Combinatorics (Foundations)

Gudfit

Contents

0
Foundations

In this chapter, we re-establish the algebraic and logical foundations of counting, beginning with the language of sets and extending to the structure of product spaces and power sets.

0.1 The Language of Sets

We begin by formalising the operations on sets and their algebraic properties. We adopt an intuitive definition of a **set** as a collection of distinct objects, denoted as elements. If x is an element of a set E, we write $x \in E$. The set containing no elements is the **empty set**, denoted by $\varnothing$.

Let Ω denote a universal set. For subsets $A, B \subseteq \Omega$, we define the fundamental operations.

I hope you've read my set theory notes as it gives a better introduction to the foundations needed.

> **Definition 0.1.** *Set Operations.*
> Let A and B be sets.
>
> (i) The **union** $A \cup B$ is the set of elements in A or B:
>
> $$x \in A \cup B \iff (x \in A) \vee (x \in B).$$
>
> (ii) The **intersection** $A \cap B$ is the set of elements in both A and B:
>
> $$x \in A \cap B \iff (x \in A) \wedge (x \in B).$$
>
> (iii) The **difference** $A \setminus B$ consists of elements in A but not in B:
>
> $$x \in A \setminus B \iff (x \in A) \wedge (x \notin B).$$
>
> (iv) The **symmetric difference** $A \Delta B$ contains elements in exactly one of the sets:
>
> $$x \in A \Delta B \iff (x \in A \cup B) \wedge (x \notin A \cap B).$$
>
> (v) The **complement** A^c (or $\bar{A}$) is the set of elements in Ω not in A:
>
> $$x \in A^c \iff x \in \Omega \setminus A.$$

定義

To manipulate these structures algebraically, we introduce the characteristic function (or indicator function).

> **Definition 0.2.** ***Characteristic Function.***
> The **characteristic function** of a subset $A \subseteq \Omega$ is the function $1_A : \Omega \to \{0,1\}$ defined by:
>
> $$1_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$
>
> 定義

The correspondence between subsets and their characteristic functions allows us to translate set-theoretic logical operations into arithmetic operations.

> **Proposition 0.1.** ***Calculus of Characteristic Functions.***
> Let $A, B \subseteq \Omega$. For all $x \in \Omega$:
> (i) **Intersection:** $1_{A\cap B}(x) = 1_A(x) \cdot 1_B(x)$.
> (ii) **Complement:** $1_{A^c}(x) = 1 - 1_A(x)$.
> (iii) **Union:** $1_{A\cup B}(x) = 1_A(x) + 1_B(x) - 1_A(x)1_B(x)$.
> (iv) **Difference:** $1_{A\setminus B}(x) = 1_A(x)(1 - 1_B(x))$.
> (v) **Symmetric Difference:** $1_{A\Delta B}(x) = 1_A(x) + 1_B(x) - 2 \cdot 1_A(x)1_B(x)$.
>
> 命題

> *Note*
>
> $A\Delta B = (A \cup B) \setminus (A \cap B)$, and pointwise $1_{A\Delta B} = |1_A - 1_B|$.

> *Proof*
>
> These identities are verified pointwise. We demonstrate the union and symmetric difference cases; the others follow similarly.
>
> - For the union: If $x \notin A \cup B$, then $1_A(x) = 0$ and $1_B(x) = 0$, so the RHS is 0. If $x \in A \setminus B$, RHS is $1 + 0 - 0 = 1$. If $x \in B \setminus A$, RHS is $0 + 1 - 0 = 1$. If $x \in A \cap B$, RHS is $1 + 1 - 1 = 1$. Thus the identity holds.
> - For the symmetric difference: If $x \in A\Delta B$, then exactly one of $1_A(x), 1_B(x)$ equals 1, so the RHS is 1. If $x \in A \cap B$ or $x \notin A \cup B$, then the RHS is 0. Thus the identity holds.
>
> ■

> **Example 0.1.** Algebraic Proof of Set Identities. We use characteristic functions to prove the identity $(A \setminus B) \setminus C = A \setminus (B \cup C)$.

$$\begin{aligned}1_{(A\setminus B)\setminus C} &= 1_{A\setminus B}(1-1_C)\\ &= 1_A(1-1_B)(1-1_C)\\ &= 1_A(1-1_B-1_C+1_B1_C)\\ &= 1_A[1-(1_B+1_C-1_B1_C)]\\ &= 1_A(1-1_{B\cup C})\\ &= 1_{A\setminus(B\cup C)}.\end{aligned}$$

Since the characteristic functions are identical, the sets are equal.

範例

Cartesian Products and Power Sets

We now define the construction of sets from components, essential for defining counting spaces.

Definition 0.3. ***Cartesian Product.***
The **Cartesian product** of two sets A and B, denoted $A \times B$, is the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$:

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

More generally, for a sequence of sets $A_1, \dots, A_n$, the product is the set of n-tuples:

$$\prod_{i=1}^{n} A_i = A_1 \times \cdots \times A_n = \{(a_1, \dots, a_n) : a_i \in A_i \text{ for all } i\}.$$

If $A_i = A$ for all i, we denote the product as A^n.

定義

Unlike sets, where order and repetition are irrelevant ($\{a, b\} = \{b, a\}$), in ordered pairs the order is strict: $(a, b) = (b, a)$ if and only if $a = b$.

Example 0.2. Combinatorial Configuration. Consider a simplified model of a "menu" where one chooses a main dish from $M = \{m_1, m_2\}$ and a side from $S = \{s_1, s_2, s_3\}$. The set of all possible meals is the Cartesian product $M \times S$.

$$M \times S = \{(m_1, s_1), (m_1, s_2), (m_1, s_3), (m_2, s_1), (m_2, s_2), (m_2, s_3)\}.$$

The size of the set is $|M \times S| = |M| \cdot |S| = 2 \cdot 3 = 6$. This multiplicative principle underlies much of combinatorics.

範例

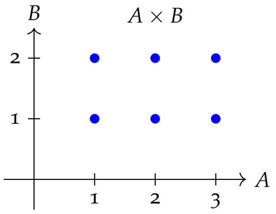

Figure 1: Visualisation of the product $\{1, 2, 3\} \times \{1, 2\}$. The set consists of $3 \times 2 = 6$ distinct points.

Definition 0.4. ***Power Set.***
The **power set** of a set E, denoted $\mathcal{P}(E)$, is the set of all subsets of E:

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

定義

Example 0.3. Power Set of a Pair. Let $E = \{1, 2\}$. The subsets of E are the empty set, the singleton sets, and E itself.

$$\mathcal{P}(E) = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}.$$

Observe that $|\mathcal{P}(E)| = 4$.

範例

Proposition 0.2. ***Monotonicity of the Power Set.***
Let A and B be sets. If $A \subseteq B$, then $\mathcal{P}(A) \subseteq \mathcal{P}(B)$.

命題

Proof

Let $X \in \mathcal{P}(A)$. By definition, this means X is a subset of A ($X \subseteq A$). Since subset inclusion is transitive, if $X \subseteq A$ and $A \subseteq B$, then $X \subseteq B$. Thus, X is a subset of B, which implies $X \in \mathcal{P}(B)$.

■

Example 0.4. Intersection of Power Sets. Consider the relationship between power sets and intersection. We claim that for any sets A and B:

$$\mathcal{P}(A) \cap \mathcal{P}(B) = \mathcal{P}(A \cap B).$$

範例

Proof

Let $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$. Then $X \in \mathcal{P}(A)$ and $X \in \mathcal{P}(B)$, meaning $X \subseteq A$ and $X \subseteq B$. By definition of intersection, $X \subseteq A \cap B$, so $X \in \mathcal{P}(A \cap B)$. Conversely, if $X \in \mathcal{P}(A \cap B)$, then $X \subseteq A \cap B$. This implies $X \subseteq A$ and $X \subseteq B$, so $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$.

■

0.2 *Mappings and Functions*

We move from static sets to dynamic relationships between them. A mapping assigns elements of one set to another in a deterministic fashion.

Definition 0.5. ***Mapping.***
Let E and F be sets. A **mapping** (or function) $f : E \to F$ is a subset $f \subseteq E \times F$ such that for every $x \in E$, there exists a unique $y \in F$

with $(x, y) \in f$. We write $y = f(x)$ or $x \mapsto y$.

- E is the **source** (or domain).
- F is the **target** (or codomain).
- y is the **image** of x.
- x is an **antecedent** (or pre-image) of y.

定義

The behaviour of a function regarding the uniqueness and existence of antecedents classifies it into three fundamental types.

Definition 0.6. ***Injectivity, Surjectivity, Bijectivity.***
Let $f : E \to F$ be a mapping.

1. **Injective** (One-to-one): Distinct elements map to distinct images.

$$\forall x, y \in E, \quad f(x) = f(y) \implies x = y.$$

2. **Surjective** (Onto): Every element in the target has at least one antecedent.

$$\forall y \in F, \quad \exists x \in E \text{ such that } y = f(x).$$

Equivalently, the image set $f(E) = \{f(x) : x \in E\}$ equals F.

3. **Bijective**: The mapping is both injective and surjective.

$$\forall y \in F, \quad \exists! x \in E \text{ such that } y = f(x).$$

定義

Example 0.5. Cardinality via Mappings. Let $E = \mathcal{P}(\{1, 2\})$ be the power set of $\{1, 2\}$, and let $F = \{0, 1, 2\}$. Define $f : E \to F$ by $f(A) = |A|$ (the cardinality of the subset).

- The elements of E are $\varnothing, \{1\}, \{2\}, \{1, 2\}$.
- The mappings are: $f(\varnothing) = 0$, $f(\{1\}) = 1$, $f(\{2\}) = 1$, $f(\{1, 2\}) = 2$.

This function is **surjective** because 0, 1, and 2 all appear as images. It is **not injective** because $f(\{1\}) = f(\{2\})$ but $\{1\} \neq \{2\}$.

範例

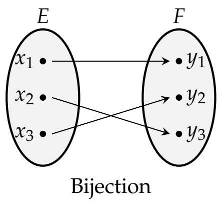

Figure 2: A bijective map requires $|E| = |F|$. Every y has exactly one incoming arrow.

When mappings are composable, properties of the individual functions transfer to the composite.

Lemma 0.1. ***Composition Properties.***
Let $f : E \to F$ and $g : F \to G$ be mappings. Let $g \circ f : E \to G$ be the composite defined by $(g \circ f)(x) = g(f(x))$.

1. If f and g are injective, $g \circ f$ is injective.
2. If f and g are surjective, $g \circ f$ is surjective.
3. If f and g are bijective, $g \circ f$ is bijective.

引理

Proof

We verify the first statement. Suppose $(g \circ f)(x) = (g \circ f)(y)$. Then $g(f(x)) = g(f(y))$. Since g is injective, $f(x) = f(y)$. Since f is injective, $x = y$. The other proofs are analogous.

■

Note

If f is bijective, there exists a unique inverse mapping $f^{-1} : F \to E$ such that $f^{-1} \circ f = \mathrm{id}_E$ and $f \circ f^{-1} = \mathrm{id}_F$.

Binary Relations

A **binary relation** $\mathcal{R}$ on a set E is a subset of $E \times E$. If $(x, y) \in \mathcal{R}$, we write $x\mathcal{R}y$. The structure of E is often understood by the properties of relations defined upon it.

Definition 0.7. ***Properties of Relations.***
A relation $\mathcal{R}$ on E is:

- **Reflexive:** $\forall x \in E, x\mathcal{R}x$.
- **Symmetric:** $\forall x, y \in E, x\mathcal{R}y \implies y\mathcal{R}x$.
- **Antisymmetric:** $\forall x, y \in E, (x\mathcal{R}y \wedge y\mathcal{R}x) \implies x = y$.
- **Transitive:** $\forall x, y, z \in E, (x\mathcal{R}y \wedge y\mathcal{R}z) \implies x\mathcal{R}z$.

定義

These properties combine to form two crucial structures: equivalences and orders.

Equivalence Relations and Partitions

An **equivalence relation** is a relation that is reflexive, symmetric, and transitive. It generalises the notion of equality by treating different objects as "the same" under a specific criterion.

Definition 0.8. ***Equivalence Class.***
Let $\mathcal{R}$ be an equivalence relation on E. The **equivalence class** of x, denoted $\mathcal{R}[x]$ (or $[x]$), is the set of all elements related to x:

$$\mathcal{R}[x] = \{y \in E : x\mathcal{R}y\}.$$

The **quotient set** $E/\mathcal{R}$ is the set of all equivalence classes.

定義

Example 0.6. Rational Number Construction. Consider the set $E = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. We define a relation $\sim$ by:

$$(a, b) \sim (c, d) \iff ad = bc.$$

This is an equivalence relation. The class of $(1, 2)$ is the set

$\{(1,2),(2,4),(3,6),\dots\}$, which represents the rational number $1/2$. The set of rationals $\mathbb{Q}$ is formally the quotient set $E/\sim$.

範例

There is a bijection between equivalence relations and partitions. Recall that a partition is a collection of disjoint non-empty subsets covering E.

Proposition 0.3. ***The Fundamental Theorem of Equivalence Relations.***

Let $\mathcal{R}$ be an equivalence relation on E. Then the quotient set $E/\mathcal{R}$ forms a partition of E. Conversely, any partition of E induces a unique equivalence relation.

命題

Proof

Since $\mathcal{R}$ is reflexive, $x \in \mathcal{R}[x]$, so the union of classes covers E, and no class is empty. We must show the classes are pairwise disjoint. Suppose $\mathcal{R}[x] \cap \mathcal{R}[y] \neq \varnothing$. Let z be an element in the intersection. Then $x\mathcal{R}z$ and $y\mathcal{R}z$. By symmetry $z\mathcal{R}y$, and by transitivity $x\mathcal{R}y$. If $w \in \mathcal{R}[y]$, then $y\mathcal{R}w$, so $x\mathcal{R}w$ (transitivity via $x\mathcal{R}y$), implying $w \in \mathcal{R}[x]$. Thus $\mathcal{R}[y] \subseteq \mathcal{R}[x]$. By symmetry, $\mathcal{R}[x] = \mathcal{R}[y]$. Thus, two classes are either disjoint or identical.
Conversely, let $\{E_i\}_{i \in I}$ be a partition of E and define $x\mathcal{R}y$ if x and y lie in the same block E_i. Each x lies in some E_i, so $x\mathcal{R}x$. If $x\mathcal{R}y$, then $y\mathcal{R}x$. If $x\mathcal{R}y$ and $y\mathcal{R}z$, then x, y, z lie in the same block, so $x\mathcal{R}z$. Hence $\mathcal{R}$ is an equivalence relation and its classes are exactly the blocks. Uniqueness follows because any relation with the same partition relates precisely the pairs lying in the same block.

■

Order Relations

An **order relation** (or partial order), denoted $\leq$, is a relation that is reflexive, antisymmetric, and transitive. The pair $(E, \leq)$ is called a **partially ordered set** (or poset).

- If for all $x, y \in E$, either $x \leq y$ or $y \leq x$, the order is **total**.
- We define $x < y$ if $x \leq y$ and $x \neq y$.

Example 0.7. Divisibility Lattice. Let $E = \{1,2,3,4,6,12\}$. Define $a \leq b$ if a divides b (written $a \mid b$). Reflexivity ($a \mid a$), antisymmetry ($a \mid b$ and $b \mid a \implies a = b$ for positive integers), and transitivity hold. This is not a total order: $2 \nmid 3$ and $3 \nmid 2$, so 2 and 3 are incomparable.

範例

In a poset, we distinguish between "greatest" and "maximal".

- An element $M \in E$ is the **greatest element** if $x \leq M$ for all $x \in E$.
- An element $m \in E$ is **maximal** if there is no $y \in E$ such that $m < y$.

In the example above, 12 is the greatest element. Consider $E \setminus \{12\} = \{1, 2, 3, 4, 6\}$. Here, 4 and 6 are both maximal, but neither is the greatest.

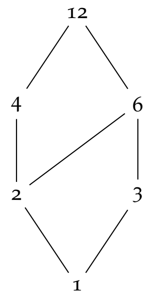

Figure 3: Hasse diagram of the divisors of 12 ordered by divisibility. Lines indicate the covering relation.

The Natural Numbers

We denote the set of natural numbers by $\mathbb{N} = \{0, 1, 2, \ldots\}$ and $\mathbb{N}^* = \mathbb{N} \setminus \{0\} = \{1, 2, 3, \ldots\}$. For $a, b \in \mathbb{N}$, the discrete interval is defined as:

$$[a, b] = \{n \in \mathbb{N} : a \leq n \text{ and } n \leq b\}.$$

If $a > b$, $[a, b] = \varnothing$.

We accept the following axiom as the foundation of arithmetic proofs.

> **Axiom 1.** Well-Ordering Principle Every non-empty subset of $\mathbb{N}$ possesses a smallest element.
>
> 公理

This principle underpins the method of mathematical induction.

> **Theorem 0.1.** ***Principle of Mathematical Induction.***
> Let $\mathcal{P}(n)$ be a proposition depending on $n \in \mathbb{N}$. If:
> 1. $\mathcal{P}(0)$ is true (Base Case), and
> 2. $\forall n \in \mathbb{N}, \mathcal{P}(n) \implies \mathcal{P}(n+1)$ (Inductive Step),
>
> then $\mathcal{P}(n)$ is true for all $n \in \mathbb{N}$.
>
> 定理

> *Proof*
>
> Let $F = \{n \in \mathbb{N} : \mathcal{P}(n) \text{ is false}\}$. We wish to show $F = \varnothing$. Suppose for contradiction that $F \neq \varnothing$. By the Well-Ordering Principle, F has a smallest element n_0. Since $\mathcal{P}(0)$ is true, $n_0 \neq 0$, so $n_0 - 1 \in \mathbb{N}$. Since n_0 is the *smallest* counterexample, $\mathcal{P}(n_0 - 1)$ must be true. By the inductive step, $\mathcal{P}(n_0 - 1) \implies \mathcal{P}(n_0)$. Thus $\mathcal{P}(n_0)$ is true, contradicting $n_0 \in F$. Therefore, $F = \varnothing$.
>
> ■

0.3 *Permutations and Cardinality of Finite Sets*

A permutation is, fundamentally, a rearrangement of a set's elements.

> **Definition 0.9.** ***Permutation.***
> Let S be a finite set. A **permutation** of S is a bijection $\sigma : S \to S$. When

> $S = \{1, 2, \dots, n\}$, the set of all permutations of S is denoted by S_n, the **symmetric group** of degree n.
>
> 定義

Geometrically, one may visualise a permutation $\sigma \in S_n$ as placing n distinct items into n fixed positions. If $\sigma(j)$ denotes the item at position j, the permutation is uniquely identified by the sequence of values $\sigma(1), \sigma(2), \dots, \sigma(n)$. This is the **one-line notation**:

$$\sigma = [\sigma(1), \sigma(2), \dots, \sigma(n)].$$

For instance, if $n = 3$ and σ maps $1 \mapsto 3$, $2 \mapsto 2$, and $3 \mapsto 1$, we write $\sigma = [3, 2, 1]$.
The **composition** of two permutations $\sigma, \rho \in S_n$ is the function composition $\sigma \circ \rho$, defined by $(\sigma \circ \rho)(j) = \sigma(\rho(j))$. This corresponds to successively applying two rearrangements.

Cycle Decomposition

While one-line notation captures the static arrangement, **cycle notation** reveals the dynamical structure of the permutation under iteration.

> **Definition 0.10.** ***Cycle Decomposition.***
> A **cycle** of length k is a list of distinct elements $(x_1, \dots, x_k)$ such that $\sigma(x_i) = x_{i+1}$ for $1 \leq i < k$ and $\sigma(x_k) = x_1$. This cycle denotes the permutation that cyclically permutes these x_i and fixes all other elements. Every permutation can be decomposed uniquely into disjoint cycles, up to reordering the cycles and cyclic rotation within each cycle.
>
> 定義

> **Example 0.8.** Decomposition Example. Consider $\sigma \in S_7$ given in one-line notation by $[7, 3, 5, 6, 2, 4, 1]$.
> We compute the images: $\sigma(1) = 7$, $\sigma(2) = 3$, $\sigma(3) = 5$, $\sigma(4) = 6$, $\sigma(5) = 2$, $\sigma(6) = 4$, $\sigma(7) = 1$. Tracing the orbits:
>
> - $1 \mapsto 7 \mapsto 1$: This forms the cycle $(1, 7)$.
> - $2 \mapsto 3 \mapsto 5 \mapsto 2$: This forms the cycle $(2, 3, 5)$.
> - $4 \mapsto 6 \mapsto 4$: This forms the cycle $(4, 6)$.
>
> Thus, the cycle decomposition is $\sigma = (2, 3, 5)(1, 7)(4, 6)$. (See figure 4).
>
> 範例

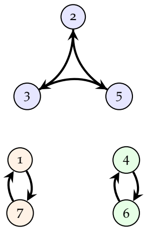

Figure 4: Cycle decomposition for $\sigma = (2, 3, 5)(1, 7)(4, 6) \in S_7$. The element 4 maps to 6, which maps back to 4.

This decomposition allows us to classify permutations by their structure:

- An **involution** is a permutation σ such that $\sigma^2 = \text{id}$. In cycle notation, this means every cycle has length 1 or 2.

- A permutation is **fixed point free** (or a derangement) if $\sigma(x) \neq x$ for all x. In cycle notation, no cycles of length 1 appear.

0.4 *Cardinality of Finite Sets*

A set E is **finite** if it is empty or if there exists a bijection between E and a discrete interval $[1, n] = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}^*$. To use n as a measure of size, we must ensure it is unique.

> *Remark.*
>
> Here $[1, n]$ is the discrete interval in $\mathbb{N}$ from *The Natural Numbers*, a finite set of consecutive integers. Thus $[1, n] = \{1, 2, \ldots, n\}$, not the real interval $[1, n] \subset \mathbb{R}$.

> **Theorem 0.2.** ***Uniqueness of Cardinality.***
> Let $n, m \in \mathbb{N}^*$. There exists an injection from $[1, n]$ into $[1, m]$ if and only if $n \leq m$. Consequently, if there is a bijection between $[1, n]$ and $[1, m]$, then $n = m$.
>
> 定理

The direct implication is trivial: if $n \leq m$, the inclusion map $i \mapsto i$ is injective. For the converse, we proceed by induction on n.

> ***Base Case ($n = 1$).***
>
> Since $m \geq 1$, the inequality $1 \leq m$ holds immediately.
>
> 証明終

> ***Inductive Step.***
>
> Assume the statement holds for some $n \geq 1$. Let $f : [1, n+1] \to [1, m]$ be an injection. Since the domain has at least two elements, $m \geq 2$, so $m - 1 \in \mathbb{N}^*$. We construct an injection $g : [1, n] \to [1, m-1]$.
>
> - If $f(n+1) = m$, let g be the restriction $f|_{[1,n]}$. Its image lies in $[1, m] \setminus \{m\} = [1, m-1]$.
>
> - If $f(n+1) = k \neq m$, define the transposition $\tau = (k, m) \in S_m$ which swaps k and m. Then $(\tau \circ f)(n+1) = m$. The function $g = (\tau \circ f)|_{[1,n]}$ maps $[1, n]$ injectively into $[1, m-1]$.
>
> By the inductive hypothesis, $n \leq m - 1$, which implies $n + 1 \leq m$. For the second part, if a bijection exists, we have injections in both directions. Thus $n \leq m$ and $m \leq n$, implying $n = m$.
>
> 証明終

> **Definition 0.11.** *Cardinality.*
> If E is a finite set in bijection with $[1, n]$, we define the **cardinality** of E as $|E| = n$. If $E = \varnothing$, we set $|E| = 0$.

定義

The uniqueness theorem leads to a powerful combinatorial tool: to count a set, we need only match it perfectly with a set we already know how to count.

> **Proposition 0.4.** ***Bijection Principle.***
> If two finite sets E and F are in bijection, then $|E| = |F|$.
>
> 命題

> *Proof*
>
> Let $f : E \to F$ be a bijection. Let $n = |F|$, so there exists a bijection $g : F \to [1, n]$. The composite $g \circ f : E \to [1, n]$ is a bijection (*lemma* 0.1). By definition, $|E| = n = |F|$.
>
> ■

Bijective proof are among the most satisfying in combinatorics, but in many cases are hard to come by, and subtle when they are found.

The Pigeonhole Principle

The contrapositive of the injection theorem yields the famous Pigeonhole Principle.

> **Corollary 0.1.** *Pigeonhole Principle.* Let E and F be finite sets.
>
> 1. An injection $E \to F$ exists if and only if $|E| \leq |F|$.
> 2. If $|E| > |F|$, then for any map $f : E \to F$, there exists some $y \in F$ such that $|f^{-1}(\{y\})| \geq 2$.
>
> 推論

> *Proof*
>
> Let $n = |E|$ and $m = |F|$. If an injection $f : E \to F$ exists, then the composite with the bijections to intervals gives an injection $[1, n] \to [1, m]$, which implies $n \leq m$. If $n > m$, no injection exists. Thus, any function must map two distinct inputs to the same output.
>
> ■

This principle provides non-constructive existence proofs for elements with specific properties.

> **Example 0.9.** Coprime Pairs. Let $A \subseteq [1, 2n]$ be a subset of size $n + 1$. We claim A contains two integers that are relatively prime.
>
> 範例

> *Proof*
>
> Partition the set $[1, 2n]$ into n "pigeonholes" defined by pairs of consecutive integers: $H_i = \{2i - 1, 2i\}$ for $i = 1, \ldots, n$. Since $|A| = n + 1$ and there are only n such sets, by *corollary* 0.1, one set H_k must contain two elements from A. The only way to contain two elements

from H_k is to contain both $2k - 1$ and $2k$. Consecutive integers are always coprime ($\gcd(x, x+1) = 1$). Thus, the pair exists.

■

Example 0.10. Divisibility in Subsets. Let $A \subseteq [1, 2n]$ with $|A| = n+1$. We show that A contains distinct integers a, b such that $a \mid b$.

範例

Proof

Every positive integer x can be uniquely written as $x = 2^k \cdot m$, where m is the "odd part" of x. For elements in $[1, 2n]$, the odd part m must be in the set of odd numbers $\{1, 3, 5, \ldots, 2n - 1\}$. There are exactly n such odd numbers. Define the pigeonholes as these n odd values. Since we select $n+1$ numbers, two distinct integers $x, y \in A$ must share the same odd part m. Let $x = 2^u m$ and $y = 2^v m$. If $u < v$, then $x \mid y$. If $v < u$, then $y \mid x$.

■

Example 0.11. The Friends Problem. In any gathering of $n \geq 2$ people, there are at least two people who have the same number of friends present at the gathering (assuming friendship is symmetric). Let $P = \{p_1, \ldots, p_n\}$ be the set of people. Let $f(p_i)$ be the number of friends of person p_i. The possible values for $f(p_i)$ are integers in the set $\{0, 1, \ldots, n - 1\}$. However, it is impossible for one person to have $n - 1$ friends (everyone else) and another to have 0 friends (no one).

- If someone has $n - 1$ friends, no one can have 0 friends. The possible values are $\{1, \ldots, n-1\}$.
- If someone has 0 friends, no one can have $n - 1$ friends. The possible values are $\{0, \ldots, n-2\}$.

In either case, the values of f range over a set of size $n - 1$. Since there are n people, by corollary 0.1, at least two people map to the same value.

範例

Example 0.12. Periodicity of Permutations. We can use the Pigeonhole Principle to prove a structural property of the symmetric group.

Let $\sigma \in S_n$. Consider the sequence of powers: $\sigma^1, \sigma^2, \sigma^3, \ldots$. Since S_n is finite, this infinite sequence must contain repetitions. We will later see that $|S_n| = n!$. Thus, there exist integers $j > i \geq 1$ such that $\sigma^i = \sigma^j$. Multiplying by the inverse permutation $(\sigma^{-1})^i$, we obtain:

$$\mathrm{id} = \sigma^{j-i}.$$

Let $k = j - i$. Then $\sigma^k = \mathrm{id}$. This integer k is called a period of σ.

The smallest such positive k is the **order** of the permutation.

範例

Corollary 0.2. *Inclusion Principle.* Let $F \subseteq E$ be finite sets. Then $|F| \leq |E|$, with equality holding if and only if $F = E$.

推論

Proof

The inclusion map $i : F \to E$ defined by $i(x) = x$ is an injection. By *corollary* 0.1, $|F| \leq |E|$. If $|F| = |E|$ but $F \neq E$, there exists $e \in E \setminus F$. Then $F \subseteq E \setminus \{e\}$, so $|F| \leq |E| - 1$, a contradiction.

■

0.5 *Infinite Sets and Cardinalities*

We now extend our scope to sets that cannot be enumerated by any finite interval $[1, n]$.

Definition 0.12. ***Infinite Sets.***
A set E is **infinite** if it is not finite. That is, for every $n \in \mathbb{N}^*$, there is no bijection between E and $[1, n]$.

定義

To compare the sizes of infinite sets, we cannot rely on counting. Instead, we appeal to the Bijection Principle (*proposition* 0.4) as the definition of size itself.

Definition 0.13. ***Equipotence.***
Two sets E and F are **equipotent** (denoted $E \sim F$) if there exists a bijection $f : E \to F$.

定義

It is often easier to construct injections in both directions than a single bijection. The following fundamental result assures us that these conditions are equivalent.

Theorem 0.3. ***Cantor-Bernstein Theorem.***
Let E and F be sets. If there exist injections $f : E \to F$ and $g : F \to E$, then E and F are equipotent.

定理

Note

The proof of this theorem requires a careful iterative construction of fixed points and is beyond the scope of this chapter, but the result is a standard tool in set theory.

Countable Sets

The smallest order of infinity is that of the natural numbers.

> **Definition 0.14.** ***Countability.***
> A set E is **countably infinite** if it is equipotent to $\mathbb{N}$. A set is **countable** if it is either finite or countably infinite.
>
> 定義

Intuitively, a set is countably infinite if its elements can be arranged in a sequence $x_0, x_1, x_2, \ldots$ indexed by $\mathbb{N}$.

> **Proposition 0.5.** ***Basic Countable Sets.***
> The following sets are countably infinite:
> 1. The positive integers $\mathbb{N}^*$.
> 2. The even natural numbers $2\mathbb{N} = \{0, 2, 4, \ldots\}$.
> 3. The integers $\mathbb{Z}$.
>
> 命題

> *Proof*
>
> We exhibit explicit bijections for each case.
>
> 1. The map $f : \mathbb{N} \to \mathbb{N}^*$ defined by $x \mapsto x + 1$ is a bijection (its inverse is $y \mapsto y - 1$).
> 2. The map $f : \mathbb{N} \to 2\mathbb{N}$ defined by $x \mapsto 2x$ is a bijection.
> 3. We can enumerate $\mathbb{Z}$ by alternating between non-negative and negative integers: $0, -1, 1, -2, 2, \ldots$. Formally, define $g : \mathbb{Z} \to \mathbb{N}$ by:
>
> $$g(x) = \begin{cases} 2x & \text{if } x \geq 0, \\ -2x - 1 & \text{if } x < 0. \end{cases}$$
>
> If $x \geq 0$, $g(x)$ maps to the even numbers $\{0, 2, 4, \ldots\}$. If $x < 0$, let $x = -k$ where $k \geq 1$; then $g(x) = 2k - 1$, mapping to the odd numbers $\{1, 3, 5, \ldots\}$. Since even and odd numbers partition $\mathbb{N}$, g is a bijection.
>
> ■

Perhaps surprisingly, increasing the dimension of the set does not increase its cardinality.

> **Proposition 0.6.** ***Countability of the Plane.***
> The Cartesian product $\mathbb{N} \times \mathbb{N}$ is countably infinite.
>
> 命題

> *Proof*
>
> We require a bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ (or to a set known to be count-

able, like $\mathbb{N}^*$). Consider the function $h : \mathbb{N} \times \mathbb{N} \to \mathbb{N}^*$ defined by:

$$h(x, y) = 2^x(2y + 1).$$

By the Fundamental Theorem of Arithmetic, every positive integer n can be uniquely written as a power of 2 multiplied by an odd number.

- **Injectivity:** If $h(x, y) = h(a, b)$, then $2^x(2y + 1) = 2^a(2b + 1)$. Equating the powers of 2 gives $x = a$, and equating the odd parts gives $2y + 1 = 2b + 1 \implies y = b$.
- **Surjectivity:** For any $n \in \mathbb{N}^*$, factor out 2s until the remainder is odd to find the unique pre-image (x, y).

Since $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}^*$ and $\mathbb{N}^* \sim \mathbb{N}$, the product is countably infinite.

■

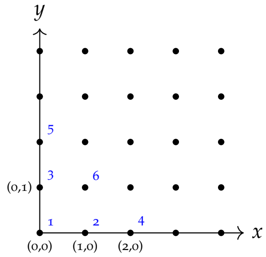

Figure 5: The lattice $\mathbb{N} \times \mathbb{N}$. The function $h(x, y) = 2^x(2y + 1)$ enumerates these points by mapping them to unique integers $1, 2, 3, \ldots$.

This result implies that taking subsets does not "break" countability.

Proposition 0.7. ***Subsets of Countable Sets.***
Every subset $E \subseteq \mathbb{N}$ is either finite or countably infinite.

命題

Proof

Suppose $E \subseteq \mathbb{N}$ is not finite. We rely on the Well-Ordering Principle (every non-empty subset of $\mathbb{N}$ has a least element) to recursively construct a bijection $f : \mathbb{N} \to E$. Define $f(0) = \min E$. Define $f(1) = \min(E \setminus \{f(0)\})$. Recursively, for $n \geq 1$:

$$f(n) = \min\left(E \setminus \{f(0), f(1), \ldots, f(n-1)\}\right).$$

This mapping is well-defined because E is infinite, so the set $E \setminus \{f(0), \ldots, f(n-1)\}$ is never empty.

- **Injectivity:** By construction, f is strictly increasing. If $n < m$, $f(m)$ is selected from a set excluding $f(n)$, so $f(n) < f(m)$.
- **Surjectivity:** Suppose $y \in E$ is not in the image of f. Let $S = \{k \in E : k \leq y\}$. Since $E \subseteq \mathbb{N}$, S is finite. However, the sequence $f(0) < f(1) < \ldots$ must eventually exceed y, which implies y would have been selected as a minimum at some step $k < |S|$. This is a contradiction.

■

Corollary 0.3. *Rational Numbers.* The set of rational numbers $\mathbb{Q}$ is countably infinite.

推論

Proof

The inclusion map $n \mapsto n$ is an injection $\mathbb{N} \to \mathbb{Q}$, so $\mathbb{Q}$ is at least infinite. Conversely, we define an injection $\mathbb{Q} \to \mathbb{Z} \times \mathbb{N}^*$. Any rational r can be uniquely written as an irreducible fraction p/q with $q > 0$. Map $r \mapsto (p, q)$. We already established bijections $\mathbb{Z} \sim \mathbb{N}$ and $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$. Composing these maps gives an injection $\mathbb{Q} \to \mathbb{N}$. By *proposition 0.7*, the image is countable, so $\mathbb{Q}$ is countable.

■

Uncountable Sets

Not all infinite sets are created equal. Georg Cantor proved that the "infinity" of the real numbers is strictly larger than that of the natural numbers. The core of this discovery is the analysis of power sets.

Theorem 0.4. ***Cantor's Theorem.***
For any set E, the sets E and $\mathcal{P}(E)$ are not equipotent. In particular, there is no surjection from E to $\mathcal{P}(E)$.

定理

Proof

Suppose for contradiction that there exists a surjection $f : E \to \mathcal{P}(E)$. Consider the "diagonal" set A defined by elements that are not members of their own image:

$$A = \{x \in E : x \notin f(x)\}.$$

Since $A \subseteq E$, we have $A \in \mathcal{P}(E)$. Because f is surjective, there must exist some $a \in E$ such that $f(a) = A$. We ask: does a belong to A?

$$a \in A \iff a \notin f(a) \iff a \notin A.$$

This is a contradiction. Thus, no such surjection exists.

■

Since $\mathcal{P}(\mathbb{N})$ is not countable, there exist infinities beyond the countable. This hierarchy continues indefinitely: $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| < |\mathcal{P}(\mathcal{P}(\mathbb{N}))| \ldots$.

Example 0.13. Binary Sequences. Let Σ^∞ be the set of all infinite binary sequences $(a_n)_{n \in \mathbb{N}}$ where $a_n \in \{0, 1\}$. This set is in bijection with $\mathcal{P}(\mathbb{N})$ via the characteristic function map: a subset $S \subseteq \mathbb{N}$ corresponds to the sequence where $a_n = 1$ if $n \in S$ and 0 otherwise. Consequently, the set of infinite binary sequences is uncountable.

範例

0.6 Exercises

1. **The Algebra of Sets.** Let A, B, C be subsets of a universal set Ω.
   (a) Using the properties of the characteristic function 1_X, prove that the symmetric difference is associative $(A\Delta B)\Delta C = A\Delta(B\Delta C)$.
   (b) Prove that the intersection distributes over the symmetric difference: $A \cap (B\Delta C) = (A \cap B)\Delta(A \cap C)$.
   (c) Conclude that $(\mathcal{P}(\Omega), \Delta, \cap)$ forms a commutative ring with identity. What is the additive identity and the multiplicative identity of this ring?
2. **Images and Pre-images.** Let $f : E \to F$ be a mapping, and let $A, B \subseteq E$.
   (a) Prove that $f(A \cup B) = f(A) \cup f(B)$.
   (b) Prove that $f(A \cap B) \subseteq f(A) \cap f(B)$, also construct a counter-example to show that equality does not hold in general for part (b).
   (c) Prove that $f(A \cap B) = f(A) \cap f(B)$ for all subsets A, B if and only if f is injective.
3. **Canonical Decomposition of a Map.** Let $f : E \to F$ be an arbitrary mapping. We define a relation $\sim_f$ on E by $x \sim_f y \iff f(x) = f(y)$.
   (a) Verify that $\sim_f$ is an equivalence relation on E.
   (b) Let $E/\sim_f$ be the quotient set and $\pi : E \to E/\sim_f$ be the canonical projection $x \mapsto [x]$. Construct a bijective map $\bar{f} : E/\sim_f \to \mathrm{Im}(f)$ such that $f = \iota \circ \bar{f} \circ \pi$, where $\iota : \mathrm{Im}(f) \hookrightarrow F$ is the inclusion map.
   (c) **Application.** If E is finite, use this decomposition to prove that $|\mathrm{Im}(f)| = |E|/k$ if and only if every fiber $f^{-1}(\{y\})$ has the same size k.
4. **Conjugation and Cycle Structure.** Let $\sigma, \tau \in S_n$ be permutations.
   (a) Prove that if σ has the cycle decomposition $(c_1, c_2, \ldots, c_k)$, then the conjugate permutation $\tau \circ \sigma \circ \tau^{-1}$ has the cycle decomposition $(\tau(c_1), \tau(c_2), \ldots, \tau(c_k))$.
   (b) Use this to show that two permutations are conjugate (i.e., $\rho = \tau\sigma\tau^{-1}$ for some τ) if and only if they have the same number of cycles of each length.
   (c) Find a $\tau \in S_4$ such that $\tau \circ (1,2)(3,4) \circ \tau^{-1} = (1,3)(2,4)$.
5. **The Subset Sum Problem.** Let S be a subset of $[1, 14]$ with $|S| = 6$.
   (a) Show that the number of distinct non-empty subsets of S is 63.
   (b) Calculate the maximum possible sum of elements of a subset of S.
   (c) Use the Pigeonhole Principle to prove that there exist two dis-

tinct disjoint subsets $A, B \subseteq S$ such that the sum of elements in A equals the sum of elements in B.

(d) Generalise this result: finding a condition on n and k such that any subset of size k from $[1, n]$ contains two disjoint subsets with equal sums.

6. **Derangements and Inclusion-Exclusion.** Recall that for sets A, B, the characteristic function satisfies $1_{A\cup B} = 1 - (1 - 1_A)(1 - 1_B)$.

(a) Generalise this identity to a finite collection of sets $A_1, \ldots, A_n \subseteq \Omega$. Show that the size of their union is given by:

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{\varnothing \neq I \subseteq [1,n]} (-1)^{|I|-1} \left|\bigcap_{i\in I} A_i\right|.$$

(b) Let S_n be the symmetric group. For each $k \in [1, n]$, let $F_k = \{\sigma \in S_n \mid \sigma(k) = k\}$ be the set of permutations fixing k. Calculate the cardinality of the intersection of any m such sets.

(c) A permutation is a *derangement* if it has no fixed points. Let D_n denote the number of derangements in S_n. Use the Inclusion-Exclusion Principle to prove:

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

7. **Algebraic and Transcendental Numbers.** A real number α is called *algebraic* if it is a root of a non-zero polynomial $P(x) = a_n x^n + \cdots + a_0$ with integer coefficients ($a_i \in \mathbb{Z}$). The *height* of such a polynomial is defined as $h(P) = n + \sum_{i=0}^{n} |a_i|$.

(a) Prove that for any integer $H \geq 1$, the set of polynomials $P \in \mathbb{Z}[x]$ satisfying $h(P) \leq H$ is finite.

(b) Using the fact that a countable union of finite sets is countable, prove that the set $\mathbb{Z}[x]$ of all polynomials with integer coefficients is countably infinite.

(c) Since a polynomial of degree n has at most n real roots, prove that the set of all algebraic numbers $\mathcal{A}$ is countable.

(d) A real number is *transcendental* if it is not algebraic. Assuming the result that $\mathbb{R}$ is uncountable, prove that the set of transcendental numbers is uncountable.

(e) **Reflection.** Reconcile the following: "Most" numbers are transcendental (in the sense of cardinality), yet it is generally much harder to prove a specific number is transcendental than to prove it is algebraic.

1

Principles of Counting

The objective of enumerative combinatorics is to determine the cardinality of specific sets, often described by parameters. Rather than listing elements exhaustively, we seek to express these cardinalities as functions of those parameters. In this chapter, we formalise the algebraic principles that allow us to reduce complex counting problems to elementary set operations.

1.1 *The Addition Principle*

We begin with the most intuitive property of counting: if two sets have no elements in common, the count of their union is the sum of their individual counts.

> **Proposition 1.1.** ***Addition Principle.***
> Let E and F be disjoint finite sets (i.e., $E \cap F = \varnothing$). Then:
>
> $$|E \cup F| = |E| + |F|.$$
>
> 命题

> *Proof*
>
> If either set is empty, the conclusion follows from the definition of cardinality. Otherwise, let $|E| = n$ and $|F| = m$. By definition, there exist bijections $f : E \to [1, n]$ and $g : F \to [1, m]$. We construct a function $h : E \cup F \to [1, n+m]$ defined by:
>
> $$h(x) = \begin{cases} f(x) & \text{if } x \in E, \\ g(x) + n & \text{if } x \in F. \end{cases}$$
>
> Since the image of E under h is $[1, n]$ and the image of F is $\{n + 1, \ldots, n + m\}$, the ranges are disjoint and cover $[1, n + m]$. Since f and g are bijections, h is a bijection. Thus, $|E \cup F| = n + m$.
>
> ■

This inductive pattern yields the following corollary.

We provide a prove in the later chapters.

Corollary 1.1. *Generalised Addition Principle.* Let $E_1, \ldots, E_n$ be pairwise disjoint finite sets. Then:

$$\left| \bigcup_{i=1}^{n} E_i \right| = \sum_{i=1}^{n} |E_i|.$$

推論

When sets are not disjoint, simply adding their cardinalities overcounts the elements in their intersection. To correct this, we subtract the intersection.

Corollary 1.2. *Inclusion-Exclusion Principle (Two Sets).* Let E and F be finite sets. Then:

$$|E \cup F| = |E| + |F| - |E \cap F|.$$

推論

Proof

Let $G = E \cap F$. We may decompose the sets into disjoint components:

$$E = (E \setminus G) \sqcup G \quad \text{and} \quad F = (F \setminus G) \sqcup G,$$

where $\sqcup$ denotes a disjoint union. Similarly, the union decomposes as:

$$E \cup F = (E \setminus G) \sqcup (F \setminus G) \sqcup G.$$

Applying *proposition 1.1* to these disjoint unions:

$$|E| = |E \setminus G| + |G| \implies |E \setminus G| = |E| - |G|.$$

$$|F| = |F \setminus G| + |G| \implies |F \setminus G| = |F| - |G|.$$

$$|E \cup F| = |E \setminus G| + |F \setminus G| + |G|.$$

Substituting the expressions for the differences:

$$|E \cup F| = (|E| - |G|) + (|F| - |G|) + |G| = |E| + |F| - |G|.$$

Since $G = E \cap F$, the result follows.

■

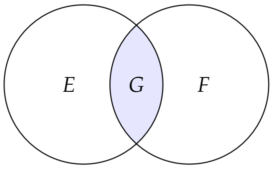

Figure 1.1: Decomposition of $E \cup F$ into three disjoint sets: $E \setminus G$, $F \setminus G$, and the intersection $G = E \cap F$.

The decomposition underlies the next example.

Example 1.1. Lattice Points in Overlapping Rectangles. Consider two discrete rectangular regions in $\mathbb{N}^2$:

$$A = [1,3] \times [1,2] \quad \text{and} \quad B = [2,4] \times [2,3].$$

We wish to calculate $|A \cup B|$.

First, we determine the size of each set using the definition of the

Cartesian product:

- $|A| = |[1,3]| \cdot |[1,2]| = 3 \cdot 2 = 6.$
- $|B| = |[2,4]| \cdot |[2,3]| = 3 \cdot 2 = 6.$

Next, we identify the intersection $A \quad \cap \quad B$. A point (x, y) is in the intersection if $x \in [1,3] \cap [2,4]$ and $y \in [1,2] \cap [2,3]$.

$$[1,3] \cap [2,4] = \{2,3\}, \quad [1,2] \cap [2,3] = \{2\}.$$

Thus, $A \cap B = \{2,3\} \times \{2\}$, and $|A \cap B| = 2 \cdot 1 = 2$. By *corollary 1.2*:

$$|A \cup B| = 6 + 6 - 2 = 10.$$

範例

The Shepherd's Principle

We formalise a technique often summarised as "to count the sheep, count the legs and divide by four". This principle relates the cardinality of a domain to the cardinality of a codomain via the structure of the mapping between them.

Proposition 1.2. ***Sum of Fibres.***
Let E and F be finite sets and let $f : E \to F$ be a mapping. Then:

$$|E| = \sum_{y \in F} |f^{-1}(y)|.$$

命題

Proof

The set E is partitioned by the fibres of f. Specifically,

$$E = \bigcup_{y \in F} f^{-1}(y).$$

If $y \neq z$, then $f^{-1}(y) \cap f^{-1}(z) = \varnothing$, as no element can map to both y and z. By the *Generalised Addition Principle*, the cardinality of E is the sum of the cardinalities of these disjoint fibres.

■

This immediately provides a bound for surjective mappings.

Corollary 1.3. *Surjection Bound.* If $f : E \to F$ is a surjective map between finite sets, then $|F| \leq |E|$.

推論

Proof

Since f is surjective, every fibre is non-empty, so $|f^{-1}(y)| \geq 1$ for all

$y \in F$. Using *proposition 1.2*:

$$|E| = \sum_{y \in F} |f^{-1}(y)| \geq \sum_{y \in F} 1 = |F|.$$

■

The most powerful application arises when the map is regular, i.e., every fibre has the same size.

Corollary 1.4. *Shepherd's Principle.* Let $f : E \to F$ be a map between finite sets. If there exists a constant $p \in \mathbb{N}^*$ such that $|f^{-1}(y)| = p$ for all $y \in F$, then:

$$|E| = p \cdot |F|.$$

推論

Proof

Applying *proposition 1.2*:

$$|E| = \sum_{y \in F} p = p \cdot \sum_{y \in F} 1 = p \cdot |F|.$$

■

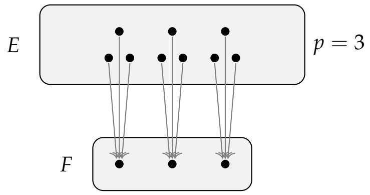

Figure 1.2: Visualisation of the Shepherd's Principle with $p = 3$. Each element in F "pulls back" to exactly 3 elements in E.

Example 1.2. Counting via Projections. Let S be the set of ordered pairs $(i, j) \in [1, 4] \times [1, 4]$ such that $i < j$. We wish to find $|S|$. Consider the mapping $f : S \to [1, 3]$ defined by $f(i, j) = i$. The image is indeed $[1, 3]$ because if $i = 4$, no $j \in [1, 4]$ satisfies $4 < j$. Let us examine the fibres:

- For $y = 1$: $f^{-1}(1) = \{(1, 2), (1, 3), (1, 4)\}$, so $|f^{-1}(1)| = 3$.
- For $y = 2$: $f^{-1}(2) = \{(2, 3), (2, 4)\}$, so $|f^{-1}(2)| = 2$.
- For $y = 3$: $f^{-1}(3) = \{(3, 4)\}$, so $|f^{-1}(3)| = 1$.

Here the fibres are not of uniform size, so we apply *proposition 1.2*:

$$|S| = 3 + 2 + 1 = 6.$$

範例

The Multiplication Principle

The Shepherd's Principle allows us to derive the cardinality of Cartesian products.

Proposition 1.3. ***Multiplication Principle.***
Let E and F be finite sets. Then:

$$|E \times F| = |E| \cdot |F|.$$

命題

Proof

Consider the projection mapping $\pi : E \times F \to F$ defined by $\pi(x,y) = y$. For any fixed $y \in F$, the fibre is:

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

The map $x \mapsto (x,y)$ is a bijection from E to $\pi^{-1}(y)$. Thus, $|\pi^{-1}(y)| = |E|$ for every $y \in F$. By the *Shepherd's Principle* with $p = |E|$:

$$|E \times F| = |E| \cdot |F|.$$

■

By induction, this extends to finite sequences of sets.

Corollary 1.5. *Generalised Multiplication.* Let $E_1, \dots, E_n$ be finite sets. Then:

$$|E_1 \times E_2 \times \cdots \times E_n| = \prod_{i=1}^{n} |E_i|.$$

推論

Proof

We proceed by induction on n. The base case $n = 1$ is trivial. Assume the result holds for $n-1$. Let $A = E_1 \times \cdots \times E_{n-1}$. Then $E_1 \times \cdots \times E_n$ is naturally identified with $A \times E_n$. By *proposition 1.3*:

$$|A \times E_n| = |A| \cdot |E_n|.$$

By the inductive hypothesis, $|A| = |E_1| \dots |E_{n-1}|$, completing the proof.

■

This leads to the counting of strings or tuples of fixed length.

Corollary 1.6. *Cardinality of Powers.* If E is a finite set, then

$$|E^n| = |E|^n.$$

推論

Proof

Apply *corollary 1.5* with $E_1 = E_2 = \cdots = E_n = E$. Then

$$|E^n| = |E_1 \times \cdots \times E_n| = \prod_{i=1}^{n} |E_i| = \underbrace{|E| \cdot |E| \cdots |E|}_{n \text{ factors}} = |E|^n.$$

■

Example 1.3. Alphabetical Combinations. How many distinct words of length 5 can be formed using the standard 26-letter alpha-

bet $\mathcal{A} = \{a, b, \ldots, z\}$?
A word of length 5 is an element of $\mathcal{A}^5$. By *corollary* 1.6:

$$|\mathcal{A}^5| = |\mathcal{A}|^5 = 26^5 = 11,881,376.$$

範例

Example 1.4. Decimal Representations. How many non-negative integers have at most 3 digits?
These integers correspond to the set $\{0, 1, \ldots, 999\}$. Each can be uniquely represented as a triplet $(d_2, d_1, d_0) \in \{0, \ldots, 9\}^3$ by padding with leading zeros (e.g., $42 \mapsto 042$). The set of digits $D = \{0, \ldots, 9\}$ has cardinality 10. Thus, the total count is $|D^3| = 10^3 = 1000$.

範例

Example 1.5. Restricted Digits. How many integers with at most 3 digits consist entirely of even digits?
The set of allowed digits is $E = \{0, 2, 4, 6, 8\}$, with $|E| = 5$. As before, we pad with leading zeros to identify each integer with a triple in E^3. We are counting elements of E^3.

$$|E^3| = 5^3 = 125.$$

範例

Example 1.6. Counting Functions. Let $A = \{a, b, c\}$ and $B = \{0, 1\}$. How many distinct functions $f : A \to B$ exist?
A function is uniquely determined by the triplet of values $(f(a), f(b), f(c))$. Since each value must belong to B, the set of all such functions is in bijection with the Cartesian product $B \times B \times B = B^3$. By *corollary* 1.6:

$$|B^3| = |B|^3 = 2^3 = 8.$$

In general, the number of mappings from a set of size n to a set of size m is m^n.

範例

1.2 *Arrangements and Permutations*

We now apply these principles to solve classical enumeration problems. These configurations serve as the building blocks for more complex combinatorial structures. We begin by counting sequences where order matters.
Consider the construction of a sequence of length k using symbols

from a finite set (an alphabet).

> **Proposition 1.4.** ***Words of Length k.***
> Let Σ be a finite alphabet with $|\Sigma| = n$. The number of words of length k (sequences of k elements from Σ) is n^k.
>
> 命題

> *Proof*
>
> A word of length k is an element of the Cartesian product Σ^k. By *corollary 1.6*, $|\Sigma^k| = |\Sigma|^k = n^k$.
>
> ■

> *Remark.*
>
> Order is part of the data: for instance, 042 and 240 are different words of length 3 over $\{0, \dots, 9\}$.

If we require the elements in the sequence to be distinct, we are counting **arrangements** (or permutations of subsets).

> **Definition 1.1.** ***Falling Factorial Power.***
> For $n \in \mathbb{N}$ and $k \in \mathbb{N}$, the **falling factorial power**, denoted $n^{\underline{k}}$ (or sometimes $(n)_k$), is the product of k terms starting at n and decreasing by 1:
>
> $$n^{\underline{k}} = \prod_{i=0}^{k-1} (n - i) = n(n-1)\cdots(n-k+1).$$
>
> If $k > n$, then $n^{\underline{k}} = 0$. By convention, $n^{\underline{0}} = 1$.
>
> 定義

> **Proposition 1.5.** ***Ordered Selections.***
> The number of ways to choose k distinct elements from a set of size n and arrange them in a sequence is $n^{\underline{k}}$.
>
> 命題

> *Proof*
>
> We construct the sequence element by element.
> - There are n choices for the first element.
> - There are $n - 1$ choices for the second element.
> - ...
> - There are $n - (k - 1)$ choices for the k-th element.
>
> By *corollary 1.5*, the total count is $n(n-1)\cdots(n-k+1) = n^{\underline{k}}$.
>
> ■

In the specific case where $k = n$, we are arranging the entire set.

> **Definition 1.2.** ***Factorial.***
> The **factorial** of a non-negative integer n is defined as $n! = n^{\underline{n}}$. Ex-

plicitly:

$$n! = \prod_{i=1}^{n} i = 1 \cdot 2 \cdots n.$$

We define $0! = 1$.

定義

Corollary 1.7. *Permutations.* The number of ways to arrange n distinct objects in a row (i.e., the number of bijections from a set of size n to itself) is $n!$.

推論

Example 1.7. Signal Flags. Suppose a ship has 10 distinct signal flags.

- If the ship hoists a sequence of 3 flags on a mast, the number of possible signals is $10^{\underline{3}} = 10 \cdot 9 \cdot 8 = 720$.
- If the ship arranges all 10 flags in a row for inspection, the number of arrangements is $10! = 3,628,800$.

範例

Circular Arrangements

Counting arrangements on a circle differs from the linear case because absolute positions do not exist; only relative order matters. Two circular configurations are considered identical if one can be obtained from the other by rotation.

Proposition 1.6. ***Circular Permutations.***
The number of ways to arrange n distinct objects around a circle is $(n-1)!$.

命題

Proof

We apply *corollary* 1.4. Let L be the set of linear arrangements of the n objects, so $|L| = n!$. Let C be the set of distinct circular arrangements. Consider the map $f : L \to C$ that takes a linear arrangement $(x_1, \ldots, x_n)$ and wraps it into a circle. For any specific circular arrangement, there are exactly n linear arrangements that produce it (corresponding to starting the read-out at any of the n positions). Thus, the map f is n-to-one. By *corollary* 1.4:

$$|L| = n \cdot |C| \implies |C| = \frac{|L|}{n} = \frac{n!}{n} = (n-1)!.$$

■

Alternatively, one may fix a distinguished element ("the head") at the "top" of the circle to break the rotational symmetry. The remaining

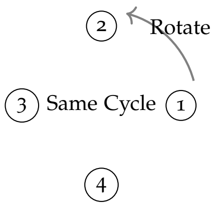

Figure 1.3: For $n = 4$, the linear sequences $(1,2,3,4)$, $(2,3,4,1)$, $(3,4,1,2)$, and $(4,1,2,3)$ all represent the same circular arrangement.

$n-1$ elements can then be arranged linearly in the remaining $n-1$ positions in $(n-1)!$ ways.

1.3 Counting Mappings

We now turn to the enumeration of functions between finite sets E and F, classifying them by their properties (arbitrary, injective, surjective, bijective).

Total Mappings and Subsets

Let $\mathcal{F}(E,F)$ denote the set of all functions $f : E \to F$.

Proposition 1.7. ***Number of Mappings.***
Let E and F be finite sets. Then:

$$|\mathcal{F}(E,F)| = |F|^{|E|}.$$

命題

Proof

Let $E = \{x_1, \ldots, x_k\}$ where $k = |E|$. A function $f \in \mathcal{F}(E,F)$ is uniquely determined by the tuple of its values $(f(x_1), \ldots, f(x_k))$. The map $\phi : \mathcal{F}(E,F) \to F^k$ defined by $f \mapsto (f(x_1), \ldots, f(x_k))$ is a bijection. By *corollary* 1.6, $|F^k| = |F|^k = |F|^{|E|}$.

■

This result provides a combinatorial proof for the cardinality of the power set.

Theorem 1.1. ***Cardinality of the Power Set.***
For any finite set E, the number of subsets is $|\mathcal{P}(E)| = 2^{|E|}$.

定理

Proof

Recall from *proposition* 0.1 that subsets of E are in bijection with their characteristic functions. The map $A \mapsto 1_A$ is a bijection from $\mathcal{P}(E)$ to $\mathcal{F}(E,\{0,1\})$. Applying the previous proposition with $F = \{0,1\}$:

$$|\mathcal{P}(E)| = |\mathcal{F}(E,\{0,1\})| = 2^{|E|}.$$

■

Injections and Bijections

Let $\mathcal{F}_{\text{inj}}(E,F)$ denote the set of injective mappings from E to F.

Proposition 1.8. ***Number of Injections.***
Let $|E| = k$ and $|F| = n$.

$$|\mathcal{F}_{\text{inj}}(E, F)| = n^{\underline{k}} = \frac{n!}{(n-k)!}.$$

If $k > n$, the value is 0.

命題

Proof

Let $E = \{x_1, \ldots, x_k\}$. To define an injection, we must assign distinct images in F to the elements of E. This is equivalent to choosing an ordered sequence of k distinct elements from F (where the i-th element of the sequence is the image of x_i). By *proposition* 1.5, the number of such ordered selections is $n^{\underline{k}}$.

■

Example 1.8. Server Allocation. A data centre needs to assign 3 distinct processing jobs (J_1, J_2, J_3) to a cluster of 50 available servers. No server may handle more than one job.
This is an injection from the set of jobs to the set of servers. The number of possible assignments is $50^{\underline{3}} = 50 \times 49 \times 48 = 117,600$.

範例

Corollary 1.8. *Number of Bijections.* Let E and F be finite sets with $|E| = |F| = n$. The number of bijections from E to F is $n!$.

推論

Proof

By *lemma* 0.1, a map between sets of equal finite cardinality is bijective if and only if it is injective. Thus, we count the injections:

$$n^{\underline{n}} = n(n-1)\cdots(1) = n!.$$

■

Surjections and Partitions

Counting surjective mappings is more subtle. We rely on the relationship between functions and partitions.

Definition 1.3. ***Stirling Numbers of the Second Kind.***
The **Stirling number of the second kind**, denoted $S(n, k)$, is the number of ways to partition a set of n elements into k non-empty subsets.

定義

Proposition 1.9. ***Surjections via Partitions.***
Let E and F be finite sets with $|E| = n$ and $|F| = k$. The number of surjective functions from E to F is:

$$|\mathcal{F}_{\text{surj}}(E,F)| = k! \cdot S(n,k).$$

命題

Proof

Every surjection $f : E \to F$ induces a partition of E into k non-empty fibres $\{f^{-1}(y) : y \in F\}$. Conversely, given a partition of E into k parts, we can construct a surjection by assigning each part to a unique element of F. There are $k!$ ways to assign these k parts to the k elements of F (permutations of F). Thus each partition yields exactly $k!$ surjections, and the total count is $k! \cdot S(n,k)$.

■

To compute $S(n,k)$, we use a recurrence relation derived by considering the placement of a specific element.

Theorem 1.2. ***Recurrence for Stirling Numbers.***
For $1 \leq k \leq n$:

$$S(n,k) = S(n-1,k-1) + k \cdot S(n-1,k).$$

The boundary conditions are $S(n,1) = 1$ and $S(n,n) = 1$.

定理

Proof

Let $E = \{1,\ldots,n\}$. Consider the last element n. In any partition of E into k parts, there are two mutually exclusive possibilities:

Type 1. The element n forms a singleton set $\{n\}$. Removing this set leaves a partition of $\{1,\ldots,n-1\}$ into $k-1$ parts. There are $S(n-1,k-1)$ such partitions.

Type 2. The element n belongs to a set with other elements. If we remove n from its block, we are left with a partition of $\{1,\ldots,n-1\}$ into k parts. To reconstruct the original partition, we could have added n to any of the k existing blocks. Thus, there are $k \cdot S(n-1,k)$ such partitions.

By *proposition 1.1*, $S(n,k) = S(n-1,k-1) + kS(n-1,k)$.

■

Example 1.9. Study Groups. We wish to split a group of 4 students into 2 non-empty study teams. The order of teams does not matter, only the grouping. This is $S(4,2)$.

Using the recurrence:

$$\begin{aligned} S(4,2) &= S(3,1) + 2S(3,2) \\ &= 1 + 2(S(2,1) + 2S(2,2)) \\ &= 1 + 2(1 + 2(1)) \\ &= 1 + 2(3) = 7. \end{aligned}$$

The 7 partitions correspond to:

- 4 splits of type $3+1$ (one student works alone).
- 3 splits of type $2+2$ (pairs).

範例

Analytical Properties of the Factorial

The factorial function grows extremely rapidly. While precise evaluation requires computation, we can establish useful bounds using elementary inequalities.

Theorem 1.3. ***Bounds on the Factorial.***
For all $n \geq 1$:

$$\sqrt{n}^n \leq n! \leq \left(\frac{n+1}{2}\right)^n.$$

定理

We require the following lemma.

Lemma 1.1. ***AM-GM Inequality.***
For non-negative real numbers a, b:

$$\sqrt{ab} \leq \frac{a+b}{2},$$

with equality if and only if $a = b$.

引理

Proof

This follows from the square of a real number being non-negative:

$$0 \leq (\sqrt{a} - \sqrt{b})^2 = a - 2\sqrt{ab} + b \implies 2\sqrt{ab} \leq a + b.$$

■

Proof of theorem 1.3

Consider the square of the factorial:

$$(n!)^2 = (1 \cdot 2 \cdots n) \cdot (n \cdot (n-1) \cdots 1) = \prod_{k=1}^{n} k(n+1-k).$$

Taking the square root:

$$n! = \prod_{k=1}^{n} \sqrt{k(n+1-k)}.$$

Upper Bound. Apply the AM-GM inequality to each term k and $n+1-k$:

$$\sqrt{k(n+1-k)} \le \frac{k+(n+1-k)}{2} = \frac{n+1}{2}.$$

Multiplying these n inequalities yields $n! \le \left(\frac{n+1}{2}\right)^n$.

Lower Bound. We observe that the function $f(x) = x(n+1-x)$ is a parabola opening downwards, with minimum values at the endpoints of the interval $[1, n]$. For $k \in [1, n]$, $k(n+1-k) \ge 1(n+1-1) = n$. Thus, $\sqrt{k(n+1-k)} \ge \sqrt{n}$. Multiplying these n terms yields $n! \ge (\sqrt{n})^n$.

■

For large n, the behaviour of $n!$ is described precisely by Stirling's Formula.

Theorem 1.4. ***Stirling's Formula.***
As $n \to \infty$,

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

meaning that

$$\lim_{n\to\infty} \frac{n!}{n^n e^{-n} \sqrt{2\pi n}} = 1,$$

and more precisely, for all $n \ge 1$,

$$n^n e^{-n} \sqrt{2\pi n} \exp\left(\frac{1}{12n+1}\right) < n! < n^n e^{-n} \sqrt{2\pi n} \exp\left(\frac{1}{12n}\right).$$

The inequalities quantify the error in the limit.

定理

Note

The proof of Stirling's formula requires analytical techniques (such as the Gamma function or integral approximations) beyond the scope of this chapter.

1.4 *Exercises*

1. **Cycle Structures and Involutions.** The text establishes that the cycle notation is unique up to the ordering of disjoint cycles and the cyclic shift of elements within a cycle (e.g., $(1, 2, 3)$ is identical

to $(2, 3, 1)$).

(a) Prove that the number of distinct cycles of length k that can be formed from a fixed subset of k elements is $(k-1)!$.

(b) Let $\sigma \in S_9$. Determine the number of permutations consisting of exactly three disjoint cycles: one of length 4, one of length 3, and one of length 2.

(c) Recall that an *involution* is a permutation composed solely of cycles of length 1 (fixed points) and length 2 (transpositions). Determine the number of involutions in S_5.

2. **The Divisor Function.** Let $S_1, \dots, S_t$ be pairwise disjoint finite sets with cardinalities $|S_i| = a_i$.

(a) Prove that the number of subsets of the union $\bigcup_{i=1}^{t} S_i$ that contain at most one element from each S_i is given by

$$P = \prod_{i=1}^{t} (a_i + 1).$$

(b) Let n be a positive integer with prime factorisation $n = p_1^{a_1} \dots p_t^{a_t}$. Let $\tau(n)$ denote the number of positive divisors of n. Using the result above, prove that $\tau(n) = \prod_{i=1}^{t}(a_i + 1)$.

(c) Deduce that $\tau(n)$ is odd if and only if n is a perfect square.

3. **Binary Expansion and Sums.**

(a) By interpreting the sum as a count of non-empty subsets of specific types, or by using the geometric series formula, prove:

$$\sum_{k=0}^{n} 2^k = 2^{n+1} - 1.$$

(b) Evaluate the sum $S = \sum_{k=1}^{n} (n-k) 2^{k-1}$.

> *Remark.*
>
> Consider the total cardinality of all subsets of $\{1, \dots, n\}$ containing at least two elements, or count pairs (A, x) where $A \subseteq \{1, \dots, n\}$ and $x \in A$ is the second largest element.

4. **Administrative Inconsistency.** The chair of a mathematics department decrees that every student must enrol in exactly 4 of the 7 available courses. The registrars report the following enrolment numbers for the courses: 51, 30, 30, 20, 25, 12, and 18. Use the Shepherd's Principle (or simple counting of student-course pairs) to demonstrate that the registrars' data must be erroneous.

5. **Finite Mappings.** Let E be a finite set with $|E| = n$.

(a) A function $f : E \to E$ is called a *retraction* (or idempotent) if

$f \circ f = f$. Prove that f is a retraction if and only if $f(x) = x$ for all $x \in \mathrm{Im}(f)$.

(b) Determine the number of retractions on E.

For (b): Classify by the size of the image set k.

(c) Prove that for finite sets, a map $f : E \to E$ is injective if and only if it is surjective. Give a counter-example for infinite sets.

6. **Counting Relations.** Let E be a set with $|E| = n$. Recall that a binary relation is a subset of $E \times E$. Determine the number of relations on E that are:

   (a) Reflexive.
   (b) Symmetric.
   (c) Both reflexive and symmetric.
   (d) **Neither** symmetric nor antisymmetric.

7. **Contiguous Permutations.** We wish to count the number of permutations σ of $\{1, \ldots, n\}$ such that the set of elements $\{\sigma(1), \ldots, \sigma(k)\}$ forms a set of consecutive integers (an interval) for every $k = 1, \ldots, n$.

   (a) List all such permutations for $n = 3$ and $n = 4$.
   (b) Prove that the total number of such permutations is 2^{n-1}.

   > *Remark.*
   >
   > Consider the possible values for the last element $\sigma(n)$ relative to the preceding set.

8. **Euler's Summation Identity.** Let $\varphi(n)$ be the Euler totient function.

   (a) Let $S = \{1, \ldots, n\}$. Partition S into disjoint sets $A_d = \{k \in S : (k, n) = d\}$ where d runs through the divisors of n.
   (b) Prove that $|A_d| = \varphi(n/d)$.
   (c) Deduce the identity $\sum_{d|n} \varphi(d) = n$.

9. **Visual Sums.**

   (a) By arranging unit squares into a staircase shape (a "Young diagram") for $1, \ldots, n$ and joining two such staircases, give a geometric proof that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.
   (b) By considering the concentric "L-shaped" gnomons of a square of side length $n(n+1)/2$, or otherwise, prove that

$$\sum_{k=1}^{n} k^3 = \left(\sum_{k=1}^{n} k \right)^2 .$$

10. **Translational Pigeonhole.** Let A be a subset of $\{1, 2, \ldots, 100\}$ with cardinality $|A| = 55$.

(a) Consider the set $B = \{x + 9 : x \in A\}$. What is the range of values in B?

(b) Apply the Pigeonhole Principle (via intersection cardinality) to A and B to prove that there exist distinct elements $x, y \in A$ such that $|x - y| = 9$.

(c) Does the property hold if $|A| = 54$? Construct a counterexample or prove it does.

11. **Partial Functions.** A partial function from E to F is a function defined on a subset $D \subseteq E$ (the domain of definition) mapping to F.

    (a) Let $|E| = n$ and $|F| = m$. Prove that the number of partial functions from E to F is $(m+1)^n$.

    (b) Explain this result by adjoining a special "undefined" element $\perp$ to F.

12. **Stirling Calculations.**

    (a) Using the recurrence relation $S(n,k) = S(n-1,k-1) + kS(n-1,k)$, compute the table of Stirling numbers of the second kind for $n = 1$ to $n = 5$.

    (b) A "rhyme scheme" for a poem of n lines can be modeled as a partition of the set of lines $\{1, \ldots, n\}$ into rhyming groups. If we distinguish the order of appearance of rhyme sounds (e.g., AABB is distinct from BBAA), the number of schemes is the Bell number $B_n = \sum_k S(n,k)$. Calculate B_5.

13. **Non-Attacking Rooks.**

    (a) In how many ways can 8 rooks be placed on a standard 8 × 8 chessboard such that no two rooks share a row or column?

    (b) Generalise this to placing k rooks on an $n \times n$ board ($k \leq n$) such that no two attack each other.

    (c) Suppose the board has a "hole" at position $(1,1)$ (i.e., no rook can be placed there). How many ways can n non-attacking rooks be placed on this defective $n \times n$ board?

For (c): Use the subtraction principle: Total arrangements minus those where a rook is at $(1,1)$.

14. **Symmetric Boolean Functions.** A Boolean function is a map $f : \{0,1\}^n \to \{0,1\}$.

    (a) Determine the total number of Boolean functions of n variables.

    (b) A Boolean function is *symmetric* if its value depends only on the number of 1s in the input (i.e., the weight of the input vector). Determine the number of symmetric Boolean functions of n variables.

2

Combinatorial Coefficients

We now focus on the specific problem of counting k-element subsets, which yields the binomial coefficients.

2.1 Binomial Numbers

Recall that $\mathcal{P}_k(E)$ denotes the set of all subsets of E with cardinality k. We formalise the size of this set as a distinct combinatorial quantity.

Definition 2.1. ***Binomial Coefficient.***
Let $n, k \in \mathbb{N}$. The **binomial coefficient** $\binom{n}{k}$ (read "n choose k") is the number of subsets of size k of a set of size n:

$$\binom{n}{k} = |\mathcal{P}_k([1,n])|.$$

If $k < 0$ or $k > n$, we define $\binom{n}{k} = 0$.

定義

We previously determined the number of k-arrangements (ordered sequences of distinct elements) to be $n^{\underline{k}}$. Since a subset is an unordered selection, we can derive the formula for $\binom{n}{k}$ by "forgetting" the order.

Theorem 2.1. ***Factorial Formula.***
For $n \in \mathbb{N}$ and $0 \leq k \leq n$:

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n!}{k!(n-k)!}.$$

定理

Proof

Let $E = [1,n]$. Let $\mathcal{A}_k$ be the set of k-arrangements of E, and let $\mathcal{S}_k = \mathcal{P}_k(E)$. Consider the mapping $\phi : \mathcal{A}_k \to \mathcal{S}_k$ defined by:

$$\phi(x_1, \ldots, x_k) = \{x_1, \ldots, x_k\}.$$

For any subset $A \in \mathcal{S}_k$, the fibre $\phi^{-1}(A)$ consists of all possible orderings of the elements of A. By the result on permutations, $|\phi^{-1}(A)| = k!$. Since the fibres have constant size, we apply the *Shepherd's Principle*:

$$|\mathcal{A}_k| = k! \cdot |\mathcal{S}_k|.$$

Substituting $|\mathcal{A}_k| = \frac{n!}{(n-k)!}$, we obtain $|\mathcal{S}_k| = \frac{n!}{k!(n-k)!}$.

■

Fundamental Identities

While the factorial formula allows for computation, combinatorial reasoning often provides more insight into the properties of these numbers.

Proposition 2.1. ***Symmetry.***
For $0 \leq k \leq n$,

$$\binom{n}{k} = \binom{n}{n-k}.$$

命题

Proof

Consider the mapping $f : \mathcal{P}_k([1,n]) \to \mathcal{P}_{n-k}([1,n])$ defined by taking the complement:

$$f(A) = [1,n] \setminus A.$$

This map is a bijection (its inverse is the same operation). Thus, the cardinalities are equal. Combinatorially, choosing k elements to include is equivalent to choosing $n-k$ elements to exclude.

■

One of the most essential recurrences for binomial coefficients allows their construction without direct factorial computation.

Theorem 2.2. ***Pascal's Identity.***
For $n \in \mathbb{N}^*$ and $k \in \mathbb{Z}$:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

定理

Proof

If $k < 0$ or $k > n$, the identity holds trivially (0 = 0). If $k = 0$, both sides equal 1. Assume $1 \leq k \leq n$. Let $E = [1,n]$. We partition $\mathcal{P}_k(E)$ based on whether the specific element n is included in the subset.

Subsets containing n. These are of the form $A' \cup \{n\}$, where $A' \subseteq [1, n-1]$ has size $k-1$. There are $\binom{n-1}{k-1}$ such sets.

Subsets not containing n. These are subsets of $[1, n-1]$ of size k.

There are $\binom{n-1}{k}$ such sets.

By the *Addition Principle*, $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

■

The Binomial Theorem

The name "binomial coefficient" arises from the expansion of powers of a binomial $(a+b)$.

Theorem 2.3. ***Newton's Binomial Theorem.***
Let R be a commutative ring and let $a, b \in R$. For any $n \in \mathbb{N}$:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

定理

Proof

We offer a combinatorial proof. Consider the product:

$$(a+b)^n = \underbrace{(a+b)\cdot(a+b)\cdots(a+b)}_{n \text{ factors}}.$$

Expanding this product involves choosing one term (either a or b) from each of the n factors. A term of the form $a^k b^{n-k}$ is generated whenever we choose a from exactly k factors and b from the remaining $n - k$ factors. The number of ways to choose the k factors contributing an a is precisely the number of subsets of indices of size k, which is $\binom{n}{k}$. Summing over all possible values of k yields the result.

■

Remark.

One may also prove this by induction using *theorem 2.2*. For the inductive step:

$$\begin{aligned}(a+b)^{n+1} &= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}\\ &= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1}.\end{aligned}$$

Re-indexing the sums to align powers of a and b recovers the coefficient $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$.

This theorem immediately yields sums over rows of Pascal's triangle.

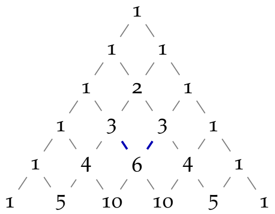

Figure 2.1: Pascal's Triangle. Lines connect each entry to the two above it whose sum it equals, illustrating Pascal's Identity. The highlighted example shows $3+3=6$.

Corollary 2.1. *Sum of Coefficients.* For $n \in \mathbb{N}$:

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

推論

Proof

Set $a = 1, b = 1$ in *theorem 2.3* to obtain $(1+1)^n = 2^n$. Set $a = -1, b = 1$ to obtain $(-1+1)^n = 0$.

■

Double Counting and Identities

A powerful method for proving combinatorial identities is double counting (or Fubini's Principle): counting the size of a set (often a subset of a Cartesian product) in two different ways.

Proposition 2.2. ***The Captain's Identity.***
For $n, k \in \mathbb{N}^*$:

$$k\binom{n}{k} = n\binom{n-1}{k-1}.$$

命題

Proof

Let E be a set of size n. Consider the set of pairs consisting of a committee of size k and a designated chairperson from within that committee:

$$C = \{(x, S) : S \subseteq E, |S| = k, x \in S\}.$$

We compute $|C|$ in two ways:

Choose the committee, then the chair. There are $\binom{n}{k}$ ways to choose the set S. Once S is chosen, there are k choices for $x \in S$. Thus $|C| = k\binom{n}{k}$.

Choose the chair, then the rest of the committee. There are n ways to choose the element $x \in E$. The remaining $k-1$ members of S must be chosen from $E \setminus \{x\}$, which has size $n-1$. There are $\binom{n-1}{k-1}$ ways to do this. Thus $|C| = n\binom{n-1}{k-1}$.

Equating the two expressions yields the identity.

■

Example 2.1. Mean of the Binomial Distribution. We calculate the sum $\sum_{k=0}^{n} k\binom{n}{k}$.

Using the identity $k\binom{n}{k} = n\binom{n-1}{k-1}$:

$$\begin{aligned}\sum_{k=0}^{n} k\binom{n}{k} &= \sum_{k=1}^{n} n\binom{n-1}{k-1}\\ &= n\sum_{j=0}^{n-1}\binom{n-1}{j} \quad \text{(letting } j = k-1\text{)}\\ &= n \cdot 2^{n-1}.\end{aligned}$$

This confirms that the "average" size of a subset weighted by binomial counts is $n/2$, which aligns with the symmetry of the binomial distribution.

範例

Example 2.2. Vandermonde's Convolution. Consider two disjoint sets A and B with $|A| = r$ and $|B| = m$. We wish to choose a committee of size k from $A \cup B$.
Directly, this is $\binom{r+m}{k}$. Alternatively, any such committee contains j members from A and $k-j$ members from B, for some $0 \leq j \leq k$. By the *Multiplication Principle*, for a fixed j, there are $\binom{r}{j}\binom{m}{k-j}$ such committees. Summing over j gives the identity:

$$\sum_{j=0}^{k}\binom{r}{j}\binom{m}{k-j} = \binom{r+m}{k}.$$

範例

2.2 *Multisets*

In many contexts, we wish to select elements where repetition is allowed. This gives rise to the concept of a multiset.

Definition 2.2. ***Multiset.***
Let E be a finite set. A **multiset** on E is a mapping $f : E \to \mathbb{N}$. The value $f(x)$ represents the multiplicity of element x. The **size** of the multiset is the sum of multiplicities $\sum_{x\in E} f(x)$.

定義

We often denote a multiset of size k as a "k-combination with repetition". If $E = \{x_1, \ldots, x_n\}$, a multiset of size k corresponds to a solution to the Diophantine equation:

$$a_1 + a_2 + \cdots + a_n = k, \quad a_i \in \mathbb{N},$$

where $a_i = f(x_i)$.

Theorem 2.4. ***Multiset Counting*.**
The number of multisets of size k from a set of n elements is:

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k}.$$

定理

Proof

Let $E = \{1, \dots, n\}$. A multiset of size k can be represented as a non-decreasing sequence of integers:

$$1 \le x_1 \le x_2 \le \cdots \le x_k \le n.$$

We map this sequence to a strictly increasing sequence $y_1 < y_2 < \cdots < y_k$ defined by:

$$y_i = x_i + (i-1).$$

Since $1 \le x_1$ and $x_k \le n$, we have:

$$1 \le y_1 < y_2 < \cdots < y_k \le n + k - 1.$$

This transformation is a bijection between the set of multisets on E of size k and the set of subsets of $\{1, \dots, n + k - 1\}$ of size k. The cardinality of the latter is $\binom{n+k-1}{k}$.

■

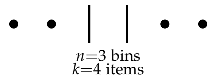

Figure 2.2: The "Stars and Bars" method. The configuration • • || • • corresponds to the solution $x_1 = 2, x_2 = 0, x_3 = 2$. There are k stars and $n-1$ bars.

Example 2.3. Integer Solutions. Find the number of non-negative integer solutions to $x_1 + x_2 + x_3 + x_4 = 10$.
Here $n = 4$ (the number of variables) and $k = 10$. By *theorem* 2.4, the number of solutions is:

$$\binom{4+10-1}{10} = \binom{13}{10} = \binom{13}{3} = \frac{13 \cdot 12 \cdot 11}{3 \cdot 2 \cdot 1} = 286.$$

範例

Example 2.4. Distributing Indistinguishable Items. Suppose we wish to distribute k indistinguishable coins into n distinguishable boxes.
This is identical to choosing a multiset of size k from the set of boxes (choosing box i means placing a coin in it). Assume $k \ge n$. If $k < n$, there are no such distributions. If we require that every box must contain at least one coin, we first place one coin in each box. We then distribute the remaining $k - n$ coins arbitrarily. The number of ways is:

$$\binom{n + (k-n) - 1}{k-n} = \binom{k-1}{k-n} = \binom{k-1}{n-1}.$$

範例

Multinomial Coefficients

The binomial coefficient $\binom{n}{k}$ counts the ways to partition a set of n elements into two disjoint subsets of sizes k and $n-k$. The **multinomial coefficient** generalizes this to partitions into r subsets.

Definition 2.3. ***Multinomial Coefficient.***
Let $n \in \mathbb{N}$ and let $k_1, \ldots, k_r$ be non-negative integers such that $\sum_{i=1}^{r} k_i = n$. The **multinomial coefficient**, denoted

$$\binom{n}{k_1, k_2, \ldots, k_r},$$

is the number of ordered r-tuples $(A_1, \ldots, A_r)$ of pairwise disjoint subsets of a set E (where $|E| = n$) such that $|A_i| = k_i$ for all $1 \leq i \leq r$.

定義

Note

If any $k_i = 0$, the corresponding subset A_i is the empty set. If $r = 2$, we recover the binomial coefficient:

$$\binom{n}{k, n-k} = \binom{n}{k}.$$

Proposition 2.3. ***Multinomial Formula.***
For $n = k_1 + \cdots + k_r$, the multinomial coefficient is given by:

$$\binom{n}{k_1, \ldots, k_r} = \frac{n!}{k_1! k_2! \cdots k_r!}.$$

命題

Proof

We proceed by induction on r. For $r = 1$, $\binom{n}{n} = \frac{n!}{n!} = 1$, which is correct.

For $r > 1$, determining the tuple $(A_1, \ldots, A_r)$ is equivalent to first choosing the subset A_r of size k_r from E, and then partitioning the remaining set $E \setminus A_r$ (of size $n - k_r$) into $r-1$ subsets of sizes $k_1, \ldots, k_{r-1}$. The number of ways to choose A_r is $\binom{n}{k_r}$. By the inductive hypothesis, the number of ways to partition the remainder is $\binom{n-k_r}{k_1, \ldots, k_{r-1}}$. Using the multiplication principle:

$$\begin{aligned}\binom{n}{k_1, \ldots, k_r} &= \binom{n}{k_r} \times \binom{n-k_r}{k_1, \ldots, k_{r-1}} \\ &= \frac{n!}{k_r!(n-k_r)!} \times \frac{(n-k_r)!}{k_1! \cdots k_{r-1}!} \\ &= \frac{n!}{k_1! \cdots k_r!}.\end{aligned}$$

■

Remark.

This quantity also counts the number of permutations of a multiset of size n containing distinct elements $x_1, \dots, x_r$ with respective multiplicities $k_1, \dots, k_r$. Geometrically, this corresponds to lattice paths in $\mathbb{Z}^r$.

Example 2.5. Lattice Paths in Higher Dimensions. Consider a particle starting at the origin $0 \in \mathbb{Z}^d$. A step consists of adding a standard basis vector e_i to the current position. How many paths of length n end at the coordinate $(a_1, \dots, a_d)$, where $\sum a_i = n$?
A path is a sequence of n steps. To land at $(a_1, \dots, a_d)$, the path must contain exactly a_1 steps in direction e_1, a_2 steps in direction e_2, and so forth. The distinct paths correspond to the distinct permutations of the multiset of steps. Thus, the number of paths is:

$$\binom{n}{a_1, \dots, a_d} = \frac{n!}{a_1! \cdots a_d!}.$$

範例

Theorem 2.5. ***The Multinomial Theorem.***
Let R be a commutative ring and $x_1, \dots, x_r \in R$. For any $n \in \mathbb{N}$:

$$(x_1 + \cdots + x_r)^n = \sum_{\substack{k_1+\cdots+k_r=n \\ k_i \geq 0}} \binom{n}{k_1, \dots, k_r} x_1^{k_1} \cdots x_r^{k_r}.$$

定理

Proof

Consider the expansion of the product $\prod_{j=1}^{n}(x_1 + \cdots + x_r)$. Each term in the expansion is formed by selecting one variable x_i from each of the n factors. If we select the variable x_1 exactly k_1 times, x_2 exactly k_2 times, and so on, we generate the monomial $x_1^{k_1} \cdots x_r^{k_r}$.
The number of ways to produce this specific monomial is the number of ways to assign the "indices" of the factors to the variables $x_1, \dots, x_r$ such that x_i receives k_i indices. This is precisely the definition of the multinomial coefficient $\binom{n}{k_1, \dots, k_r}$.

■

Example 2.6. Coefficient of a Polynomial. Determine the coefficient of $x^2y^3z^4$ in the expansion of $(x + y + z)^9$.
Here $n = 9, k_x = 2, k_y = 3, k_z = 4$. Since $2 + 3 + 4 = 9$, the term exists. The coefficient is:

$$\binom{9}{2, 3, 4} = \frac{9!}{2!3!4!} = \frac{362880}{2 \cdot 6 \cdot 24} = \frac{362880}{288} = 1260.$$

範例

2.3 *Analysis of Binomial Coefficients*

We now turn to the magnitude and "shape" of the binomial coefficients $\binom{n}{k}$ for fixed n. Understanding these growth properties is essential for asymptotic analysis in number theory and probability.

Bounds and Estimation

While exact calculation is possible via factorials, bounds are often more useful for analysis.

> **Lemma 2.1.** ***Exponential Bound.***
> For all $x \in \mathbb{R}$, $1 + x \leq e^x$.
>
> 引理

> **Theorem 2.6.** ***Standard Bounds.***
> For $1 \leq k \leq n$:
> $$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k.$$
>
> 定理

Proof

Lower Bound: Recall $\binom{n}{k} = \frac{n}{k} \cdot \frac{n-1}{k-1} \cdots \frac{n-k+1}{1} = \prod_{i=0}^{k-1} \frac{n-i}{k-i}$. Since $n \geq k$, we have $n - i \geq k - i > 0$, and the function $f(t) = \frac{n-t}{k-t}$ is non-decreasing for $t \in [0, k)$. Thus $\frac{n-i}{k-i} \geq \frac{n}{k}$. The product satisfies:

$$\binom{n}{k} \geq \prod_{i=0}^{k-1} \frac{n}{k} = \left(\frac{n}{k}\right)^k.$$

Upper Bound: By the Binomial Theorem, for any $x > 0$:

$$(1+x)^n = \sum_{j=0}^{n} \binom{n}{j} x^j \geq \binom{n}{k} x^k.$$

Thus, $\binom{n}{k} \leq \frac{(1+x)^n}{x^k}$. If $k = n$, the bound is immediate. Assume $1 \leq k < n$ and choose $x = \frac{k}{n-k}$.

$$\binom{n}{k} \leq \frac{\left(1 + \frac{k}{n-k}\right)^n}{\left(\frac{k}{n-k}\right)^k} = \left(\frac{n}{k}\right)^k \left(1 + \frac{k}{n-k}\right)^{n-k}.$$

Using the lemma $1 + y \leq e^y$ with $y = \frac{k}{n-k}$:

$$\left(1 + \frac{k}{n-k}\right)^{n-k} \leq e^k.$$

Substituting this back yields $\binom{n}{k} \leq (\frac{n}{k})^k e^k = (\frac{en}{k})^k$.

■

Unimodality and Identities

For a fixed n, the sequence of coefficients $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$ increases to a maximum and then decreases.

Proposition 2.4. ***Unimodality.***
The sequence $k \mapsto \binom{n}{k}$ satisfies:

- Increasing for $k < \frac{n-1}{2}$.
- Maximal at $k = \lfloor \frac{n}{2} \rfloor$ and $k = \lceil \frac{n}{2} \rceil$.
- Decreasing for $k > \frac{n-1}{2}$.

命題

Proof

Consider the ratio of consecutive terms:

$$\rho_k = \frac{\binom{n}{k+1}}{\binom{n}{k}} = \frac{n!}{(k+1)!(n-k-1)!} \cdot \frac{k!(n-k)!}{n!} = \frac{n-k}{k+1}.$$

The sequence increases when $\rho_k > 1$, which corresponds to $n - k > k + 1 \iff 2k < n - 1$. Equality $\rho_k = 1$ holds if $2k = n - 1$ (only possible if n is odd).

■

The summation of binomial coefficients along different diagonals of Pascal's triangle yields the "Hockey Stick Identity".

Proposition 2.5. ***Upper Summation Identity.***
For $n \geq k \geq 0$:

$$\sum_{m=k}^{n} \binom{m}{k} = \binom{n+1}{k+1}.$$

命題

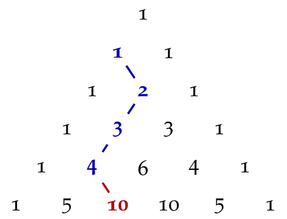

Figure 2.3: Visualisation of the Upper Summation Identity $\sum_{m=1}^{4} \binom{m}{1} = \binom{5}{2}$. The blue sum equals the red entry.

Proof

We count the number of subsets of $\{1, \dots, n+1\}$ of size $k+1$. Let this collection be $\mathcal{S}$. We know $|\mathcal{S}| = \binom{n+1}{k+1}$. Alternatively, partition $\mathcal{S}$ based on the largest element of the subset. Let $A \in \mathcal{S}$ and let $x = \max(A)$. Since $|A| = k+1$, the smallest possible maximum is $x = k+1$ (where $A = \{1, \dots, k+1\}$). The largest is $n+1$. For a fixed maximum $m+1$ (where $k \leq m \leq n$), the remaining k elements must be chosen from $\{1, \dots, m\}$. There are $\binom{m}{k}$ ways to do this. Summing over all possible values of m:

$$|\mathcal{S}| = \sum_{m=k}^{n} \binom{m}{k}.$$

■

The Principle of Inclusion-Exclusion

The Addition Principle states that $|A \cup B| = |A| + |B|$ only if A and B are disjoint. If they overlap, we must subtract the intersection. The **Sieve Formula**, or Principle of Inclusion-Exclusion (PIE), generalizes this correction to n sets.

> **Theorem 2.7.** ***The Sieve Formula.***
> Let $A_1, \ldots, A_n$ be finite subsets of a set E. Then:
>
> $$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{\varnothing \neq I \subseteq [1,n]} (-1)^{|I|+1} \left|\bigcap_{i \in I} A_i\right|.$$

定理

Proof

We utilize characteristic functions (indicator variables) introduced in the Foundations chapter.
Let f_S denote the characteristic function of a set S. The characteristic function of the union $A = \bigcup A_i$ is $1 - f_{\overline{A}}$, where $\overline{A} = \bigcap \overline{A_i}$.

$$f_{\overline{A}}(x) = \prod_{i=1}^{n} f_{\overline{A_i}}(x) = \prod_{i=1}^{n} (1 - f_{A_i}(x)).$$

Expanding this product:

$$\prod_{i=1}^{n} (1 - f_{A_i}(x)) = \sum_{I \subseteq [1,n]} (-1)^{|I|} \prod_{i \in I} f_{A_i}(x) = \sum_{I \subseteq [1,n]} (-1)^{|I|} f_{\bigcap_{i \in I} A_i}(x).$$

Summing over all $x \in E$ to convert functions to cardinalities:

$$|E \setminus A| = \sum_{I \subseteq [1,n]} (-1)^{|I|} \left|\bigcap_{i \in I} A_i\right|.$$

Isolating $|A|$ (note that the term for $I = \varnothing$ is $|E|$):

$$|E| - |A| = |E| + \sum_{\varnothing \neq I} (-1)^{|I|} \left|\bigcap_{i \in I} A_i\right|.$$

Rearranging yields the result.

■

Applications of the Sieve

The Sieve Formula is particularly effective when the intersection of sets is easier to calculate than their union.

> **Theorem 2.8.** ***Number of Surjections.***
> Let E and F be sets with $|E| = n$ and $|F| = k$. The number of surjec-

tive maps from E to F is:

$$|\mathcal{F}_{\text{surj}}(E,F)| = \sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}j^n.$$

定理

Proof

Let $F = \{1,\dots,k\}$. Let S be the set of all functions $E \to F$, so $|S| = k^n$. A function is *not* surjective if its image misses at least one element of F. Let P_i be the set of functions whose image does not contain i (i.e., range is contained in $F \setminus \{i\}$). We seek $|S| - |\bigcup_{i=1}^{k} P_i|$. For any index set $I \subseteq \{1,\dots,k\}$, the intersection $\bigcap_{i\in I} P_i$ is the set of functions mapping E into $F \setminus I$. The size of the target is $k - |I|$, so:

$$\left|\bigcap_{i\in I} P_i\right| = (k-|I|)^n.$$

By *theorem 2.7*:

$$\left|\bigcup P_i\right| = \sum_{\varnothing\neq I}(-1)^{|I|+1}(k-|I|)^n.$$

Grouping by the size of I (let $j = k - |I|$ be the size of the allowed image):

$$\left|\bigcup P_i\right| = \sum_{m=1}^{k}\binom{k}{m}(-1)^{m+1}(k-m)^n.$$

Subtracting this from total functions k^n yields the formula.

■

Example 2.7. Derangements. A **derangement** of $[1,n]$ is a permutation $\sigma \in S_n$ such that $\sigma(i) \neq i$ for all i (it has no fixed points). Let D_n be the number of derangements.
Let $A_i = \{\sigma \in S_n : \sigma(i) = i\}$. We seek $n! - |\bigcup A_i|$. The intersection of any k sets $A_{i_1} \cap \cdots \cap A_{i_k}$ fixes k specific points. The remaining $n-k$ points can be permuted arbitrarily. Thus, $|\bigcap_{i\in I} A_i| = (n-|I|)!$. Applying the Sieve:

$$|\bigcup A_i| = \sum_{k=1}^{n}\binom{n}{k}(-1)^{k+1}(n-k)!.$$

Simplifying $\binom{n}{k}(n-k)! = \frac{n!}{k!}$:

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

For large n, $D_n \approx n!/e$.

範例

2.4 *Exercises*

1. **The Gap Method.** We wish to arrange the letters of the word `MISSISSIPPI`.

   (a) Calculate the total number of distinct permutations using the Multinomial Coefficient formula.

   (b) Calculate the number of distinct permutations such that no two `I`'s are adjacent.

2. **Multinomial Coefficients and Expansions.**

   (a) Determine the coefficient of the term $x^2y^3z^4$ in the expansion of $(x - 2y + 3z)^9$.

   (b) Using the Multinomial Theorem, give a combinatorial proof that:
   $$\sum_{n_1+n_2+n_3=n} \binom{n}{n_1, n_2, n_3} (-1)^{n_2} = 1.$$

Consider the expansion of $(A + B + C)^n$ for specific values of A, B, C.

3. **Indistinguishable Groupings.** Consider a group of 12 distinct graduate students.

   (a) The students are to be assigned to 3 distinct laboratories (Lab A, Lab B, Lab C) such that Lab A receives 5 students, Lab B receives 4, and Lab C receives 3. In how many ways can this be done?

   (b) The students are to be partitioned into 3 study groups of sizes 5, 4, and 3. The study groups have no names or designations (they are indistinguishable beyond their size). In how many ways can this be done?

   (c) The students are to be paired off into 6 teams of 2 members each. The teams are indistinguishable. Calculate the number of possible partitions.

4. **Inclusion-Exclusion Logic.** In the "Hands with At Most Two Suits" example, we calculated the size of a set by subtracting the overcounted intersection (the single-suit hands). Generalize this logic to solve the following:

   A 6-digit PIN code $d_1d_2d_3d_4d_5d_6$ is constructed using digits from the set $\{1, 2, 3\}$. How many such PIN codes contain **at least one** of each digit?

5. **Constrained Integer Partitions (Investment Strategies).** We possess \$20,000 to invest across 4 distinct opportunities. Investments must be in integral units of \$1,000. Furthermore, each opportunity imposes a minimum entry threshold: the minimum investments are \$2,000, \$2,000, \$3,000, and \$4,000 respectively.

   Using the stars and bars transformation ($y_i = x_i - m_i$), determine the number of distinct investment strategies available if:

(a) An investment must be made in **every** opportunity (i.e., all minimums must be met).

(b) Investments must be made in **at least** 3 of the 4 opportunities.

For part (b), consider the mutually exclusive cases based on which opportunities are active. Be careful to sum the case of 4 active opportunities and the cases of exactly 3 active opportunities.

6. **The Chairperson Identities.** We explore identities of the form $\sum k^p \binom{n}{k}$ via double counting.

(a) **Second Moment:** Verify the identity

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

For (a): Consider the selection of a committee, a chairperson, and a secretary (who may be the same person as the chair).

(b) **Third Moment:** By considering a committee, a chairperson, a secretary, and a treasurer (where roles may overlap), argue combinatorially that:

$$\sum_{k=1}^{n} k^3 \binom{n}{k} = 2^{n-3} n^2(n+3).$$

7. **Inequalities and Decompositions.**

(a) **Inequality Constraints:** Determine the number of integer vectors $(x_1, \dots, x_n)$ such that $x_i \geq 0$ for all i, and

$$\sum_{i=1}^{n} x_i \leq k.$$

For (a): Introduce a "slack" variable $y \geq 0$ such that $\sum x_i + y = k$ to convert the inequality into an equality, then apply the Stars and Bars theorem.

(b) **Analytic Decomposition:** Give an algebraic verification of the identity:

$$\binom{n}{2} = \binom{k}{2} + k(n-k) + \binom{n-k}{2}, \quad \text{for } 1 \leq k \leq n.$$

Explain the combinatorial significance of this decomposition in terms of choosing 2 items from a set partitioned into two groups of size k and $n-k$.

8. **Parliamentary Deadlock.** In the parliament of a certain country, there are 151 seats filled by members of three distinct political parties.

(a) How many possible distributions of seats (n_1, n_2, n_3) are there such that $n_1 + n_2 + n_3 = 151$?

(b) A "hung parliament" occurs if no single party holds an absolute majority (strictly more than half the seats). Determine the number of distributions that result in a hung parliament.

For (b): Use the Principle of Inclusion-Exclusion or complementary counting. Note that it is impossible for two parties to simultaneously hold a majority.

9. **Triangular Sums and Double Counting.**

(a) Establish the identity $\sum_{i=1}^{n} i(n-i) = \binom{n+1}{3}$ by counting the same collection of subsets, but this time classifying them by the *middle* element (in terms of magnitude).

(b) Generalise part (b) to find a summation formula for $\binom{n+1}{2m+1}$ by considering the median element of a subset of size $2m+1$.

10. The Fibonacci Subsets. Let n be a positive integer. We wish to count the number of subsets of $\{1, 2, \ldots, n\}$ that contain no pair of consecutive integers. Let $f(n,k)$ denote the number of such subsets of size k.

(a) Construct a bijection between these non-consecutive k-subsets and the standard k-subsets of $\{1, \ldots, n-k+1\}$.

(b) Deduce that $f(n,k) = \binom{n-k+1}{k}$.

(c) Let F_n be the Fibonacci sequence defined by $F_0 = 0, F_1 = 1, F_n = F_{n-1} + F_{n-2}$. Prove that the total number of non-consecutive subsets of $\{1, \ldots, n\}$ is given by F_{n+2}.

For (a): Consider the map $x_i \mapsto x_i - (i-1)$ applied to the elements of the subset arranged in increasing order.

For (c): Sum the result from (b) and prove the sum satisfies the Fibonacci recurrence relation by classifying subsets based on whether they contain the element n.

3
Partitions and Allocations

We have previously explored compositions, where the order of summands distinguishes one configuration from another. If we relax this constraint and consider summands up to reordering, we enter the theory of integer partitions.

3.1 Integer Partitions

A composition of a natural number n is a sequence of positive integers summing to n. For instance, $(1,3,1)$ and $(3,1,1)$ are distinct compositions of 5. If we regard these as identical (distinguishing them only by the multisets of their parts), we obtain partitions.

> **Definition 3.1.** ***Integer Partition.***
> A **partition** of a positive integer n is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ of positive integers such that:
>
> $$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0 \quad \text{and} \quad \sum_{i=1}^{k} \lambda_i = n.$$
>
> The terms λ_i are called the **parts** of the partition. We define $p(n)$ as the total number of distinct partitions of n.
>
> 定義

> *Note*
>
> By convention, $p(0) = 1$ (the empty sum).

> **Example 3.1.** Partitions of Small Integers. For $n = 4$, the possible partitions are:
> - 4
> - $3+1$
> - $2+2$
> - $2+1+1$
> - $1+1+1+1$

> Thus, $p(4) = 5$. For $n = 5$, we list: $5, 4+1, 3+2, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1$. Hence $p(5) = 7$.
>
> 範例

Unlike the binomial coefficients or compositions, there is no elementary closed-form expression for $p(n)$. The growth of $p(n)$ is governed by the Hardy-Ramanujan asymptotic formula:

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right) \quad \text{as } n \to \infty.$$

While the analytic derivation of this formula is advanced, we can establish profound algebraic properties of partitions using bijective combinatorial proofs.

Odd and Distinct Parts

Consider the partitions of $n = 5$.

- Partitions into **odd parts**: 5, $3+1+1$, $1+1+1+1+1$. (Count: 3)
- Partitions into **distinct parts**: 5, $4+1$, $3+2$. (Count: 3)

This equality is not coincidental.

> **Theorem 3.1.** ***Euler's Partition Theorem.***
> For any integer $n \geq 1$, the number of partitions of n into odd parts equals the number of partitions of n into distinct parts.
>
> 定理

Let $\mathcal{O}_n$ be the set of partitions of n into odd parts, and $\mathcal{D}_n$ be the set of partitions of n into distinct parts. We construct a bijection $\phi : \mathcal{O}_n \to \mathcal{D}_n$.

> *Construction of ϕ.*
>
> Let $\lambda \in \mathcal{O}_n$. We may write the sum as $\sum_{k \text{ odd}} b_k \cdot k$, where b_k is the multiplicity of the odd part k in λ. We express each multiplicity b_k in its binary representation:
>
> $$b_k = \sum_{j=0}^{m} c_{k,j} 2^j, \quad \text{where } c_{k,j} \in \{0, 1\}.$$
>
> Substituting this back into the sum:
>
> $$n = \sum_{k \text{ odd}} \left(\sum_{j=0}^{m} c_{k,j} 2^j\right) k = \sum_{k \text{ odd}} \sum_{j : c_{k,j} = 1} (2^j \cdot k).$$
>
> The terms in this expanded sum are of the form $2^j \cdot k$. By the Fundamental Theorem of Arithmetic, every integer m can be uniquely written as $m = 2^j \cdot k$ where k is odd. Thus, all terms $2^j \cdot k$ generated are distinct. We define $\phi(\lambda)$ to be the partition consisting of these

parts $2^j \cdot k$, arranged in decreasing order. Since they sum to n and are distinct, $\phi(\lambda) \in \mathcal{D}_n$.

証明終

Construction of ϕ^{-1}.

Let $\mu \in \mathcal{D}_n$ with distinct parts $\mu_1, \ldots, \mu_r$. For each part μ_i, factor out the highest power of 2 to write $\mu_i = 2^{a_i} \cdot k_i$, where k_i is odd. The inverse map decomposes the part μ_i into 2^{a_i} copies of the odd number k_i. Collecting all such copies for all i yields a partition into odd parts. Since the binary representation is unique, ϕ is a bijection.

証明終

Example 3.2. Application of the Bijection. Consider the partition of $n = 44$ into odd parts:

$$\lambda = (13, 13, 5, 5, 5, 3).$$

Here, the multiplicities are:

- Part 13: appears 2 times. $2 = 2^1$. Terms: $2^1 \cdot 13 = 26$.
- Part 5: appears 3 times. $3 = 2^1 + 2^0$. Terms: $2^1 \cdot 5 = 10$, $2^0 \cdot 5 = 5$.
- Part 3: appears 1 time. $1 = 2^0$. Terms: $2^0 \cdot 3 = 3$.

The corresponding partition into distinct parts is $\mu = (26, 10, 5, 3)$. Sum check: $26 + 10 + 5 + 3 = 44$.

範例

Example 3.3. Reverse Mapping. Consider the partition into distinct parts $\mu = (12, 10, 2)$ for $n = 24$.

We decompose each part into odd components:

- $12 = 4 \times 3 = 2^2 \times 3 \implies$ four copies of 3.
- $10 = 2 \times 5 = 2^1 \times 5 \implies$ two copies of 5.
- $2 = 2 \times 1 = 2^1 \times 1 \implies$ two copies of 1.

The corresponding partition into odd parts is $(5, 5, 3, 3, 3, 3, 1, 1)$.

範例

Ferrers Diagrams

A graphical representation of partitions, known as the Ferrers diagram, facilitates the proof of complex identities via geometric transformations.

Definition 3.2. ***Ferrers Diagram.***
The **Ferrers diagram** of a partition $\lambda = (\lambda_1, \ldots, \lambda_k)$ is a collection of dots left-justified in k rows, such that the i-th row contains λ_i dots.

定義

For example, the diagram for $\lambda = (7, 5, 5, 3, 1)$ is:

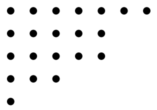

Figure 3.1: Ferrers diagram for $(7, 5, 5, 3, 1)$.

By reflecting the diagram across the main diagonal (swapping rows and columns), we obtain a new partition.

Definition 3.3. ***Conjugate Partition.***
The **conjugate** of a partition λ, denoted λ', is the partition defined by the column lengths of the Ferrers diagram of λ. Formally, λ'_j is the number of parts of λ that are greater than or equal to j.

定義

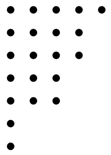

Figure 3.2: The conjugate partition $\lambda' = (5, 4, 4, 3, 3, 1, 1)$.

The conjugate of $(7, 5, 5, 3, 1)$ shown in *figure 3.1* is $(5, 4, 4, 3, 3, 1, 1)$, shown in *figure 3.2*. Since the reflection is an involution, the conjugate map is a bijection on the set of partitions of n. This symmetry yields several identities immediately.

Proposition 3.1. ***Conjugate Identities.***

1. For $n \geq k \geq 1$, the number of partitions of n into at least k parts is equal to the number of partitions of n in which the largest part is at least k.
2. For $n \geq 1$, the number of partitions of n in which the first two parts are equal is equal to the number of partitions of n in which all parts are at least 2.

命題

Proof

We rely on the geometric properties of the Ferrers diagram under conjugation.

1. Let λ be a partition. The number of parts is the number of rows in its diagram, which becomes the number of columns in the conjugate λ'. The largest part of λ' is the length of its first row, which corresponds to the length of the first column of λ (the number of rows). Thus, λ has $\geq \quad k$ parts if and only if λ' has largest part $\geq k$.
2. A partition λ has all parts $\geq \quad 2$ if and only if every row in its diagram has length at least 2. Under conjugation, the number of rows of length at least 1 is λ'_1 and the number of rows of length at least 2 is λ'_2. Thus all parts $\geq$ 2 if and only if $\lambda'_1 = \lambda'_2$. This is exactly the condition that the first two parts are equal. Thus the map $\lambda \mapsto \lambda'$ is a bijection between the set of partitions with all

> parts ≥ 2 and the set of partitions with the first two parts equal.
>
> ■

The Twelvefold Way

We now this section by synthesising our study of counting into the "Twelvefold Way", a classification attributed to Gian-Carlo Rota. This framework considers the number of ways to place balls into boxes under various conditions.
Let N be a set of balls ($|N| = n$) and K a set of boxes ($|K| = k$). We enumerate the functions $f : N \to K$ (or equivalent structures) based on:

Distinguishability : Are the balls distinguishable? Are the boxes distinguishable?

Restrictions :

- **None**: Any number of balls per box.
- **Injective**: At most one ball per box (requires $n \leq k$).
- **Surjective**: At least one ball per box (requires $n \geq k$).

Equivalence of functions is defined by the distinguishability. For example, if balls are indistinguishable, functions f and g are equivalent if $g = f \circ \sigma$ for some permutation σ of N.
We denote the scenarios by pairs (Balls, Boxes).

Distinguishable Balls, Distinguishable Boxes

Here we count standard functions $f : N \to K$.

No restriction. Each of the n balls can be placed in any of the k boxes independently.

$$|K^N| = k^n.$$

Injective. This corresponds to permutations of subsets. We place balls sequentially into distinct boxes.

$$k^{\underline{n}} = \frac{k!}{(k-n)!}.$$

Surjective. We partition the n balls into exactly k non-empty subsets (fibres), then assign these subsets to the k distinct boxes. Using the Stirling numbers of the second kind:

$$k!S(n,k).$$

Indistinguishable Balls, Distinguishable Boxes

Because the balls are identical, only the *count* of balls in each box matters. This is the problem of weak compositions or multisets.

No restriction. We seek integer solutions to $x_1 + \cdots + x_k = n$ with $x_i \geq 0$. By *theorem* 2.4:

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}.$$

Injective. Each box receives 0 or 1 ball. We simply choose which n boxes contain a ball.

$$\binom{k}{n}.$$

Surjective. We seek solutions to $x_1 + \cdots + x_k = n$ with $x_i \geq 1$. This is the number of compositions of n into k parts. By distributing one ball to each box initially, we distribute the remaining $n-k$ balls freely:

$$\binom{(n-k)+k-1}{k-1} = \binom{n-1}{k-1}.$$

Distinguishable Balls, Indistinguishable Boxes

Since boxes are indistinguishable, the specific assignment of values does not matter, only the partitioning of the domain N.

Surjective. This is the definition of the Stirling numbers of the second kind. We partition n items into k non-empty sets.

$$S(n,k).$$

No restriction. We partition the balls into j non-empty sets, where $1 \leq j \leq k$.

$$\sum_{j=1}^{k} S(n,j).$$

Injective. If $n \leq k$, we place each ball in a separate box. Since boxes are indistinguishable, there is only 1 way to do this. If $n > k$, it is impossible (0).

Indistinguishable Balls, Indistinguishable Boxes

Here we partition the integer n (the total number of balls) into parts defined by the box contents.

Surjective. We partition n into exactly k parts (no empty boxes allowed). We denote this by $p(n,k)$.

$$p(n,k).$$

No restriction. We partition n into at most k parts (some boxes may be empty).

$$\sum_{j=1}^{k} p(n,j).$$

> *Note*
>
> By *proposition 3.1*, this is equal to the number of partitions of n where the largest part is at most k.

Injective. As with distinguishable balls, if $n \leq k$, we place one ball in each of n boxes; indistinguishability makes this unique.

$$\begin{cases} 1 & \text{if } n \leq k, \\ 0 & \text{if } n > k. \end{cases}$$

> **Example 3.4.** Server Load Balancing. Consider a system with $n = 4$ jobs and $k = 3$ servers.
>
> 1. If jobs and servers are distinguishable (e.g., jobs are unique tasks, servers have different hardware):
>   - Total assignments: $3^4 = 81$.
>   - Surjective (all servers active): $3!S(4,3) = 6 \times \binom{4}{2} = 6 \times 6 = 36$.
> 2. If jobs are identical (standard computational units) but servers are distinguishable:
>   - Total assignments: $\binom{4+3-1}{2} = \binom{6}{2} = 15$.
> 3. If jobs are distinct but servers are identical (we only care about which jobs are grouped):
>   - Surjective: $S(4,3) = 6$. The grouping is of type $\{2,1,1\}$.
> 4. If both are indistinguishable (we only care about load distribution):
>   - Partitions of 4 into at most 3 parts: 4 (one server), $3 \; + \; 1$ (two servers), $2+2$ (two servers), $2+1+1$ (three servers). Total = 4.
>
> 範例

3.2 *Generating Functions*

The objective of this section is to translate combinatorial problems regarding sequences into algebraic problems regarding formal power series. By encoding a sequence $(a_n)_{n\in\mathbb{N}}$ as the coefficients of a series $A(X) = \sum a_n X^n$, we can determine properties of the sequence—such as closed forms or asymptotic behaviour—by manipulating the function $A(X)$ within a ring structure.

3.3 The Ring of Formal Power Series

To define formal power series rigorously, we first recall the structure of polynomials. Let $\mathcal{A}$ be a commutative ring (typically $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$).

The ring of polynomials $\mathcal{A}[X]$ consists of sequences $P = (p_n)_{n\in\mathbb{N}}$ with coefficients in $\mathcal{A}$ that have **finite support**; that is, there exists some N such that $p_n = 0$ for all $n > N$. We conventionally write such a sequence as a sum:

$$P(X) = \sum_{n=0}^{N} p_n X^n.$$

The operations of addition and multiplication in $\mathcal{A}[X]$ are defined to satisfy the laws of commutativity, associativity, and distributivity, distinguishing $\mathcal{A}[X]$ as a sub-ring of the structure we are about to define.

Formal Power Series

We extend the concept of a polynomial by removing the restriction of finite support.

> **Definition 3.4.** *Formal Power Series.*
> Let $\mathcal{A}$ be a commutative ring. A **formal power series** with coefficients in $\mathcal{A}$ is a sequence $(a_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{A}$. We denote this series by the formal sum:
>
> $$A(X) = \sum_{n=0}^{\infty} a_n X^n = a_0 + a_1 X + a_2 X^2 + \dots$$
>
> The set of all such series is denoted by $\mathcal{A}[[X]]$.

定義

> *Remark.*
>
> The variable X is strictly a placeholder; it serves to index the positions of the coefficients. Unlike in analysis, we do not assign a numerical value to X, and issues of convergence do not arise. The series is a purely algebraic object defined by its sequence of coefficients.

We equip $\mathcal{A}[[X]]$ with algebraic operations analogous to those for polynomials.

> **Definition 3.5.** *Operations.*
> Let $A(X) = \sum_{n\geq 0} a_n X^n$ and $B(X) = \sum_{n\geq 0} b_n X^n$.
> 1. **Sum:** The sum $A(X) + B(X)$ is the series whose n-th coefficient is

the sum of the n-th coefficients of A and B:

$$A(X) + B(X) = \sum_{n \geq 0} (a_n + b_n) X^n.$$

2. **Cauchy Product:** The product $A(X) \cdot B(X)$ is the series defined by the convolution of the sequences:

$$A(X) \cdot B(X) = \sum_{n \geq 0} c_n X^n, \quad \text{where } c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

定義

The definition of the product is motivated by the distributive expansion of polynomials: the term X^n in the product arises from pairing $a_k X^k$ with $b_{n-k} X^{n-k}$ for all possible k.

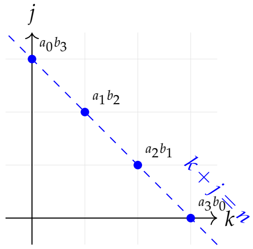

Figure 3.3: The coefficient c_n of the Cauchy product sums all terms $a_k b_j$ where the indices sum to n. Here $n = 3$.

Example 3.5. Convolution of the Constant Sequence. Let $A(X) = \sum_{n \geq 0} X^n$. This corresponds to the sequence $(1, 1, 1, \dots)$.
Consider the square $C(X) = A(X)^2$. By the definition of the Cauchy product, the coefficient c_n is:

$$c_n = \sum_{k=0}^{n} a_k a_{n-k} = \sum_{k=0}^{n} 1 \cdot 1 = \sum_{k=0}^{n} 1 = n + 1.$$

Thus:

$$\left(\sum_{n \geq 0} X^n \right)^2 = \sum_{n \geq 0} (n+1) X^n = 1 + 2X + 3X^2 + \dots$$

Combinatorially, c_n counts the number of ways to write n as a sum of two non-negative integers (where order matters).

範例

Theorem 3.2. ***Ring Structure.***
The set $\mathcal{A}[[X]]$ equipped with the sum and Cauchy product forms a commutative ring.
- The additive identity is the zero series $0 = (0, 0, \dots)$.
- The multiplicative identity is the unity series $1 = (1, 0, 0, \dots)$.

Furthermore, $\mathcal{A}[X]$ is a subring of $\mathcal{A}[[X]]$.

定理

Proof

Let $A(X), B(X), C(X) \in \mathcal{A}[[X]]$ with coefficients (a_n), (b_n), (c_n) in $\mathcal{A}$. By the definitions of coefficientwise addition and the Cauchy product above, associativity, commutativity, the additive identity, and additive inverses hold termwise because they hold in $\mathcal{A}$. For multiplication, the coefficient of X^n in $(AB)C$ equals

$$\sum_{i=0}^{n} \left(\sum_{k=0}^{i} a_k b_{i-k} \right) c_{n-i},$$

while the coefficient of X^n in $A(BC)$ equals

$$\sum_{k=0}^{n} a_k \left(\sum_{j=0}^{n-k} b_j c_{n-k-j} \right).$$

Reindexing with $j = i - k$ matches the two sums, so the product is associative. Commutativity and distributivity follow by the same coefficient checks. The unity series has constant term 1 and all other coefficients 0, so it is a multiplicative identity. Finally, a polynomial is exactly a series with finite support, so $\mathcal{A}[X]$ is closed under these operations and is a subring.

■

Invertibility

In the ring of polynomials $\mathcal{A}[X]$, very few elements have multiplicative inverses (typically only the constant polynomials that are units in $\mathcal{A}$). In contrast, the ring of formal power series $\mathcal{A}[[X]]$ allows us to invert a much broader class of elements.

Theorem 3.3. ***Invertibility Criterion.***
A formal power series $A(X) = \sum_{n\geq 0} a_n X^n \in \mathcal{A}[[X]]$ is invertible if and only if its constant term a_0 is invertible in the coefficient ring $\mathcal{A}$. If this condition holds, the inverse is unique.

定理

$(\implies)$

Suppose $A(X)$ has an inverse $B(X) = \sum b_n X^n$ such that $A(X)B(X) = 1$. The constant term of the product is given by the Cauchy formula for $n = 0$:

$$c_0 = a_0 b_0.$$

Since $A(X)B(X) = 1$, we must have $c_0 = 1$, so $a_0 b_0 = 1$. Thus a_0 is invertible in $\mathcal{A}$.

証明終

$(\impliedby)$

Suppose a_0 is invertible. We seek a sequence (b_n) satisfying $A(X)B(X) = 1$. This yields the system of linear equations:

$$\sum_{k=0}^{n} a_k b_{n-k} = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \geq 1. \end{cases}$$

For $n = 0$, we have $a_0 b_0 = 1$, which determines $b_0 = a_0^{-1}$. For $n \geq 1$, we isolate the term involving b_n (which occurs when $k = 0$):

$$a_0 b_n + \sum_{k=1}^{n} a_k b_{n-k} = 0.$$

Since a_0 is invertible, we can uniquely solve for b_n in terms of the preceding coefficients $b_0, \ldots, b_{n-1}$:

$$b_n = -a_0^{-1} \sum_{k=1}^{n} a_k b_{n-k}.$$

By induction, the sequence (b_n) is uniquely determined.

証明終

Example 3.6. Inverting a Polynomial. Consider the polynomial $P(X) = 1 - X - X^2 \in \mathbb{Z}[[X]]$. Here $a_0 = 1$, which is invertible in $\mathbb{Z}$, so $P(X)^{-1}$ exists. Let $B(X) = \sum b_n X^n$ be the inverse.
Using the recurrence $b_n = -a_0^{-1} \sum_{k=1}^{n} a_k b_{n-k}$ derived in the proof:

- $b_0 = 1^{-1} = 1$.
- $b_1 = -1(a_1 b_0) = -1(-1 \cdot 1) = 1$.
- $b_2 = -1(a_1 b_1 + a_2 b_0) = -1(-1 \cdot 1 + -1 \cdot 1) = 2$.
- $b_3 = -1(a_1 b_2 + a_2 b_1 + a_3 b_0) = -1(-1 \cdot 2 + -1 \cdot 1 + 0) = 3$.
- $b_4 = -1(a_1 b_3 + a_2 b_2) = -1(-1 \cdot 3 + -1 \cdot 2) = 5$.

The coefficients $1, 1, 2, 3, 5, \ldots$ are the Fibonacci numbers. Thus, the generating function for the Fibonacci sequence is $(1 - X - X^2)^{-1}$.

範例

The most fundamental inverse is that of the linear polynomial $1 - \alpha X$. This yields the geometric series formula, which is valid formally even without convergence considerations.

Corollary 3.1. *Geometric Series.* For any $\alpha \in \mathcal{A}$, the series $1 - \alpha X$ is invertible, and:

$$\frac{1}{1 - \alpha X} = \sum_{n=0}^{\infty} \alpha^n X^n = 1 + \alpha X + \alpha^2 X^2 + \ldots$$

推論

Proof

We verify that the product is unity.

$$\begin{aligned}(1-\alpha X)\sum_{n=0}^{\infty}\alpha^n X^n &= \sum_{n=0}^{\infty}\alpha^n X^n - \sum_{n=0}^{\infty}\alpha\cdot\alpha^n X^{n+1}\\ &= \sum_{n=0}^{\infty}\alpha^n X^n - \sum_{m=1}^{\infty}\alpha^m X^m \quad (\text{letting } m=n+1)\\ &= 1+\sum_{n=1}^{\infty}(\alpha^n-\alpha^n)X^n\\ &= 1.\end{aligned}$$

■

3.4 *Generalised Binomial Theorem*

The algebraic nature of formal power series allows us to extend this to negative integer exponents.

Definition 3.6. ***Generalised Binomial Coefficient.***
For any $r \in \mathbb{R}$ and $k \in \mathbb{N}$, we define the **generalised binomial coefficient** $\binom{r}{k}$ by:

$$\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!}.$$

If $k=0$, we define $\binom{r}{0}=1$ (following the convention for empty products).

定義

This definition recovers the standard coefficient when $r \in \mathbb{N}$. For negative integers, it exhibits a regular sign-alternating pattern related to multiset counting.

Example 3.7. Negative Integers.

- For $r=-1$:

$$\binom{-1}{k} = \frac{(-1)(-2)\cdots(-k)}{k!} = \frac{(-1)^k k!}{k!} = (-1)^k.$$

- For $r=-m$ where $m \in \mathbb{N}$:

$$\begin{aligned}\binom{-m}{k} &= \frac{(-m)(-m-1)\cdots(-m-k+1)}{k!}\\ &= (-1)^k\frac{m(m+1)\cdots(m+k-1)}{k!}\\ &= (-1)^k\binom{m+k-1}{k}.\end{aligned}$$

範例

Note

This identity connects negative binomial coefficients to multiset coefficients, often denoted $\left(\!\binom{m}{k}\!\right) = \binom{m+k-1}{k}$.

Theorem 3.4. ***Generalised Binomial Theorem.***
For any integer $r \in \mathbb{Z}$, the formal power series $(1+X)^r$ is the generating function for the sequence of coefficients $\binom{r}{k}$:

$$(1+X)^r = \sum_{k=0}^{\infty} \binom{r}{k} X^k.$$

定理

Proof

If $r \geq 0$, this is *theorem* 2.3. Let $r = -m$ with $m \in \mathbb{N}$. By *corollary* 3.1 with $\alpha = -1$,

$$(1+X)^{-1} = \sum_{n=0}^{\infty} (-1)^n X^n.$$

Hence

$$(1+X)^{-m} = \left(\sum_{n=0}^{\infty} (-1)^n X^n \right)^m.$$

The coefficient of X^k in this product equals $(-1)^k$ times the number of m-tuples of non-negative integers summing to k, which is $\binom{m+k-1}{k}$ by *theorem* 2.4. By the identity for negative integers given above, this coefficient is $\binom{-m}{k}$.

■

Example 3.8. Vandermonde's Identity. This is the same identity as the Vandermonde's Convolution example in the previous chapter.

範例

Proof

By *theorem* 3.4,

$$(1+X)^r(1+X)^s = (1+X)^{r+s}.$$

The coefficient of X^n on the left is

$$\sum_{k=0}^{n} \binom{r}{k} \binom{s}{n-k},$$

and the coefficient of X^n on the right is $\binom{r+s}{n}$.

■

A particularly useful case for *theorem* 3.4 arises for negative integer exponents, relating to multiset counting.

Corollary 3.2. *Negative Binomial Expansion.* For $m \in \mathbb{N}$ and $\beta \in \mathbb{C}$:

$$(1-\beta X)^{-m} = \sum_{n=0}^{\infty} \binom{n+m-1}{m-1} \beta^n X^n.$$

推論

Proof

We expand $(1 + (-\beta X))^{-m}$. The coefficient of X^n is:

$$\binom{-m}{n}(-\beta)^n = \frac{(-m)(-m-1)\cdots(-m-n+1)}{n!}(-1)^n \beta^n.$$

Factoring $(-1)^n$ from the numerator:

$$(-1)^n \frac{m(m+1)\cdots(m+n-1)}{n!}(-1)^n \beta^n = \binom{m+n-1}{n}\beta^n.$$

Using the symmetry $\binom{N}{k} = \binom{N}{N-k}$, we have $\binom{m+n-1}{n} = \binom{m+n-1}{m-1}$. ■

Example 3.9. Convolution via Binomials. Consider the product of two geometric series derivatives.
We wish to find the coefficient of X^n in $(1-X)^{-2} \cdot (1-X)^{-2} = (1-X)^{-4}$. Directly using the corollary with $m = 4, \beta = 1$:

$$[X^n](1-X)^{-4} = \binom{n+3}{3}.$$

Alternatively, this is the convolution of $(n+1)$ with itself: $\sum_{k=0}^{n}(k+1)(n-k+1)$. The identity $\sum_{k=0}^{n}(k+1)(n-k+1) = \binom{n+3}{3}$ is thus established algebraically.

範例

3.5 *Linear Recurrence Relations*

We now apply formal power series to solve linear recurrence relations of the form:

$$a_{n+k} = c_1 a_{n+k-1} + \cdots + c_k a_n + f(n),$$

where c_i are constants and $f(n)$ is a known term.

Rational Generating Functions

We first treat the homogeneous case ($f(n) = 0$).

Theorem 3.5. ***Rational Generating Functions.***
Let (a_n) be a sequence satisfying the recurrence $a_n = \sum_{j=1}^{k} c_j a_{n-j}$ for $n \geq k$. The generating function $A(X) = \sum a_n X^n$ is a rational function of the form:

$$A(X) = \frac{P(X)}{Q(X)},$$

where $Q(X) = 1 - \sum_{j=1}^{k} c_j X^j$ and $P(X)$ is a polynomial of degree at most $k-1$ determined by the initial conditions.

定理

Proof

We multiply the recurrence $a_n - \sum_{j=1}^{k} c_j a_{n-j} = 0$ by X^n and sum over $n \geq k$:

$$\sum_{n\geq k} a_n X^n - \sum_{j=1}^{k} c_j X^j \sum_{n\geq k} a_{n-j} X^{n-j} = 0.$$

Let $A(X) = \sum_{n\geq 0} a_n X^n$. We can rewrite the sums as:

$$(A(X) - A_{k-1}(X)) - \sum_{j=1}^{k} c_j X^j (A(X) - A_{k-j-1}(X)) = 0,$$

where $A_m(X) = \sum_{i=0}^{m} a_i X^i$ is the truncated series. Rearranging terms to isolate $A(X)$:

$$A(X)\left(1 - \sum_{j=1}^{k} c_j X^j\right) = A_{k-1}(X) - \sum_{j=1}^{k} c_j X^j A_{k-j-1}(X).$$

The right-hand side is a finite sum of polynomials of degree less than k, hence a polynomial $P(X)$ of degree at most $k-1$. The term in the bracket is $Q(X)$. Thus $A(X) = P(X)/Q(X)$.

■

To extract the coefficients a_n from $P(X)/Q(X)$, we employ partial fraction decomposition. We require the following lemma for the ring $\mathbb{C}[X]$.

Lemma 3.1. ***Bezout's Identity for Polynomials.***
Let $f(X), g(X) \in \mathbb{C}[X]$ be non-zero polynomials with no common factors. There exist polynomials $u(X), v(X) \in \mathbb{C}[X]$ such that:

$$u(X)f(X) + v(X)g(X) = 1.$$

引理

Proof

This follows from the Euclidean algorithm for polynomials. Since $\mathbb{C}[X]$ is a Euclidean domain, the greatest common divisor can be expressed as a linear combination of the elements. Since f and g are coprime, $\gcd(f, g) = 1$.

■

Proposition 3.2. ***Partial Fraction Decomposition.***
Let $P(X)/Q(X)$ be a rational function with $\deg(P) < \deg(Q)$. Factor $Q(X)$ over $\mathbb{C}$ as:

$$Q(X) = \prod_{i=1}^{r} (1 - \alpha_i X)^{m_i},$$

where α_i are distinct non-zero complex numbers. Then there exist unique constants A_{ij} such that:

$$\frac{P(X)}{Q(X)} = \sum_{i=1}^{r} \sum_{j=1}^{m_i} \frac{A_{ij}}{(1 - \alpha_i X)^j}.$$

命題

Proof

We proceed by induction on the number of distinct factors. Suppose $Q(X) = Q_1(X)Q_2(X)$ where Q_1, Q_2 are coprime. By Bezout's Identity, there exist u, v such that $uQ_1 + vQ_2 = 1$. Multiply by P/Q:

$$\frac{P}{Q_1 Q_2} = \frac{P(uQ_1 + vQ_2)}{Q_1 Q_2} = \frac{Pu}{Q_2} + \frac{Pv}{Q_1}.$$

By polynomial division, we can reduce the numerators so that the degree condition is satisfied. Repeating this separates all distinct factors $(1 - \alpha_i X)^{m_i}$. It remains to decompose a term like $R(X)/(1 - \alpha X)^m$. Since the polynomials $(1 - \alpha X)^j$ for $j = 0, \ldots, m-1$ form a basis for polynomials of degree $< m$, we can expand $R(X)$ in this basis, yielding the inner sum.

■

Combining these results yields the explicit solution for any linear recurrence.

Theorem 3.6. ***Explicit Solution.***
For a sequence defined by a linear recurrence with characteristic denominator $Q(X) = \prod_{i=1}^{r} (1 - \alpha_i X)^{m_i}$, the n-th term is given by:

$$a_n = \sum_{i=1}^{r} \sum_{j=1}^{m_i} A_{ij} \binom{n + j - 1}{j - 1} \alpha_i^n.$$

定理

Proof

Apply the linear operator $[X^n]$ (coefficient extraction) to the partial fraction decomposition of $A(X)$. From *corollary* 3.1 and *theorem* 3.4:

$$[X^n]\frac{1}{(1-\alpha_i X)^j} = \binom{n+j-1}{j-1}\alpha_i^n.$$

Summing these contributions gives the result.

■

Corollary 3.3. *Asymptotic Behaviour*. Let α_1 be the factor with largest modulus among the α_i in $Q(X) = \prod_{i=1}^{r}(1-\alpha_i X)^{m_i}$ (equivalently, $1/\alpha_1$ is the root of Q with smallest modulus). If α_1 is unique and has multiplicity m_1, then as $n \to \infty$:

$$a_n \sim \frac{A_{1m_1}}{(m_1-1)!} n^{m_1-1}\alpha_1^n.$$

推論

Proof

The term with the largest base α_i dominates the sum. Among terms with the same base, the term with the highest power of n (from the binomial coefficient $\binom{n+m-1}{m-1} \approx \frac{n^{m-1}}{(m-1)!}$) dominates.

■

Example 3.10. Repeated Roots. Consider the recurrence $a_n = 4a_{n-1} - 4a_{n-2}$ for $n \geq 2$, with $a_0 = 1, a_1 = 4$.
The generating function denominator is $Q(X) = 1 - 4X + 4X^2 = (1-2X)^2$. Using the formula derived in *theorem* 3.5:

$$A(X) = \frac{a_0 + (a_1 - 4a_0)X}{1 - 4X + 4X^2} = \frac{1}{(1-2X)^2}.$$

Here, $\alpha_1 = 2$ with multiplicity $m_1 = 2$. Using the expansion for negative powers:

$$a_n = \binom{n+1}{1}2^n = (n+1)2^n.$$

Check: $a_2 = 3 \cdot 4 = 12$. Recurrence: $4(4) - 4(1) = 12$. Correct.

範例

Non-Homogeneous Recurrences

If the recurrence contains a polynomial term $f(n)$, we can solve it by incorporating the generating function for $f(n)$.

Lemma 3.2. ***Stirling Number Identity.***

For any integer $\ell \geq 0$:

$$\sum_{n=0}^{\infty} n^{\ell} X^n = \sum_{k=0}^{\ell} k! S(\ell,k) \frac{X^k}{(1-X)^{k+1}},$$

where $S(\ell,k)$ are the Stirling numbers of the second kind.

引理

Proof

We differentiate the geometric series. Let $D = X\frac{d}{dX}$. Then $D(X^n) = nX^n$, so $n^{\ell}X^n = D^{\ell}(X^n)$. We apply D^{ℓ} to $(1-X)^{-1}$. Using the identity $n^{\ell} = \sum_{k=0}^{\ell} S(\ell,k)(n)_k$, we can express the operator D^{ℓ} in terms of derivatives. Explicitly,

$$\sum_{n\geq 0} n^{\ell}X^n = \sum_{n\geq 0}\sum_{k=0}^{\ell} S(\ell,k)(n)_k X^n = \sum_{k=0}^{\ell} S(\ell,k) \sum_{n\geq 0} (n)_k X^n.$$

Note that $\sum (n)_k X^n = X^k \frac{d^k}{dX^k} \sum X^n = X^k k!(1-X)^{-(k+1)}$. Substituting this back yields the result.

■

This lemma ensures that if $f(n)$ is a polynomial, its generating function is rational. The method of partial fractions then applies to the sum $P/Q + F(X)$.

Example 3.11. A Non-Homogeneous Example. Solve $a_n = 2a_{n-1} + n$ for $n \geq 1$ with $a_0 = 0$.
Multiply by X^n and sum:

$$A(X) = 2XA(X) + \sum_{n\geq 1} nX^n.$$

Using the lemma for $\ell = 1$ (or simply differentiating $\sum X^n$):

$$\sum nX^n = \frac{X}{(1-X)^2}.$$

Thus:

$$A(X)(1-2X) = \frac{X}{(1-X)^2} \implies A(X) = \frac{X}{(1-2X)(1-X)^2}.$$

Partial fraction decomposition form:

$$\frac{X}{(1-2X)(1-X)^2} = \frac{A}{1-2X} + \frac{B}{1-X} + \frac{C}{(1-X)^2}.$$

Solving for constants yields $A = 2, B = -1, C = -1$.

$$a_n = 2(2^n) - 1(1^n) - \binom{n+1}{1}(1^n) = 2^{n+1} - 1 - (n+1) = 2^{n+1} - n - 2.$$

範例

3.6 Exercises

1. **Recurrences for Polynomials.** The Chebyshev polynomials $T_n(x)$ are defined by the recurrence $T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$ for $n \geq 2$, with $T_0(x) = 1$ and $T_1(x) = x$.

   (a) Determine the generating function $F(z) = \sum_{n=0}^{\infty} T_n(x)z^n$.

   (b) Using the generating function, prove the explicit formula:

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} x^{n-2k}(x^2-1)^k.$$

For (b): Alternatively, relate the generating function roots to $\cos(n\theta)$.

2. **Fibonacci Sum Identities.** Let F_n denote the Fibonacci numbers ($F_0 = 0, F_1 = 1$). Use the generating function $\mathcal{F}(z) = \frac{z}{1-z-z^2}$ to prove the following identities:

   (a) $\sum_{k=0}^{n} F_k = F_{n+2} - 1$.

   (b) $\sum_{k=0}^{n} F_{2k} = F_{2n+1} - 1$.

   (c) $\sum_{k=0}^{n} F_k F_{n-k} = \frac{nF_{n+1}+2(n+1)F_n}{5}$ (Hint: Differentiate).

3. **Domino Tilings.** Let A_n be the number of ways to tile a $2 \times n$ rectangle with 1×2 dominoes.

   (a) Establish the recurrence $A_n = A_{n-1} + A_{n-2}$.

   (b) Find the generating function $A(z) = \sum A_n z^n$.

   (c) Deduce the closed form for A_n.

4. **Ternary Words.** Let $f(n)$ be the number of words of length n over the alphabet $\{0, 1, 2\}$ that contain no adjacent zeros (i.e., "00" is forbidden).

   (a) Show that $f(n)$ satisfies the recurrence $f(n) = 2f(n-1) + 2f(n-2)$ for $n \geq 2$.

   (b) Compute the generating function for $f(n)$ and use it to find an explicit formula involving $\sqrt{3}$.

5. **Reciprocal Convolution.** Compute the sum

$$S_n = \sum_{0<k<n} \frac{1}{k(n-k)}.$$

Consider the coefficient of z^n in the square of the generating function for harmonic numbers, or use partial fraction decomposition on the term $\frac{1}{k(n-k)}$.

6. **Systems of Recurrences.** Let A_n be the number of ways to tile a $3 \times n$ rectangle with 1×2 dominoes. Let B_n be the number of tilings of a $3 \times n$ rectangle with one corner missing (a shape of area $3n - 1$).

(a) Prove the coupled recurrences:

$$\begin{aligned} A_n &= A_{n-2} + 2B_{n-1} \\ B_n &= A_{n-1} + B_{n-2} \end{aligned}$$

(b) Solve this system using generating functions to find A_n.

7. **End-to-End Evaluation.** Evaluate the sum

$$s_n = \sum_{k=0}^{n} \binom{n+k}{2k} 2^{n-k}.$$

(a) Find the generating function $S(z) = \sum_{n \geq 0} s_n z^n$.

(b) Show that $S(z)$ represents a rational function corresponding to a second-order linear recurrence.

(c) Determine the explicit formula for s_n.

8. **Harmonic Convolution.** Let $H_n = \sum_{k=1}^{n} \frac{1}{k}$ be the n-th harmonic number. Express the convolution sum

$$\sum_{k=1}^{n-1} H_k H_{n-k}$$

in terms of H_n and n.

Use the identity $\sum H_n z^n = -\frac{\ln(1-z)}{1-z}$.

9. **Partition Generating Functions.**

(a) Write down the generating function for $p_d(n)$, the number of partitions of n into distinct parts.

(b) Write down the generating function for $p_o(n)$, the number of partitions of n into odd parts.

(c) Prove Euler's Partition Theorem ($p_d(n) = p_o(n)$) by showing their generating functions are identical.

10. **Making Change.** Let c_n be the number of ways to change n pence using only 1p, 2p, and 5p coins.

(a) Write the generating function $C(z) = \sum c_n z^n$ as a rational function.

(b) Using partial fractions or series expansion, determine the asymptotic behaviour of c_n for large n. (i.e., find a quadratic polynomial $P(n)$ such that $c_n \sim P(n)$).

11. **Snake Oil Method.** Re-prove the identity

$$\sum_{k=0}^{n} F_k = F_{n+2} - 1$$

by evaluating $\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} F_k\right) z^n$ as a product of generating functions.