$$a^{\frac{p-1}{2}} \equiv (-1)^{\mu(p,a)}.$$

Using (i), we get the result. □

*Remark.* This is a powerful result, but not useful for computation.

**Proposition VIII.0.14.** *2 is a residue (resp. non-residue) for primes of the form $8k \pm 1$ (resp. $8k \pm 3$). Equivalently, $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.*

*Proof.* Let $p$ be an odd prime and $a = 2$. By Gauss' lemma, look at $\{2, 4, \dots, p-1\}$. Thus, $\mu(p, 2)$ is equal to the number of elements which are more than $\frac{p-1}{2}$.

Define $m \in \mathbb{N}$ as the unique natural number such that $2m \leq \frac{p-1}{2}$ and $2(m+1) > \frac{p-1}{2}$. Then $\mu = \frac{p-1}{2} - m$.

- If $p = 8k+1$, then $\frac{p-1}{2} = 4k$ so $m = 2k$. This gives that $\mu = 4k - 2k = 2k$ is even.
- If $p = 8k+3$, then $\frac{p-1}{2} = 4k+1$ so $m = 2k$. This gives that $\mu = 4k+1-2k = 2k+1$ is odd.
- If $p = 8k+5$, then $\frac{p-1}{2} = 4k+2$ so $m = 2k+1$. This gives $\mu = 4k+2-2k-1 = 2k+1$ is odd.
- If $p = 8k+7$, then $\frac{p-1}{2} = 4k+3$ so $m = 2k+1$. This gives $\mu = 4k+3-2k-1 = 2k+2$ is even.

Thus $\left(\frac{2}{p}\right) = (-1)^{\mu}$ is 1 for $p = 8k \pm 1$ and $-1$ for $p = 8k \pm 3$. □

**Theorem VIII.0.15** (Law of quadratic reciprocity)**.** *Let $p$ and $q$ be distinct odd primes. Then*

*(i)* $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

*(ii)* $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

*(iii)* $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.

**Lemma VIII.0.16.** *Let $p$ be an odd prime and $(a, 2p) = 1$. Then*

$$\left(\frac{a}{p}\right) = (-1)^t, \quad \textit{where } t = \sum_{j=1}^{\frac{p-1}{2}} \left\lfloor \frac{ja}{p} \right\rfloor$$

*Proof.* Note that $a$ is odd. Let $\mu$ be as in Gauss' lemma. Let $-r_1, \dots, -r_\mu$ be negative least residues and $s_1, \dots, s_k$ be positive least residues ($k = \frac{p-1}{2} - \mu$). Then

$$\sum_{j=1}^{\frac{p-1}{2}} ja = \sum_{j=1}^{\frac{p-1}{2}} p \left\lfloor \frac{ja}{p} \right\rfloor + \sum_{j=1}^{k} s_j + \sum_{j=1}^{\mu} (p - r_j).$$

But $\sum_{j=1}^{(p-1)/2} j = \frac{p^2-1}{8}$. Also, from the proof of theorem VIII.0.13, we have $\sum r_j + \sum s_j = \sum_{j=1}^{\frac{p-1}{2}} j$.

Thus we have

$$\frac{p^2-1}{8}a = p\sum_{j=1}^{\frac{p-1}{2}} \left\lfloor \frac{ja}{p} \right\rfloor + \sum s_j + \mu p - \sum r_j \tag{VIII.1}$$

$$\frac{p^2-1}{8} = \sum s_j + \sum r_j \tag{VIII.2}$$

Subtracting equation (VIII.2) from equation (VIII.1) gives

$$\begin{aligned}(a-1)\frac{p^2-1}{8} &= p\sum \left\lfloor \frac{ja}{p} \right\rfloor + p\mu - 2\sum r_j \\ &\equiv \sum \left\lfloor \frac{ja}{p} \right\rfloor + \mu \pmod 2.\end{aligned}$$

□

**Lecture 30.** Fri 22 Mar '24

## VIII.1 Generators & Relations

**Lecture 31.** Mon 25 Mar '24

**Definition VIII.1.1** (Presentation)**.** Let $S$ be an abstract set of generators (and their inverses). Let $R$ be a set of identities satisfied by these generators. Then $G = \langle S \mid R\rangle$ forms a group, where the elements are words in the alphabet $S$ and product is concatenation.

We say that $\langle S \mid R\rangle$ is a *presentation* of $G$.

This is assuming that the relations are consistent.

*Examples.*

- $D_{2n} = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\}$ can be written as
$$D_{2n} = \left\langle r, s \mid r^n = s^2 = 1, rs = sr^{-1}\right\rangle$$
- $$\mathbb{Z}/n\mathbb{Z} = \langle x \mid x^n = 1\rangle$$
- $$S_3 = \left\langle s, t \mid s^2 = t^2 = 1, sts = tst\right\rangle.$$
The last relation can also be written as $(st)^3 = 1$.

- The quaternion group $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ with the relations

$$\begin{aligned} 1x = x1 &= x \quad \forall x, \\ (-1) \pm x = \pm x(-1) &= \mp x \quad \forall x, \\ i^2 = j^2 = k^2 &= -1, \\ ijk &= -1. \end{aligned}$$

Here, when we write "1" we mean the empty word.

**Definition VIII.1.2** (Free group)**.** A group $F$ is *free* if it can be written as

$$F = \langle S \mid \emptyset \rangle \quad \text{for some set } S.$$

**Definition VIII.1.3** (Symmetric group)**.** The *symmetric group* of order $n$ is

$$S_n = \big\{ \sigma \in [n]^{[n]} \mid \sigma \text{ is a bijection} \big\}$$

*Example.* Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 7 & 1 & 2 & 6 & 4 & 5 \end{pmatrix} \in S_8$. In cycle notation, $\sigma = (1374)(285)(6)$. It is useful to omit the 1-cycles, $\sigma = (1374)(285)$. This makes products simple to compute.

If one calls the cycles $\sigma_1 = (1374)$ and $\sigma_2 = (285)$, then $\sigma = \sigma_1 \sigma_2 = \sigma_2 \sigma_1$.

**Exercise VIII.1.4.** *The product of disjoint cycles commutes in $S_n$.*

**Definition VIII.1.5** (Morphisms)**.** Let $(G, *)$ and $(H, \cdot)$ be groups. Then a map $\varphi \colon G \to H$ is called a *homomorphism* if for all $g_1, g_2 \in G$,

$$\varphi(g_1 * g_2) = \varphi(g_1) \cdot \varphi(g_2).$$

If in addition, $\varphi$ is a bijection, then $\varphi$ is called an *isomorphism.*

*Examples.*

- $\varphi \colon g \in G \mapsto g \in G$ is an isomorphism.
- $\varphi \colon g \in G \mapsto g^{-1} \in G$ is *not* a homomorphism in general. In fact, it is an isomorphism if and only if $G$ is abelian.
- $\exp \colon (\mathbb{R}, +) \to (\mathbb{R}_+, \cdot)$ is an isomorphism.
- $\operatorname{sgn} \colon S_n \to (\{\pm 1\}, \cdot)$ is a homomorphism.

- Let $n > m$ and $\varphi\colon S_m \to S_n$ given by adding singletons is a homomorphism. That is, if $\sigma = c_1 \dots c_k \in S_m$, then
$$\varphi(\sigma) = c_1 \dots c_k (m+1) \dots (n).$$
- $\varphi\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ given by $x \mapsto 2x$ is a homomorphism. It is an isomorphism iff $n$ is odd.
- $\varphi\colon \pi \in S_n \mapsto \pi^2 \in S_n$ is *not* a homomorphism (for $n \geq 3$).

*Notation.* The order of $x \in G$ is denoted $|x|$.

**Exercise VIII.1.6.** *Suppose $\varphi\colon G \to H$ is an isomorphism. Then*

*(1) $|G| = |H|$*

*(2) $G$ is abelian iff $H$ is abelian.*

*(3) $\forall x \in G$, $|x| = |\varphi(x)|$.*

**Corollary VIII.1.7.** *$\mathbb{Z}/6\mathbb{Z}$ and $S_3$ are not isomorphic.*

## VIII.2 Group Actions

Lecture 32.
Wed 27 Mar '24

**Definition VIII.2.1** (Group action)**.** A (left) *group action* of a group $G$ on a set $A$ is a map $G \times A \to A$ written $(g, a) \mapsto g \cdot a$, satisfying

(i) $g_1 \cdot (g_2 \cdot a) = (g_1 g_2) \cdot a$ for all $g_1, g_2 \in G$ and $a \in A$.

(ii) $e \cdot a = a$ for all $a \in A$.

If such a map exists, we say that $G$ *acts on* $A$. Equivalently, $A$ has the *symmetries* of $G$.

*Examples.*

- The regular $n$-gon is acted on by $D_{2n}$.
- $(g, a) \mapsto a$ is the *trivial action* of $G$ on $A$.
- Let $V$ be a vector space over a field $F$. Then $(F^*, *)$ is a group, and $F^*$ acts on $V$ by scalar multiplication.

- Let $A$ be a set and $S_A$ denote the group of permutations of $A$ (*i.e.*, bijection from $A$ to $A$). Then $S_A$ acts on $A$ by $\sigma \cdot a = \sigma(a)$.

Suppose $G$ acts on $A$. Then define for all $g \in G$, the map $\sigma_g \colon A \to A$ by

$$\sigma_g(a) = g \cdot a.$$

**Proposition VIII.2.2.** *(i) For all $g \in G$, $\sigma_g$ is a permutation of $A$. That is, $\sigma_g \in S_A$.*

*(ii) The map $\varphi \colon G \to S_A$ given by $\varphi(g) = \sigma_g$ is a group homomorphism (where the group operation on $S_A$ is composition).*

*Proof.*

(i) We will show that $\sigma_{g^{-1}}$ is the (two-sided) inverse of $\sigma_g$.

$$\begin{aligned}(\sigma_{g^{-1}} \circ \sigma_g)(a) &= \sigma_{g^{-1}}(g \cdot a)\\ &= g^{-1} \cdot (g \cdot a)\\ &= (g^{-1}g) \cdot a\\ &= e \cdot a\\ &= a.\end{aligned}$$

Interchange the roles of $g$ and $g^{-1}$ to get

$$(\sigma_g \circ \sigma_{g^{-1}})(a) = a.$$

(ii) Consider

$$\begin{aligned}\varphi(g_1g_2)(a) &= \sigma_{g_1g_2}(a)\\ &= (g_1g_2) \cdot a\\ &= g_1 \cdot (g_2 \cdot a)\\ &= \sigma_{g_1}(\sigma_{g_2}(a))\\ &= \varphi(g_1)(\varphi(g_2)(a))\\ &= (\varphi(g_1) \circ \varphi(g_2))(a).\end{aligned}$$

Since this holds for all $a \in A$,

$$\varphi(g_1g_2) = \varphi(g_1) \circ \varphi(g_2).$$

□

**Definition VIII.2.3** (Permutation representation)**.** Let $G$ act on $A$ by $g \cdot a$. The map $\varphi\colon G \to S_A$ given by $\varphi(g) = \sigma_g$, where $\sigma_g(a) = g \cdot a$, is called the *permutation representation* of $G$.

**Definition VIII.2.4** (Faith)**.** A group action is called *faithful* if every $g \in G$ gives a different permutation of $A$, *i.e.*, the permutation representation is injective.

**Definition VIII.2.5** (Kernel)**.** The *kernel* of a group action of $G$ on $A$ is the set

$$\{g \in G \mid g \cdot a = a \text{ for all } a \in A\}.$$

*Example.* The trivial action is unfaithful (except for the trivial group) and the kernel is the group itself.

**Proposition VIII.2.6.** *Any action is faithful if and only if its kernel is* $\{e\}$.

*Proof.* We know that $\varphi(e) = \sigma_e = \mathrm{id}_A$. If the action is faithful, then there is no other $g \in G$ such that $g \cdot a = a \ \forall a \in A$. Thus, the kernel is $\{e\}$.

Conversely, suppose the kernel is $\{e\}$. Let $g_1, g_2$ be such that $\sigma_{g_1} = \sigma_{g_2}$. Then $\sigma_{g_1 g_2^{-1}} = \mathrm{id}_A$, so $g_1 g_2^{-1} = e$, and $g_1 = g_2$. □

*Example.* $G$ acts on itself by left multiplication. This is faithful, because $g_1 e = g_2 e \iff g_1 = g_2$.

**Definition VIII.2.7** (Subgroup)**.** Let $G$ be a group. A non-empty subset $H \subseteq G$ is called a *subgroup*, denoted $H \leq G$, if $H$ is closed under products and inverses.

**Proposition VIII.2.8.** *Any subgroup of a group is a group.*

*Proof.* □

*Examples.*

- Every group $G$ has two subgroups, $\{e\}$ and $G$. $\{e\}$ is called the *trivial subgroup.* If $H \leq G$ and $H \neq G$, then $H$ is called a *proper subgroup.*
- $\{1, r, \ldots, r^{n-1}\} < D_{2n}$.
- $(3\mathbb{Z}, +) < (\mathbb{Z}, +)$.
- $(\mathbb{Q}^*, \cdot) \not\leq (\mathbb{R}, +)$.

**Proposition VIII.2.9.** *A non-empty subset $H \subseteq G$ is a subgroup if and only if for every $x, y \in H$, $xy^{-1} \in H$.*

*Proof.* Suppose $H$ is a subgroup. Then $x, y \in H \implies y^{-1} \in H$ by closure under inverses, and $xy^{-1} \in H$ by closure under products.

Suppose $H$ satisfies the condition. Let $h \in H$. Then $hh^{-1} = e \in H$. Now taking $x = e$ and $y = h$, we have $h^{-1} = eh^{-1} \in H$. Thus we have closure under inverses.

Let $h_1, h_2 \in H$. Then $h_2^{-1} \in H$, so $h_1 h_2 = h_1(h_2^{-1})^{-1} \in H$. □

**Exercise VIII.2.10.** *A finite non-empty subset $H$ of a group $G$ is a subgroup iff for all $x, y \in H$, $xy \in H$.*

*Proof.* If $H$ is a subgroup, this is by definition.

Suppose $H$ satisfies the condition. Let $h \in H$. $H$ is closed under products, so $\{h, h^2, \ldots\} \subseteq H$. By the pigeonhole principle, $e$ and $h^{-1}$ are in this set. □

**Lecture 33.** Fri 05 Apr '24

**Lecture 34.** Fri 05 Apr '24

**Lecture 35.** Fri 05 Apr '24

**Proposition VIII.2.11.** *Let $N \leq G$. Then the set of left cosets partitions $G$. Moreover, $uN = vN$ iff $u^{-1}v \in N$, and $uN = vN$ iff $u$ and $v$ represent the coset.*

*Proof.* Since $N \leq G$, $e \in N$ and so $g = g \cdot 1 \in gN$. Thus $G = \bigcup_{g \in G} gN$.

Suppose $u, v \in G$ such that $uN \cap vN \neq \varnothing$. Let $x = um = vn$ be an element of the intersection. Then $u = vnm^{-1}$, so $uN \subseteq vN$. Similarly, $vN \subseteq uN$, so that $uN = vN$.

Thus the set of left cosets $\{gN \mid g \in G\}$ is a partition of $G$.

Now $uN = vN \iff um = vn \iff u^{-1}v = mn^{-1} \in N$.

Finally, $v \in uN$ means that $v$ is a representative for $uN$. So $vN = uN$. Thus they both represent the same coset. □

**Definition VIII.2.12** (Normal)**.** The element $gng^{-1}$ is called the *conjugate* of $n$ by $g$, and $gNg^{-1} = \{gng^{-1} \mid n \in N\}$ is called the conjugate of $N$ by $g$.

The element $g$ is said to *normalize* $N$ if $gNg^{-1} = N$. A subgroup $N$ is said to be *normal* if every $g \in G$ normalizes $N$. We write $N \trianglelefteq G$.

**Proposition VIII.2.13.** *Let $G$ be a group and $N \leq G$. Then*

*(i) The operation $(uN) \cdot (vN) = (uv)N$ is well-defined iff $N$ is normal in $G$.*

*(ii) If so, the set of left cosets form a group with the above product, with identity $eN$ and inverse $(gN)^{-1} = g^{-1}N$.*

*Proof.*

(i) Suppose the product is well-defined. That is, for any $u' \in uN$ and $v' \in vN$, we have $u'v' \in uvN$.

Let $g \in G$ and $n \in N$. Taking the product of $eN$ with $gN$, we have $ng \in gN \implies g^{-1}ng \in N$.

Now suppose $N$ is normal in $G$. Then $g^{-1}ng \in N$ for all $g \in G$ and $n \in N$.

Let $u' = um$, $v' = vn$, $m, n \in N$. Then

$$u'v' = umvn = uv(v^{-1}mvm) = uv(m'm) \in uvN$$

since $v^{-1}mv = m' \in N$.

(ii) Associativity follows from the associativity of $G$. Identity is borrowed from $G$. Inverse is borrowed from $G$. □

*Example.* If $G$ is abelian, then every subgroup is normal. So $G/N$ is the quotient group for any $N \leq G$.

**Theorem VIII.2.14.** *Let $N \leq G$. Then the following are equivalent.*

*(i)* $N \trianglelefteq G$.

*(ii)* $N_G(N) = G$*, where* $N_G(N)$ *is the set of normalizers of* $N$ *in* $G$.

*(iii)* $gN = Ng$ *for all* $g \in G$.

*(iv) The multiplication of left cosets makes* $G/N$ *a group.*

*(v)* $gNg^{-1} \subseteq N$ *for all* $g \in G$.

**Proposition VIII.2.15.** $N \leq G$ *is normal iff it is the kernel of some homomorphism.*

*Proof.* We have shown that left and right cosets are equal for kernels of homomorphisms. This gives that kernels are normal subgroups.

Conversely, suppose $N \trianglelefteq G$. Let $H = G/N$. Define $\pi\colon G \to H$ by $\pi(g) = gN$. Then $\pi$ is a homomorphism.

What is the kernel of $\pi$?

$$\begin{aligned} \ker \pi &= \{g \in G \mid \pi(g) = eN\} \\ &= \{g \in G \mid gN = N\} \\ &= \{g \in G \mid g \in N\} = N. \end{aligned}$$ □

**Definition VIII.2.16.** Let $N \leq G$. The homomorphism $g \mapsto gN$ is the *natural projection* of $G$ onto $G/N$.