Combinatorics

Discrete Math, Spring 2026

Konstantin Chukharev

Graph Theory

- Graphs & digraphs
- Paths & connectivity
- Trees & spanning trees
- Bipartite graphs
- Matchings & Hall's theorem
- Planarity & coloring
- Network flows

Languages & Computation

- Alphabets & formal languages
- Regular expressions
- Finite automata (DFA, NFA)
- Pumping lemma
- Context-free grammars
- Pushdown automata
- Turing machines
- Decidability & complexity

Combinatorics & Recurrences

- Counting principles
- Permutations & combinations
- Inclusion–exclusion
- Partitions & Stirling numbers
- Generating functions
- Recurrence relations
- Asymptotic analysis

Combinatorics

"It is better to solve one problem five different ways, than to solve five problems one way."

— George Pólya

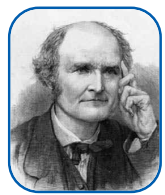

George Pólya


Abraham de Moivre


Leonhard Euler


Gottfried Wilhelm Leibniz


Herbert Wilf


Arthur Cayley


Augustin-Louis Cauchy

What is Combinatorics?

Definition 1: *Combinatorics* is the branch of mathematics that studies the existence, enumeration, and structure of *finite* discrete objects.

Three fundamental questions arise in every combinatorial problem:

Existence

Does an arrangement of the required kind *exist* at all?

Example: Does a graph with 5 vertices where every vertex has exactly 3 neighbors exist?

Enumeration

How many such arrangements are there?

Example: How many 4-digit PIN codes are possible?

Optimization

Which arrangement is *best* by some criterion?

Example: What is the shortest tour visiting all n cities exactly once?

Basic Counting Principles

"Music is the pleasure the human mind experiences from counting without being aware that it is counting."

— Gottfried Wilhelm Leibniz

Basic Counting Rules

Addition (*either/or*)

If S splits into *disjoint* parts $S_1, \ldots, S_k$, count each part separately: $|S| = |S_1| + \ldots + |S_k|$

Example: 26 vowels + 20 consonants = 46 letters.

Subtraction (complement)

Count what you *don't* want and subtract from the universe U: $\left|\overline{S}\right| = |U| - |S|$

Example: 8-bit strings *with* a leading zero: $2^8 - 2^7 = 128$.

Multiplication (*and*)

For *independent* sequential choices of sizes $n_1, \ldots, n_k$, multiply: $|S| = n_1 \cdot n_2 \cdot \ldots \cdot n_k$

Example: $26 \times 10 = 260$ two-character codes (letter then digit).

Bijection (one-to-one)

If S and T are in *bijective correspondence*, they have the same size: $|S| = |T|$

Example: $\binom{n}{k} = \binom{n}{n-k}$ via $A \mapsto [n] \setminus A$.

Addition Principle

Definition 2: We say a finite set S is *partitioned* into *parts* $S_1, ..., S_k$ if the parts are pairwise disjoint and their union is S. In other words, $S_i \cap S_j = \emptyset$ for $i \neq j$ and $S_1 \cup S_2 \cup ... \cup S_k = S$. In that case:

$$|S| = |S_1| + |S_2| + ... + |S_k|$$

Example: Let S be the set of students attending the combinatorics lecture. It can be partitioned into parts S_1 and S_2 where

$$S_1 = \text{set of students that like easy examples.}$$
$$S_2 = \text{set of students that don‘t like easy examples.}$$

If $|S_1| = 22$ and $|S_2| = 8$, then we can conclude $|S| = |S_1| + |S_2| = 30$.

Multiplication Principle

Definition 3: If S is a finite set that is the *product* of $S_1, ..., S_k$, that is, $S = S_1 \times ... \times S_k$, then

$$|S| = |S_1| \times ... \times |S_k|$$

Example: A *license plate* consists of 3 letters followed by 3 digits. The letters are chosen from $\{A, ..., Z\}$ (26 choices each) and the digits from $\{0, ..., 9\}$ (10 choices each). Since each position is filled independently:

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 = 26^3 \cdot 10^3 = 17,576,000$$

Subtraction Principle

Definition 4: Let S be a subset of a finite set T. We define the *complement* of S as $\overline{S} = T \setminus S$. Then

$$\left|\overline{S}\right| = |T| - |S|$$

Example: Let T be the set of all students at a university ($|T| = 23905$). Let $S \subseteq T$ be the set of students studying *neither* mathematics nor computer science ($|S| = 20178$). The complement $\overline{S} = T \setminus S$ consists of students studying mathematics or computer science:

$$\left|\overline{S}\right| = |T| - |S| = 23905 - 20178 = 3727$$

Bijection Principle

Definition 5: If S and T are sets, then

$$|S| = |T| \iff \text{there exists a bijection between } S \text{ and } T$$

Example: The number of k-subsets of $[n]$ equals the number of $(n-k)$-subsets: $\binom{n}{k} = \binom{n}{n-k}$. The map $S \mapsto [n] \setminus S$ is a bijection from $\binom{[n]}{k}$ to $\binom{[n]}{n-k}$: each k-subset maps to its $(n-k)$-element complement, and the map is its own inverse.

Note: The bijection principle works for *infinite* sets too — this is the foundation of Cantor's theory of cardinality.

Pigeonhole Principle

Definition 6: Let $S_1, ..., S_k$ be finite sets that are pairwise disjoint and $|S_1| + |S_2| + ... + |S_k| = n$.

$$\exists i \in \{1, ..., k\} : |S_i| \geq \left\lceil \frac{n}{k} \right\rceil \quad \text{and} \quad \exists j \in \{1, ..., k\} : |S_j| \leq \left\lfloor \frac{n}{k} \right\rfloor$$

Example: Assume there are 5 holes in the wall where pigeons nest. Say there is a set S_i of pigeons nesting in hole i. Assume there are $n = 17$ pigeons in total. Then we know:

- There is some hole with at least $d = 4$ pigeons.
- There is some hole with at most $b = 3$ pigeons.

Birthday attack: a hash function maps arbitrary inputs to n-bit digests.

- With $2^n + 1$ inputs, at least two must collide — guaranteed by the pigeonhole principle.
- With only $\approx 2^{n/2}$ random inputs, collision probability already exceeds 50%.

This is the *birthday bound* and sets the security level of hash functions in cryptography.

Double Counting

If we count the same quantity in *two different ways*, then this gives us a (perhaps non-trivial) identity.

Example (Handshaking Lemma): Assume there are n people at a party and everybody will shake hands with everybody else. How many handshakes will occur? We count this number in two ways:

1. Every person shakes $n-1$ hands and there are n people. However, two people are involved in a handshake so if we just multiply $n \cdot (n-1)$, then every handshake is counted twice. The total number of handshakes is therefore $\frac{n \cdot (n-1)}{2}$.
2. We number the people from 1 to n. To avoid counting a handshake twice, we count for person i only the handshakes with persons of lower numbers. Then the total number of handshakes is:

$$\sum_{i=1}^{n}(i-1) = \sum_{i=0}^{n-1} i = \sum_{i=1}^{n-1} i$$

The identity we obtain is therefore: $\sum_{i=1}^{n-1} i = \frac{n \cdot (n-1)}{2}$

Arrangements, Permutations, Combinations

"As for everything else, so for a mathematical theory: beauty can be perceived but not explained."

— Arthur Cayley

Ordered Arrangements

Definition 7: Denote by $[n] = \{1, ..., n\}$ the set of natural numbers from 1 to n.

Hereinafter, let X be a finite set.

Definition 8: An *ordered arrangement* of n elements of X is a *map* $s : [n] \to X$.

- Here, $[n]$ is the *domain* of s, and $s(i)$ is the *image* of $i \in [n]$ under s.
- The set $\{x \in X \mid s(i) = x \text{ for some } i \in [n]\}$ is the *range* of s.

Other common names for ordered arrangements are:

- *string* (or *word*), *e.g.* "Banana"
- *sequence*, *e.g.* "0815422372"
- *tuple*, *e.g.* $(3, 5, 2, 5, 8)$

Example:

i	1	2	3	4	5	6	7
$s(i)$	🦀	🥖	🥖	🫙	🥖	🦀	🫙

Permutations

Definition 9: A *permutation* of X is a *bijective* map $\pi : [n] \to X$.

Usually, $X = [n]$, and the set of all permutations of $[n]$ is denoted by S_n.

Example:

i	1	2	3	4	5	6	7
$\pi(i)$	2	7	1	3	5	4	6

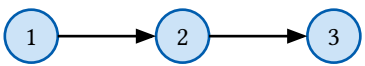
Definition 10: *k-permutation* of X is an ordered arrangement of k *distinct* elements of X, that is, an *injective* map $\pi : [k] \to X$.

The set of all k-permutations of $X = [n]$ is denoted by $P(n, k)$. In particular, $S_n = P(n, n)$.

Circular Permutations

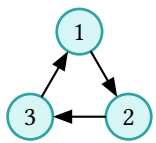
Definition 11: A *circular k-permutation* of $[n]$ arranges k distinct elements of $[n]$ in a circle. Two such arrangements are considered the same if one is a cyclic rotation of the other. The set of all circular k-permutations of $[n]$ is denoted $P_{c(n,k)}$.

Linear: all $3! = 6$ are distinct:



$(1,2,3)$, $(1,3,2)$, $(2,1,3)$,
$(2,3,1)$, $(3,1,2)$, $(3,2,1)$

Circular: rotations are identical:



$(1,2,3) \sim (2,3,1) \sim (3,1,2)$ — same circle

Only 2 distinct circles: $(1,2,3)$ and $(1,3,2)$

Counting Permutations

Theorem 1: For any natural numbers $0 \leq k \leq n$, we have

$$|P(n,k)| = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

This formula is also called the *falling factorial* and denoted $n^{\underline{k}}$ or $(n)_k$.

Proof: A permutation is an injective map $\pi : [k] \to [n]$. We count the number of ways to pick such a map, picking the images one after the other. There are n ways to choose $\pi(1)$. Given a value for $\pi(1)$, there are $(n-1)$ ways to choose $\pi(2)$ (since we may not choose $\pi(1)$ again). Continuing like this, there are $(n-i+1)$ ways to pick $\pi(i)$, and the last value we pick is $\pi(k)$ with $(n-k+1)$ possibilities.

Every k-permutation can be constructed like this in *exactly one way*. The total number of k-permutations is therefore given as the product:

$$|P(n,k)| = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

□

Counting Circular Permutations

Theorem 2: For any natural numbers $0 \leq k \leq n$, we have

$$|P_c(n,k)| = \frac{n!}{k \cdot (n-k)!}$$

Proof: We doubly count $P(n,k)$:

1. $|P(n,k)| = \frac{n!}{(n-k)!}$ which we proved before.
2. $|P(n,k)| = |P_c(n,k)| \cdot k$ because every equivalence class in $P_c(n,k)$ contains k permutations from $P(n,k)$ since there are k ways to rotate a k-permutation.

From this we get $\frac{n!}{(n-k)!} = |P_c(n,k)| \cdot k$, which implies $|P_c(n,k)| = \frac{n!}{k \cdot (n-k)!}$. □

Unordered Arrangements

Definition 12: An *unordered arrangement* of k elements of X is a *multiset* $S = \langle X, r \rangle$ of size k.

In a multiset, X is the set of *types*, and for each type $x \in X$, r_x is its *repetition number*.

Example: Let $X = \{$👒, 🦭, 🐈, 🪭, 🌵$\}$.

- An unordered arrangement of 7 elements could be $S = \{$👒, 👒, 🦭, 🐈, 🐈, 🐈, 🌵$\}^*$.
- The same multiset could be written as $S = \{2$👒$, 1$🦭$, 3$🐈$, 0$🪭$, 1$🌵$\}$.

Subsets

The most important special case of unordered arrangements is where all repetitions are 1, that is, $r_x = 1$ for all $x \in X$. Then S is simply a *subset* of X, denoted $S \subseteq X$.

Definition 13: A *k-combination* of X is an unordered arrangement of k *distinct* elements of X.

Note: The more standard term is *subset*. The term "combination" is only used to emphasize the selection process.

The set of all k-subsets of X is denoted $\binom{X}{k} = \{A \subseteq X \mid |A| = k\}$. If $|X| = n$, then

$$\binom{n}{k} := \left|\binom{X}{k}\right|$$

Example: The set of edges in a simple undirected graph consists of 2-subsets of its vertices: $E \subseteq \binom{V}{2}$.

Counting k-Combinations

Theorem 3: For $0 \leq k \leq n$, we have

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

Proof: $|P(n,k)| = \dfrac{n!}{(n-k)!} = \binom{n}{k} \cdot k!$ □

Four Counting Problems

	No repetition	**With repetition**
Ordered (k-perm., sequence)	$\dfrac{n!}{(n-k)!} = n^{\underline{k}}$	n^k
Unordered (k-comb., subset)	$\binom{n}{k} = \dfrac{n!}{k! \cdot (n-k)!}$	$\binom{k+n-1}{k}$

- n is the size of the ground set X.
- k is the size of the arrangement we want to count.
- "Ordered" means position matters: $(A, B) \neq (B, A)$.
- "Unordered" means only the selection matters: $\{A, B\} = \{B, A\}$.
- "With repetition" means the same element can appear multiple times.

Multisets

Multiset

Definition 14: A *multiset* is a modification of the concept of a set that allows for *repetitions* of its elements. Formally, it is denoted as a pair $M = \langle X, r \rangle$, where X is the *groundset* (the set of *types*) and $r : X \to \mathbb{N}_0$ is the *multiplicity function*.

Example: When the multiset is defined by enumeration, it is advisable to use the notation with the star:

$$M = \{a, b, a, a, b\}^* = \{3 \cdot a, 2 \cdot b\} \quad X = \{a, b\} \quad r_a = 3, r_b = 2$$

Example: Prime factorization of a natural number n is a multiset, *e.g.* $120 = 2^3 \cdot 3^1 \cdot 5^1$.

k-Combinations of a Multiset

Definition 15: Let X be a finite set of types, and let $M = \langle X, r \rangle$ be a finite multiset with repetition numbers $r_1, ..., r_{|X|}$. A *k-combination of M* is a multiset $S = \langle X, s \rangle$ with types in X and repetition numbers $s_1, ..., s_{|X|}$ such that $s_i \leq r_i$ for all $1 \leq i \leq |X|$, and $\sum_{i=1}^{|X|} s_i = k$.

Example: Consider $M = \{2🦫, 1🥤, 3🍉, 1💎\}$.

- $T = \{1🦫, 2🍉\}$ is a 3-combination of M.
- $T' = \{3💎\}$ is not.

Counting k-combinations of a multiset is not as simple as it might seem...

k-Permutations of a Multiset

Definition 16: Let M be a finite multiset with set of types X. A *k-permutation of M* is an ordered arrangement of k elements of M where different orderings of elements of the same type are *not distinguished*. This is an ordered multiset with types in X and repetition numbers $s_1, \ldots, s_{|X|}$ such that $s_i \leq r_i$ for all $1 \leq i \leq |X|$, and $\sum_{i=1}^{|X|} s_i = k$.

Note: There might be several elements of the same type compared to a permutation of a set (where each repetition number equals 1).

Example: Let $M = \{2🦫, 1🥤, 3🍉, 1💎\}$, then $T = (💎, 🍉, 🍉, 🦫)$ is a 4-permutation of multiset M.

Binomial Theorem

Theorem 4: The expansion of any non-negative integer power n of the binomial $(x + y)$ is a sum

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k}$$

where each $\binom{n}{k}$ is a positive integer known as a *binomial coefficient*, defined as

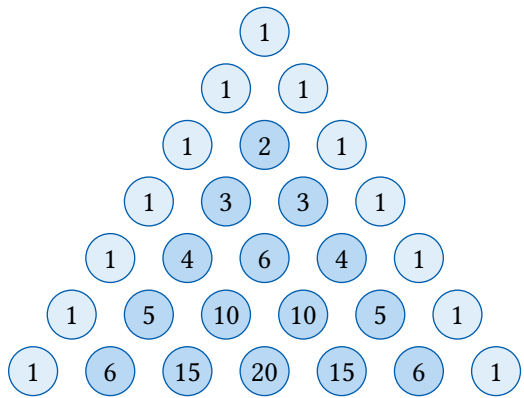
$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} = \frac{n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)}{k \cdot (k-1) \cdot (k-2) \cdot \ldots \cdot 2 \cdot 1}$$

Pascal's Triangle

Row n of *Pascal's triangle* lists $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}$. Each entry is the sum of the two entries above it: $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

Key identities:

- **Row sum:** $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$ (set $x = y = 1$)
- **Symmetry:** $\binom{n}{k} = \binom{n}{n-k}$
- **Absorption:** $n\binom{n-1}{k-1} = (k+1)\binom{n}{k+1}$
- **Vandermonde:** $\displaystyle\sum_{k} \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r}$



Multinomial Theorem

Theorem 5: The generalization of the binomial theorem:

$$(x_1 + ... + x_r)^n = \sum_{\substack{0 \le k_1,...,k_r \le n \\ k_1+...+k_r=n}} \binom{n}{k_1, ..., k_r} \cdot x_1^{k_1} \cdot ... \cdot x_r^{k_r}$$

Multinomial coefficients are defined as

$$\binom{n}{k_1, ..., k_r} = \frac{n!}{k_1! \cdot ... \cdot k_r!}$$

Note: Binomial coefficients are special cases of multinomial coefficients ($r = 2$):

$$\binom{n}{k} = \binom{n}{k_1, k_2} = \binom{n}{k, n-k} = \frac{n!}{k! \cdot (n-k)!}$$

Multinomial Theorem [2]

Proof: Expand $(x_1 + ... + x_r)^n$ by choosing one term from each of the n factors. A monomial $x_1^{k_1} \cdot ... \cdot x_r^{k_r}$ with $k_1 + ... + k_r = n$ is produced whenever we choose x_i from exactly k_i of the n factors. The number of ways to assign the n factor-positions to the r variables in groups of sizes $k_1, ..., k_r$ is $\binom{n}{k_1,...,k_r} = \frac{n!}{k_1! \cdot ... \cdot k_r!}$. □

Theorem 6: Let S be a finite multiset with k different types and repetition numbers $r_1, ..., r_k$. Let the size of S be $n = r_1 + ... + r_k$. Then the number of n-permutations of S equals

$$\binom{n}{r_1, ..., r_k}$$

Proof: In an n-permutation there are n positions that need to be assigned a type.

First, choose the r_1 positions for the first type, there are $\binom{n}{r_1}$ ways to do so. Then, assign r_2 positions for the second type, out of the $(n - r_1)$ positions that are still available, there are $\binom{n-r_1}{r_2}$ ways to do so. Continue for all k types. The total number of choices will be:

Multinomial Theorem [3]

$$\binom{n}{r_1} \cdot \binom{n-r_1}{r_2} \cdot \ldots \cdot \binom{n-r_1-r_2-\ldots-r_{k-1}}{r_k} = \binom{n}{r_1, \ldots, r_k}$$

□

k-Combinations of an *Infinite* Multiset

Example: Suppose you have a *sufficiently large* amount of each type of fruit (🍌, 🍎, 🍐) in the supermarket, and you want to buy *two* fruits. How many choices do you have?

There are exactly *six* combinations: {🍌, 🍌}, {🍌, 🍎}, {🍌, 🍐}, {🍎, 🍎}, {🍎, 🍐}, {🍐, 🍐}.

Note that your selection is *not ordered*, so {🍐, 🍎} and {🍎, 🍐} are considered the *same* choice.

k-Combinations of an *Infinite* Multiset [2]

Theorem 7: Let $k, s \in \mathbb{N}$ and let S be a multiset with s types and large repetition numbers (each $r_1, \ldots, r_s$ is *at least* k), then the number of k-combinations of S equals

$$\binom{k+s-1}{k} = \binom{k+s-1}{s-1}$$

Proof: Let $S = \langle X, r_\infty \rangle = \{\infty 🍌, \infty 🍎, \infty 🍐\}$ with $r_x = \infty$ and $|X| = s = 3$.

- Let $k = 5$ (as an example). Consider a 5-combination of S: {🍌, 🍎, 🍌, 🍐, 🍐}.
- Reorder and group: {🍌🍌 | 🍎 | 🍐🍐}.
- Convert to *dots* and *bars*: ●● | ● | ●●
- Represent as a 2-type multiset: $M = \{ k \cdot \bullet, (s-1) \cdot | \}$
- Observe: each *permutation* of k dots and $(s-1)$ bars corresponds *uniquely* to a *k-combination* of S.
- Permute the 2-type multiset: $\binom{k+s-1}{k,s-1}$ ways, by Theorem 5.

This method is also known as *Stars and Bars*. □

Compositions

Weak Compositions

Definition 17: A *weak composition* of a non-negative integer $k \geq 0$ into s parts is a *solution* to the equation $b_1 + ... + b_s = k$, where each $b_i \geq 0$.

Example: Let $k = 5$, $s = 3$. Possible non-negative integer solutions for $b_1 + b_2 + b_3 = 5$ are:

- $(b_1, b_2, b_3) = (1, 1, 3)$
- $(b_1, b_2, b_3) = (1, 3, 1)$
- $(b_1, b_2, b_3) = (2, 0, 3)$
- $(b_1, b_2, b_3) = (0, 5, 0)$
- ... (total 21 solutions)

Note: If M is a multiset over groundset $\{1, ..., s\}$ with all multiplicities infinite ($r_i = \infty$), then for $k \geq 0$, the number of sub-multisets of M of size k is exactly the number of weak compositions of k into s parts.

Counting Weak Compositions

Theorem 8: There are $\binom{k+s-1}{k,s-1}$ *weak compositions* of $k > 0$ into s parts.

Proof: Observe that $k = \overbrace{\underbrace{1+1}_{b_1} + \underbrace{...}_{b_i} + \underbrace{1+1}_{b_s}}^{k \text{ ones}}$.

Use the *stars-and-bars* method to count the number of s groups composed of k "ones". □

Example: Let $k = 3$. There are $\binom{3+3-1}{3,3-1} = \binom{5}{3} = \binom{5}{2} = 10$ ways to decompose $k = 3$ into $s = 3$ parts:

$$\begin{aligned} k = 3 &= \\ &= 0+1+2 = 0+2+1 \\ &= 1+0+2 = 1+2+0 = 1+1+1 \\ &= 2+0+1 = 2+1+0 \\ &= 3+0+0 = 0+3+0 = 0+0+3 \end{aligned}$$

Compositions

Definition 18: A *composition* of a positive integer $k \geq 1$ into s *positive* parts is a *solution* to the equation $b_1 + ... + b_s = k$, where each $b_i > 0$.

Theorem 9: There are $\binom{k-1}{s-1}$ *compositions* of $k > 0$ into s positive parts.

Proof: Line up k ones in a row. A composition $b_1 + ... + b_s = k$ with each $b_i \geq 1$ corresponds to choosing $s - 1$ of the $k - 1$ *gaps between consecutive ones* as dividers. There are $\binom{k-1}{s-1}$ such choices. □

Compositions [2]

Theorem 10: The total number of compositions of $k > 0$ into *some* number of positive parts is

$$\sum_{s=1}^{k} \binom{k-1}{s-1} = 2^{k-1}$$

Proof: Substituting $j = s - 1$:

$$\sum_{s=1}^{k} \binom{k-1}{s-1} = \sum_{j=0}^{k-1} \binom{k-1}{j} = (1+1)^{k-1} = 2^{k-1}$$

by the Binomial Theorem applied to $(1+1)^{k-1}$. □

Parallel Summation Identity

Q: How many integer solutions are there to the *inequality* $b_1 + ... + b_s \leq k$, where each $b_i \geq 0$?

Theorem 11: $\sum_{m=0}^{k} \binom{m+s-1}{m} = \binom{k+s}{k}$

Proof *(hint)*: Introduce a "dummy" variable b_{s+1} to take up the *slack* between $b_1 + ... + b_s$ and k. Construct a bijection with the solutions to $b_1 + ... + b_s + b_{s+1} = k$, where $b_i \geq 0$. □

Set Partitions

Set Partitions

Definition 19: A *partition* of a set X is a set of non-empty subsets of X such that every element of X belongs to exactly one of these subsets.

Equivalently, a family of sets P is a partition of X iff:

1. The family P does not contain the empty set: $\emptyset \notin P$.
2. The union of P is X, that is, $\bigcup_{A \in P} A = X$. The sets in P are said to *cover* X.
3. The intersection of any two distinct sets in P is empty: $\forall A, B \in P.\, (A \neq B) \to (A \cap B = \emptyset)$. The sets in P are said to be *pairwise disjoint* or *mutually exclusive*.

The sets in P are called *blocks*, *parts*, or *cells*, of the partition.

The block in P containing an element $x \in X$ is denoted by $[x]$.

Examples of Set Partitions

Example: The empty set $X = \emptyset$ has exactly one partition, $P = \emptyset$.

Example: Any singleton set $X = \{x\}$ has exactly one partition, $P = \{\{x\}\}$.

Example: For any non-empty proper subset $A \subset U$, the set A and its complement form a partition of U, namely $P = \{A, U - A\}$.

Example: The set $X = \{1, 2, 3\}$ has five partitions:

1. $\{\{1\}, \{2\}, \{3\}\}$ or $1 \mid 2 \mid 3$
2. $\{\{1\}, \{2, 3\}\}$ or $1 \mid 2\ 3$
3. $\{\{1, 2\}, \{3\}\}$ or $1\ 2 \mid 3$
4. $\{\{1, 3\}, \{2\}\}$ or $1\ 3 \mid 2$
5. $\{\{1, 2, 3\}\}$ or $1\ 2\ 3$

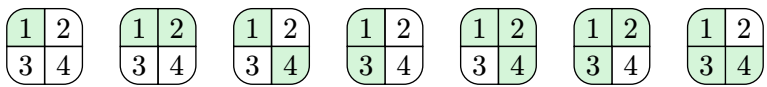
Example: The following are *not* partitions of $\{1, 2, 3\}$:

- $\{\{\}, \{1, 3\}, \{2\}\}$, because it contains the empty set.
- $\{\{1, 2\}, \{2, 3\}\}$, because the element 2 is contained in more than one block.
- $\{\{1\}, \{3\}\}$, because no block contains the element 3.

Counting Set Partitions

Definition 20: The number of partitions of a set X (of size $n = |X|$) into k non-empty blocks ("unlabeled subsets") is called a *Stirling number of the second kind* and denoted $S(n, k)$ or $\left\{ {n \atop k} \right\}$.

Example: Let $X = \{1, 2, 3, 4\}$, $k = 2$. There are 7 possible partitions:



Counting Set Partitions [2]

Theorem 12: Let $\left\{ {n \atop 0} \right\} = 0$ for $n \geq 1$, $\left\{ {0 \atop k} \right\} = 0$ for $k \geq 1$, and $\left\{ {0 \atop 0} \right\} = 1$. For $n, k \geq 1$, we have:

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \cdot \left\{ {n-1 \atop k} \right\}$$

Proof *(informal)*: Consider the element n in any partition of $[n]$ into k blocks.

There are exactly two cases:

1. $\{n\}$ forms a singleton block on its own.
   The remaining $n-1$ elements are partitioned into $k-1$ non-empty blocks: $\left\{ {n-1 \atop k-1} \right\}$ ways.
2. n joins an existing block of a partition of $[n-1]$.
   Start with any partition of $[n-1]$ into k blocks ($\left\{ {n-1 \atop k} \right\}$ ways),
   then insert n into one of the k blocks: $k \cdot \left\{ {n-1 \atop k} \right\}$ ways.

These two cases are exhaustive and disjoint. □

Bell Numbers

Definition 21: The total number of partitions of a set X of size $n = |X|$ (into an arbitrary number of non-empty blocks) is called a *Bell number* and denoted B_n.

$$B_n = \sum_{k=0}^{n} \left\{ {n \atop k} \right\}$$

Note: Consider the special case of $n = 0$. There is exactly *one* partition of $\emptyset$ into non-empty parts: $\emptyset = \bigcup_{A \in \emptyset} A \in \emptyset$. Every $A \in \emptyset$ is non-empty, since no such A exists. Thus, we have $B_0 = S(0, 0) = 1$.

Bell Numbers [2]

Theorem 13: For $n \geq 1$, we have the recursive identity for Bell numbers:

$$B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k$$

Proof: Every partition of $[n]$ has one part that contains the number n. In addition to n, this part also contains k other numbers (for some $0 \leq k \leq n-1$). The remaining $n-1-k$ elements are partitioned arbitrarily. From this correspondence, we obtain the desired identity:

$$B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_{n-1-k} = \sum_{k=0}^{n-1} \binom{n-1}{n-1-k} B_{n-1-k} = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k$$

□

Integer Partitions

Integer Partitions

The integer 4 can be written as a sum of positive integers in exactly 5 ways: $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$. Order does *not* matter: $3 + 1$ and $1 + 3$ are the same partition.

Definition 22: An *integer partition* of a positive integer $n \geq 1$ into k *positive* parts is a solution to the equation $n = a_1 + ... + a_k$, where $a_1 \geq a_2 \geq ... \geq a_k \geq 1$.

The number of partitions into exactly k parts is $p_k(n)$, defined recursively:

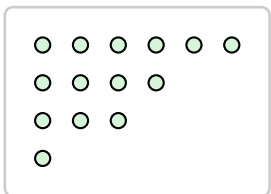
$$p_k(n) = \begin{cases} 0 & \text{if } k > n \\ 0 & \text{if } n \geq 1 \text{ and } k = 0 \\ 1 & \text{if } n = k = 0 \\ p_k(n-k) + p_{k-1}(n-1) & \text{if } 1 \leq k \leq n \end{cases}$$

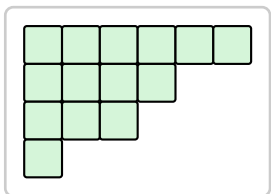
The *partition function* $p(n) = \sum_{k=0}^{n} p_k(n)$ counts all partitions of n.

Ferrers Diagrams and Young Tableaux

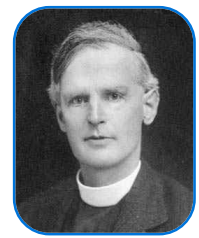
Example: Consider an integer partition: $14 = 6 + 4 + 3 + 1$.

Ferrer Diagram **Young Tableaux**





Norman Ferrer



Alfred Young

Inclusion–Exclusion

"There are very few things which we know, which are not capable of being reduced to a mathematical reasoning."

— Abraham de Moivre, The Doctrine of Chances (1718)

Motivating Example

Example: How many integers in $\{1, \ldots, 100\}$ are divisible by 2 *or* by 3?

- Let $A = \{k \in [100] \mid 2 \mid k\}$, so $|A| = 50$.
- Let $B = \{k \in [100] \mid 3 \mid k\}$, so $|B| = 33$.
- Their intersection $A \cap B = \{k \in [100] \mid 6 \mid k\}$ has $|A \cap B| = 16$.

By inclusion–exclusion:

$$|A \cup B| = |A| + |B| - |A \cap B| = 50 + 33 - 16 = 67$$

Generalizing this to an arbitrary number of sets gives the full PIE theorem.

Principle of Inclusion–Exclusion (PIE)

Theorem 14: Let X be a finite set and $P_1, ..., P_m$ *properties*.
- Define $X_i = \{x \in X \mid x \text{ has } P_i\}$, *i.e.* the set of all elements from X having a property P_i.
- Define for $S \subseteq [m]$ the set $N(S) = \{x \in X \mid \forall i \in S : x \text{ has } P_i\}$. Observe: $N(S) = \bigcap_{i \in S} X_i$.

The number of elements of X that satisfy *none* of the properties $P_1, ..., P_m$ is given by

$$|X \setminus (X_1 \cup ... \cup X_m)| = \sum_{S \subseteq [m]} (-1)^{|S|} |N(S)| \tag{1}$$

Proof: Consider any $x \in X$. If $x \in X$ has none of the properties, then $x \in N(\emptyset)$ and $x \notin N(S)$ for any other $S \neq \emptyset$. Hence *x contributes 1* to the sum (1) .

If $x \in X$ has exactly $k \geq 1$ of the properties, call this set $T \in \binom{[m]}{k}$. Then $x \in N(S)$ iff $S \subseteq T$.
The *contribution of x* to the sum (1) is $\sum_{S \subseteq T} (-1)^{|S|} = \sum_{i=0}^{k} \binom{k}{i} (-1)^i = 0$, *i.e. zero*. □

Note: In the last step, we used that for any $y \in \mathbb{R}$ we have $(1-y)^k = \sum_{i=0}^{k} \binom{k}{i} (-y)^i$ which implies (for $y = 1$) that $0 = \sum_{i=0}^{k} \binom{k}{i} (-1)^i$.

Counting the Union

Corollary 14.1:

$$\left|\bigcup_{i\in[m]} X_i\right| = \sum_{\emptyset\neq S\subseteq[m]} (-1)^{|S|-1}|N(S)|$$

Applications of PIE

Step 1: Define "bad" properties

Identify elements of a universe X to *exclude*: those having *at least one* of properties $P_1, ..., P_m$.

Target: count $X \setminus (X_1 \cup ... \cup X_m)$.

Step 2: Count $N(S)$

For each $S \subseteq [m]$, determine $N(S)$ — number of elements of X having *all* bad properties P_i for $i \in S$.

Step 3: Apply PIE

Use Theorem 14 to compute a closed formula:

$$|X \setminus (X_1 \cup ... \cup X_m)| = \sum_{S \subseteq [m]} (-1)^{|S|} N(S)$$

Counting Surjections via PIE

Theorem 15: The number of surjections from $[k]$ to $[n]$ is given by

$$\left|\left\{f : [k] \underset{\text{surj.}}{\to} [n]\right\}\right| = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^k$$

Proof: Let X be the set of all maps from $[k]$ to $[n]$.

1. *Define bad properties:* Define the "bad" property P_i for $i \in [n]$ as "i is not in the image of f", *i.e.*

$$f : [k] \to [n] \text{ has property } P_i \iff \forall j \in [k] : f(j) \neq i$$

   The *surjective* functions are exactly those functions that *do not* have any of the "bad" properties.

2. *Count $N(S)$:* We claim $N(S) = (n - |S|)^k$ for any $S \subseteq [n]$. To see this, observe that f has all properties with indices from S if and only if $f(i) \notin S$ for all $i \in [k]$. In other words, f must be a function from $[k]$ to $[n] \setminus S$, and there are $(n - |S|)^k$ of those.

Counting Surjections via PIE [2]

3. *Apply PIE:* Using Theorem 14, the number of surjections from $[k]$ to $[n]$ is

$$\begin{aligned} |X \setminus (X_1 \cup \ldots \cup X_n)| &\overset{\text{PIE}}{=} \sum_{S \subseteq [n]} (-1)^{|S|} |N(S)| \\ &= \sum_{S \subseteq [n]} (-1)^{|S|} (n - |S|)^k \\ &= \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^k \end{aligned}$$

In the last step, we used that $(-1)^{|S|}(n-|S|)^k$ only depends on the size of S, and there are $\binom{n}{i}$ sets $S \subseteq [n]$ of size i.

□

Useful Corollaries

Corollary 15.1: Consider the case $n = k$. A function from $[n]$ to $[n]$ is a *surjection* iff it is a *bijection*. Since there are $n!$ bijections on $[n]$ (namely, all permutations), we have the following identity:

$$n! = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^n$$

Corollary 15.2: A surjection from $[k]$ to $[n]$ can be seen as a partition of $[k]$ into n non-empty distinguishable (labeled) parts (the map assigns a part to each $i \in [k]$).

Since the partition of $[k]$ into n non-empty indistinguishable parts is denoted $s_n^{\text{II}}(k)$, and there are $n!$ ways to assign labels to n parts, we obtain that the number of surjections is equal to $n! s_n^{\text{II}}(k)$:

$$n! s_n^{\text{II}}(k) = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^k$$

Derangements

Theorem 16: The *derangements* D_n on n elements are permutations of $[n]$ without fixed points.

The number of derangements is given by

$$|D_n| = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)!$$

Proof: Let X be the set of all permutations of $[n]$.

1. Define the "bad" property P_i to mean "π has a fixpoint i" ($i \in [n]$):

$$\pi \in X \text{ has property } P_i \iff \pi(i) = i$$

2. We claim $N(S) = (n - |S|)!$ for any $S \subseteq [n]$.

   Indeed, $\pi \in X$ has all properties with indices from S if and only if all $i \in S$ are fixed points of π. On the other elements, *i.e.* on $[n] \setminus S$, π may be an arbitrary bijection, so there are $(n - |S|)!$ choices for π.

Derangements [2]

3. Using Theorem 14, the number of derangements is given by

$$\begin{aligned} |X \setminus (X_1 \cup \ldots \cup X_n)| &\overset{\text{PIE}}{=} \sum_{S \subseteq [n]} (-1)^{|S|} |N(S)| \\ &= \sum_{S \subseteq [n]} (-1)^{|S|} (n - |S|)! \\ &= \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)! \end{aligned}$$

In the last step, we used that $(-1)^{|S|}(n - |S|)!$ only depends on the size of S, and there are $\binom{n}{i}$ sets $S \subseteq [n]$ of size i.

□

Derangements [3]

Hat-check problem: n guests check their hats; the attendant loses the tickets.

As $n \to \infty$, the probability that no guest gets their own hat back converges to

$$\lim_{n\to\infty} \frac{|D_n|}{n!} = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} = \frac{1}{e} \approx 36.8\%$$

This probability barely changes for $n \geq 5$.

Generating Functions

"A generating function is a clothesline on which we hang up a sequence of numbers for display."

— Herbert Wilf, generatingfunctionology

Generating Functions

A generating function is a device somewhat similar to a bag. Instead of carrying many little objects detachedly, which could be embarrassing, we put them all in a bag, and then we have only one object to carry, the bag.

— George Pólya, Mathematics and Plausible Reasoning [1]

Encode a sequence (a_n) as coefficients of a formal power series $G(x) = \sum_{n=0}^{\infty} a_n x^n$.

Then manipulate $G(x)$ *algebraically* (multiply, differentiate, compose) to derive counting identities and closed forms.

Counting with Polynomials

Problem. Two fair dice. How many outcomes give sum n?

Encode one die as a polynomial: the term x^k stands for "rolling k is possible."

$$D(x) = x + x^2 + x^3 + x^4 + x^5 + x^6$$

Two independent dice $\longrightarrow$ multiply:

$$D(x)^2 = x^2 + 2x^3 + 3x^4 + 4x^5 + 5x^6 + \underline{6x^7} + 5x^8 + 4x^9 + 3x^{10} + 2x^{11} + x^{12}$$

The coefficient of x^n gives the count for sum n. For sum 7: $[x^7]D(x)^2 = 6$. ✓

One multiplication solves the problem for *every* n simultaneously.

This polynomial is a *generating function* — now let's understand what they are and why this works.

Ordinary Generating Functions

Definition 23: An *ordinary generating function* (OGF) of a sequence a_n is a *power series*

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

Example: The *sequence* $a_n = (a_0, a_1, a_2, ...)$ is *generated* by the OGF $G(x) = a_0 + a_1 x + a_2 x^2 + ...$

Example: $G(x) = 3 + 8x^2 + x^3 + \frac{1}{7}x^5 + 100x^6 + ...$ *generates* the sequence $\left(3, 0, 8, 1, 0, \frac{1}{7}, 100, 0, ...\right)$

Example: Consider a long division of 1 by $(1 - x)$, the result is an infinite power series

$$\frac{1}{1-x} = 1 + x^1 + x^2 + x^3 + ... = \sum_{n=0}^{\infty} x^n$$

Note that all coefficients are 1. Thus, the generating function of $(1, 1, 1, ...)$ is $G(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$.

The Core Generating Function

Another proof that $(1, 1, 1, ...)$ is generated by $G(x) = 1 + x + x^2 + x^3 + ... = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x} = S$:

$$\begin{array}{rcl} S & = & 1 + x + x^2 + x^3 + ... \\ x \cdot S & = & \quad x + x^2 + x^3 + ... \\ \hline S - x \cdot S & = & 1 \end{array}$$

Thus, $S = \dfrac{1}{1-x}$.

The generating function $G(x) = 1 + x + x^2 + ...$ is also known as the *Maclaurin series* of $\dfrac{1}{1-x}$.

More Examples of Generating Functions

Formula	Power series	Sequence	Description
$\frac{1}{1-x}$	$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \ldots$	$(1, 1, 1, \ldots)$	constant 1
$\frac{2}{1-x}$	$\sum_{n=0}^{\infty} 2x^n = 2 + 2x + 2x^2 + 2x^3 + \ldots$	$(2, 2, 2, \ldots)$	constant 2
$\frac{x}{1-x}$	$\sum_{n=1}^{\infty} x^n = 0 + x + x^2 + x^3 + \ldots$	$(0, 1, 1, 1, \ldots)$	right shift
$\frac{1}{1+x}$	$\sum_{n=0}^{\infty} (-1)^n x^n = 0 + 1 - x + x^2 - x^3 + \ldots$	$(1, -1, 1, \ldots)$	sign-alternating 1's
$\frac{1}{1-3x}$	$\sum_{n=0}^{\infty} 3^n x^n = 1 + 3x + 9x^2 + 27x^3 + \ldots$	$(1, 3, 9, \ldots)$	powers of 3

More Examples of Generating Functions [2]

Formula	Power series	Sequence	Description
$\frac{1}{1-x^2}$	$\sum_{n=0}^{\infty} x^{2n} = 1 + x^2 + x^4 + x^6 + ...$	$(1, 0, 1, 0, ...)$	regular gaps
$\frac{1}{(1-x)^2}$	$\sum_{n=0}^{\infty} (n+1)x^n = 1 + 2x + 3x^2 + 4x^3 + ...$	$(1, 2, 3, 4, ...)$	natural numbers

$$\begin{aligned}
\frac{1-x^{n+1}}{1-x} &= \frac{1}{1-x} - \frac{x^{n+1}}{1-x} = \\
&\triangleq (1, 1, 1, ...) - (\underbrace{0, 0, ..., 0}_{n+1 \text{ zeros}}, 1, 1, ...) = \\
&= (\underbrace{1, 1, ..., 1}_{n+1 \text{ ones}}, 0, 0, ...) = \\
&\triangleq 1 + x + x^2 + ... + x^n
\end{aligned}$$

Exercises

Example: Find the GF for odd numbers: $(1, 3, 5, 7, ...)$.

Hint: write $A(x) = 1 + 3x + 5x^2 + 7x^3 + ...$, then compute $x \cdot A(x)$ and subtract from $A(x)$.

What familiar series do you see in the result?

Example: Find the GF for $(1, 3, 7, 15, 31, 63, ...)$, which satisfies $a_n = 3a_{n-1} - 2a_{n-2}$ with $a_0 = 1$, $a_1 = 3$.

Hint: write $A(x)$, $-3xA(x)$, and $+2x^2A(x)$ stacked and add them column by column.

Most terms cancel because of the recurrence. What remains?

Solutions

Example: **Odd numbers.** Let $A(x) = 1 + 3x + 5x^2 + 7x^3 + ...$

$$\begin{array}{rcl} A(x) & = & 1 + 3x + 5x^2 + 7x^3 + 9x^4 + ... \\ x \cdot A(x) & = & \quad + x + 3x^2 + 5x^3 + 7x^4 + ... \\ \hline A - x \cdot A & = & 1 + 2x + 2x^2 + 2x^3 + 2x^4 + ... \end{array}$$

The result is $1 + 2(x + x^2 + x^3 + ...) = 1 + 2 \cdot \frac{x}{1-x}$, so:

$$A(x) = \frac{1 + 2 \cdot \frac{x}{1-x}}{1 - x} = \frac{1 - x + 2x}{(1-x)^2} = \frac{1 + x}{(1-x)^2}$$

Check: $\frac{1+x}{(1-x)^2} = (1, 1, 1, ...) + (0, 2, 4, 6, 8, ...) = (1, 3, 5, 7, 9, ...)$ ✓

Solutions [2]

Example: **Recurrence sequence.** Stack A, $-3x \cdot A$, $+2x^2 \cdot A$ and add:

$$\begin{array}{rcl} A(x) &=& 1 + 3x + 7x^2 + 15x^3 + 31x^4 + \ldots \\ -3x \cdot A(x) &=& \quad - 3x - 9x^2 - 21x^3 - 45x^4 - \ldots \\ 2x^2 \cdot A(x) &=& \qquad\quad 2x^2 + 6x^3 + 18x^4 + \ldots \\ \hline (1 - 3x + 2x^2) \cdot A(x) &=& 1 \end{array}$$

Every column beyond x^0 vanishes because $a_n - 3a_{n-1} + 2a_{n-2} = 0$ by the recurrence.

Thus, $A(x) = \dfrac{1}{1 - 3x + 2x^2}$

Solving Combinatorial Problems via Generating Functions

Example: Find the number of integer solutions to $y_1 + y_2 + y_3 = 12$ with $0 \leq y_i \leq 6$.

- Possible values for y_1 are $0 \leq y_1 \leq 6$.
  - There is a *single* way to select $y_1 = 0$. The same for other values among $1, ..., 6$.
  - There are *no* ways to select any value of y_1 higher than 6.
  - The *number of ways to select y_1 to be equal to n* forms a sequence $(1, 1, 1, 1, 1, 1, 1, 0, ...)$.
  - Write this sequence as a polynomial $x^0 + x^1 + ... + x^6$.
  - Do the same for y_2 and y_3 (*in isolation*!).
- Since all combinations of y_1, y_2 and y_3 are valid non-conflicting solutions, we can multiply those polynomials and obtain the *generating function* $G(x) = (1 + x + x^2 + ... + x^6)^3$.
  - For each n, the coefficient of x^n in $G(x)$ is the number of integer solutions to $y_1 + y_2 + y_3 = n$.
  - In particular, we are interested in the coefficient of x^{12} in $G(x)$, denoted $[x^{12}]G(x)$.
  - Use ~~pen and paper~~ Wolfram Alpha to expand $G(x)$:

$$G(x) = x^{18} + 3x^{17} + 6x^{16} + ... + \underline{28x^{12}} + ... + 6x^2 + 3x + 1$$

- The *answer* is $[x^{12}]G(x) = 28$ solutions.

Slightly More Complex Combinatorial Problem

Example: Suppose we have marbles of three different colors (🔴, 🟢, 🔵), and we want to *count* the number of ways to select 20 marbles, such that:

- There are an even number of 🔴: $1 + x^2 + x^4 + \ldots + x^{20}$.
- There are at least 12 🟢: $x^{12} + x^{13} + \ldots + x^{20}$.
- There are at most 5 🔵: $1 + x + x^2 + x^3 + x^4 + x^5$.

Multiply polynomials and find $[x^{20}]G(x)$:

$$
\begin{aligned}
&[x^{20}](1 + x^2 + x^4 + \ldots + x^{20})(x^{12} + x^{13} + \ldots + x^{20})(1 + x + x^2 + x^3 + x^4 + x^5) = \\
&= [x^{20}]\left(x^{45} + 2x^{44} + \ldots + \underline{21x^{20}} + \ldots + 2x^{13} + x^{12}\right) \\
&= 21
\end{aligned}
$$

Using Power Series in Combinatorial Problems

Example: Find the number of integer solutions to $a_1 + a_2 + a_3 = 12$ with $a_1 \geq 2, 3 \leq a_2 \leq 6, 0 \leq a_3 \leq 9$.

- Compose the generating function:

$$G(x) = (x^2 + x^3 + ...) \cdot (x^3 + x^4 + x^5 + x^6) \cdot (1 + x + x^2 + ... + x^9)$$

- Substitute the power series with the corresponding simple forms:

$$G(x) = \left(x^2 \cdot \frac{1}{1-x}\right) \cdot \left(x^3 \cdot \frac{1-x^4}{1-x}\right) \cdot \left(\frac{1-x^{10}}{1-x}\right)$$

- Expand the series:

$$G(x) = x^5 + 3x^6 + 6x^7 + 10x^8 + 14x^9 + 18x^{10} + 22x^{11} + \underbrace{26x^{12}} + 30x^{13} + 34x^{14} + 37x^{15} + 39x^{16} + 40x^{17} + ... + 40x^n + ...$$

- Sequence: $(g_n) = (0, 0, 0, 0, 0, 1, 3, 6, 10, 14, 18, 22, 26, 30, 34, 37, 39, \overline{40}, ...)$
- Answer for $n = 12$ is $[x^{12}]G(x) = 26$.

Operations on Generating Functions

Let $F(x) = \sum_{n=0}^{\infty} a_n x^n$ and $G(x) = \sum_{n=0}^{\infty} b_n x^n$ be ordinary generating functions.

All operations are performed *term-wise* on the power series.

Operation	Result
Differentiate $F(x)$ term-wise	$F'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$
Multiply $F(x)$ by a scalar $\lambda \in \mathbb{R}$ term-wise	$\lambda F(x) = \sum_{n=0}^{\infty} \lambda a_n x^n$
Add $F(x)$ and $G(x)$ term-wise	$F(x) + G(x) = \sum_{n=0}^{\infty} (a_n + b_n) x^n$
Multiply $F(x)$ and $G(x)$ term-wise (*Cauchy product*, or *convolution*)	$F(x) \cdot G(x) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) x^n$

Catalan Numbers via Generating Functions

So far, we *composed* generating functions from known polynomials and read off coefficients.

Now we turn the idea around:

1. Start from a *recursive counting argument*.
2. Derive a *functional equation* for $F(x)$.
3. Solve it.

This requires a heavier tool — *Newton's binomial theorem* — but the payoff is one of the most famous sequences in combinatorics.

Historical note: Eugène Catalan (1814–1894) studied these numbers in the context of polygon triangulation. They appear in dozens of seemingly unrelated counting problems — a phenomenon that generating functions help explain.

Well-Formed Parenthesis Expressions

Example: Find the number of *well-formed parenthesis expressions* with n pairs of parenthesis.

For example, "`(()())()`" is a well-formed parenthesis expression with 4 pairs of parenthesis.

Formally, a permutation of the multiset $\{n \cdot \texttt{"("}, n \cdot \texttt{")"}\}$ is *well-formed* if reading it from left to right and counting "$+1$" for every opening parenthesis "(" and "-1" for every closing parenthesis ")" never yields a negative number at any time.

Every well-formed expression with $n \geq 1$ pairs of paranthesis starts with "(" and there is a unique matching ")" such that the sequence in between and the sequence after are well-formed. For example:

$$\big(()\big)()() \quad \big(()(())\big) \quad \big()(()(())\big)$$

In other words, a well-formed expression with n pairs of parenthesis is obtained by putting a well-formed expression with k pairs in between "(" and ")" and then appending a well-formed expression with $n-k-1$ pairs of parenthesis. This gives the equation:

$$a_n = \sum_{k=0}^{n-1} a_k a_{n-k-1}$$

Well-Formed Parenthesis Expressions [2]

Let $F(x)$ be a generating function for a_n, then we know:

$$F(x) = \sum_{n=0}^{\infty} a_n x^n = 1 + \sum_{n=1}^{\infty} \left(\sum_{k=0}^{n-1} a_k a_{n-k-1} \right) x^n = 1 + \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k a_{n-k} \right) x^{n+1}$$

$$= 1 + x \cdot \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k a_{n-k} \right) x^n = 1 + x \cdot F(x)^2$$

Solving the equation $xF^2 - F + 1 = 0$ for F gives:

$$F(x) = \frac{1 \pm \sqrt{1 - 4x}}{2x}$$

We are only interested in the root with the *minus* sign, since $\lim_{x \to 0} F(x) = a_0 = 1$:

$$\lim_{x \to 0} F(x) = \lim_{x \to 0} \frac{1 - \sqrt{1 - 4x}}{2x} = \lim_{x \to 0} \frac{\frac{2}{\sqrt{1-4x}}}{2} = 1$$

Newton's Binomial Theorem

To extract the coefficients of $F(x) = \frac{1-\sqrt{1-4x}}{2x}$, we need to expand $\sqrt{1-4x}$ as a power series.

This requires extending the binomial theorem to *real* exponents.

Let's revisit the binomial theorem:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k = \sum_{k=0}^{\infty} \binom{n}{k} x^k \quad \forall n \in \mathbb{N}$$

where $\binom{n}{k} = 0$ for $k > n$.

Note: This shows that $(1+x)^n$ is the generating function for the series $(a_k)_{k\in\mathbb{N}}$ with $a_k = \binom{n}{k}$.

We can extend this result from natural numbers $n \in \mathbb{N}$ to any *real* number $n \in \mathbb{R}$.

Binomial Coefficients for Real Numbers

Definition 24: Let $p(n,k) = n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$, also called the *falling factorial* $n^{\underline{k}}$.

Extend the definition of binomial coefficients for real numbers $n, k \in \mathbb{R}$:

$$\binom{n}{k} = \frac{p(n,k)}{k!}$$

Note: This definition aligns with the definition of binomial coefficients for natural numbers:

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

Example: Consider the number "$-7/2$ choose 5":

$$\binom{-7/2}{5} = \frac{-7/2 \cdot -9/2 \cdot -11/2 \cdot -13/2 \cdot -15/2}{5!} = -\frac{9009}{256} \approx -35$$

Note: $p(n,0) = 1$ and for $k \geq 1$, we have $p(n,k) = (n-k+1) \cdot p(n,k-1) = n \cdot p(n-1,k-1)$. $(\star)$

Extended Newton's Binomial Theorem

Theorem 17: For all non-zero $n \in \mathbb{R}$, we have:

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

Example: Let $n = 1/2$, then we have an identity for $\sqrt{1+x}$:

$$\sqrt{1+x} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k$$

To actually *use* this fact, we need some *lemma*...

Lemma 18: For any integer $n \geq 1$, we have:

$$\binom{1/2}{n} = (-1)^{n+1} \cdot \binom{2n-2}{n-1} \cdot \frac{1}{2^{2n-1}} \cdot \frac{1}{n}$$

Extended Newton's Binomial Theorem [2]

Proof: By induction on n.

Base: $n = 1$.

$$\binom{1/2}{1} = \frac{1/2}{1!} = \frac{1}{2} = 1 \cdot 1 \cdot \frac{1}{2} \cdot 1 = \underbrace{(-1)^2}_{1} \cdot \underbrace{\binom{2-2}{1-1}}_{1} \cdot \underbrace{\frac{1}{2^{2-1}}}_{\frac{1}{2}} \cdot \underbrace{\frac{1}{1}}_{1}$$

Induction step: n to $n+1$ for $n > 1$. We use the recusion $(\star)$ $p(n,k) = n \cdot p(n-1, k-1)$:

$$\binom{1/2}{n+1} = \frac{p(1/2, n+1)}{(n+1)!} = \frac{(1/2 - (n+1) + 1) \cdot p(1/2, n)}{(n+1) \cdot n!} = -\frac{n - 1/2}{n+1}\binom{1/2}{n}$$

$$\overset{\text{IH}}{=} -\frac{n-1/2}{n+1}(-1)^{n+1} \cdot \binom{2n-2}{n-1} \cdot \frac{1}{2^{2n-1}} \cdot \frac{1}{n}$$

$$= \frac{2n}{2n} \cdot \frac{2n-1}{2n} \cdot (-1)^{n+2} \cdot \binom{2n-2}{n-1} \cdot \frac{1}{2^{2n-1}} \cdot \frac{1}{n+1}$$

$$= (-1)^{n+2} \cdot \underbrace{\frac{(2n-2)! \cdot (2n-1) \cdot (2n)}{(n-1)! \cdot (n-1) \cdot n \cdot n}}_{\binom{2n}{n}} \cdot \frac{1}{2^{2n+1}} \cdot \frac{1}{n+1}$$

□

Catalan Numbers

Proposition 19: Now we can expand $\sqrt{1+n}$ into the following series:

$$\sqrt{1+n} = \sum_{n=0}^{\infty} \binom{1/2}{n} x^n = 1 + \sum_{n=1}^{\infty} -2 \cdot \binom{2n-2}{n-1} \cdot (-1)^n \cdot \frac{1}{2^{2n}} \cdot \frac{1}{n} \cdot x^n$$

Example: Going back to the example with the number of well-formed paranthesis expressions, we get:

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

The numbers $C_n := \binom{2n}{n} \frac{1}{n+1}$ are called *Catalan numbers*.

Catalan Numbers [2]

Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ count many combinatorial structures:

- well-formed parenthesis strings of length $2n$,
- full binary trees with $n+1$ leaves,
- triangulations of a convex $(n+2)$-gon,
- monotone lattice paths from $(0,0)$ to (n,n) not crossing the diagonal.

$$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5, \quad C_4 = 14, \quad C_5 = 42, \quad \ldots$$

Recurrence Relations

"A certain man placed one pair of rabbits in a place enclosed on all sides by a wall. How many pairs of rabbits can be produced from that pair in one year?"

— Leonardo Fibonacci, Liber Abaci (1202)

Recurrence Relations

A *recurrence relation* defines a sequence by expressing each term via its predecessors.

Solving one means finding a *closed-form* formula that avoids recursion.

Example: *Arithmetic progression* as a recurrence:

$$a_n = \begin{cases} a_0 = \text{const} & \text{if } n = 0 \\ a_{n-1} + d & \text{if } n > 0 \end{cases}$$

Solution by telescoping: $a_n = a_0 + n \cdot d$.

Recursive Structure

Many combinatorial objects have *recursive structure*: a large instance decomposes into smaller instances of the same type.

A *recurrence* is a formula that captures this decomposition.

Example: How many binary strings of length n contain *no two consecutive* 1s? Let $B(n)$ denote the count.

Look at the last bit of a valid string of length n:

- **Ends in 0:** the first $n-1$ bits can be any valid string, that gives $B(n-1)$ options.
- **Ends in 1:** then the bit before it *must* be 0, so the first $n-2$ bits can be any valid string, giving $B(n-2)$ options.

Therefore $B(n) = B(n-1) + B(n-2)$, with $B(1) = 2$ and $B(2) = 3$.

This is the *Fibonacci recurrence* with shifted initial conditions.

Telescoping

Before reaching for characteristic equations, try the simplest approach: *unwind the recurrence*.

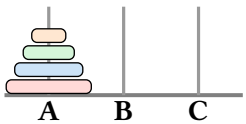
Example: Solve $T(n) = T(n-1) + n$ with $T(0) = 0$.

Unwind by substitution:

$$\begin{aligned} T(n) &= T(n-1) + n \\ &= T(n-2) + (n-1) + n \\ &= \vdots \\ &= T(0) + 1 + 2 + \ldots + n \\ &= \frac{n(n+1)}{2} \end{aligned}$$

Theorem 20: If $a_n = a_{n-1} + f(n)$, then $a_n = a_0 + \sum_{i=1}^{n} f(i)$ (telescoping sum).

Tower of Hanoi



Move n disks from peg A to peg C, one at a time, never placing a larger disk on a smaller one.

Strategy: move the top $n-1$ disks to peg B, move the largest disk to peg C, then move the $n-1$ disks from B to C.

Example: Let $T(n)$ be the minimum number of moves. Recurrence: $T(n) = 2T(n-1) + 1$, with $T(1) = 1$.

Unwind:

$$\begin{aligned} T(n) &= 2T(n-1) + 1 \\ &= 2(2T(n-2) + 1) + 1 = 4T(n-2) + 3 \\ &= 8T(n-3) + 7 \\ &= 2^k T(n-k) + \left(2^k - 1\right) \end{aligned}$$

Setting $k = n-1$: $T(n) = 2^{n-1}T(1) + 2^{n-1} - 1 = 2^n - 1$.

Linear Homogeneous Recurrence Relations

Definition 25: A *linear homogeneous* recurrence relation *of degree k* with constant coefficients is a recurrence relation of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k},$$

where $c_1, c_2, \ldots, c_k$ are constants (real or complex numbers), and $c_k \neq 0$.

Examples:

- $b_n = 2.71 b_{n-1}$ is a linear homogeneous recurrence relation of degree 1.
- $F_n = F_{n-1} + F_{n-2}$ is a linear homogeneous recurrence relation of degree 2.
- The recurrence relation $a_n = a_{n-1} + a_{n-2}^2$ is *not linear*.
- The recurrence relation $H_n = 2H_{n-1} + 1$ is *not homogeneous*.

Characteristic Equations

Fix a linear homogeneous recurrence of degree k:

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k} \quad (\star)$$

If we try a solution of the form $a_n = r^n$ (with $r \neq 0$), substituting into $(\star)$ and dividing by r^{n-k} yields:

Definition 26: The *characteristic equation* of $(\star)$ is:

$$r^k - c_1 r^{k-1} - c_2 r^{k-2} - \ldots - c_k = 0$$

Its solutions $r_1, r_2, \ldots$ are the *characteristic roots* of the recurrence. Each root yields a solution $a_n = r^n$.

Why guess r^n? Plug $a_n = r^n$ into the recurrence:

$$r^n = c_1 r^{n-1} + c_2 r^{n-2} + \ldots + c_k r^{n-k}$$

Divide by r^{n-k} (assuming $r \neq 0$) and you get exactly the characteristic equation. So r^n is a solution precisely when r is a root, and by linearity any linear combination of such solutions works too.

Distinct Roots Case

Theorem 21: Let c_1 and c_2 be real numbers. Suppose that $r^2 - c_1 r - c_2 = 0$ has two *distinct* roots r_1 and r_2. Then the sequence (a_n) is a solution of the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ if and only if $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$ for $n = 0, 1, 2, \ldots$, where α_1 and α_2 are constants.

Proof: We prove both directions.

($\Rightarrow$) Suppose $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$. We verify it satisfies $a_n = c_1 a_{n-1} + c_2 a_{n-2}$. Since r_i is a root of $r^2 - c_1 r - c_2 = 0$, we have $r_i^2 = c_1 r_i + c_2$. Multiply both sides by r_i^{n-2} to get $r_i^n = c_1 r_i^{n-1} + c_2 r_i^{n-2}$. This holds for $i = 1$ and $i = 2$ separately, so taking a linear combination preserves equality:

$$\begin{aligned} c_1 a_{n-1} + c_2 a_{n-2} &= c_1(\alpha_1 r_1^{n-1} + \alpha_2 r_2^{n-1}) + c_2(\alpha_1 r_1^{n-2} + \alpha_2 r_2^{n-2}) \\ &= \alpha_1(c_1 r_1^{n-1} + c_2 r_1^{n-2}) + \alpha_2(c_1 r_2^{n-1} + c_2 r_2^{n-2}) \\ &= \alpha_1 r_1^n + \alpha_2 r_2^n = a_n. \end{aligned}$$

Distinct Roots Case [2]

(⇐) Suppose a_n satisfies the recurrence. We show a_n *must* be of the form $\alpha_1 r_1^n + \alpha_2 r_2^n$.
The recurrence and initial conditions a_0, a_1 determine the entire sequence: each a_n is computed from a_{n-1} and a_{n-2}. So it suffices to find α_1, α_2 such that $\alpha_1 r_1^n + \alpha_2 r_2^n$ matches the two initial values:

$$\begin{cases} \alpha_1 + \alpha_2 = a_0 \\ \alpha_1 r_1 + \alpha_2 r_2 = a_1 \end{cases}$$

This is a 2×2 linear system with determinant $r_2 - r_1$. Since $r_1 \neq r_2$, the determinant is nonzero, so a unique solution (α_1, α_2) exists. The sequence $\alpha_1 r_1^n + \alpha_2 r_2^n$ matches the initial conditions and satisfies the recurrence (by the first direction), so it *equals* a_n by uniqueness. □

Worked Example

Example: Solve $a_n = a_{n-1} + 2a_{n-2}$ with $a_0 = 2$ and $a_1 = 7$.

- The *characteristic* equation is $r^2 - r - 2 = 0$.
- It has two *distinct* roots $r_1 = 2$ and $r_2 = -1$.
- The sequence (a_n) is a solution iff $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$ for $n = 0, 1, 2, \ldots$ and some constants α_1 and α_2.

$$\begin{cases} a_0 = 2 = \alpha_1 + \alpha_2 \\ a_1 = 7 = \alpha_1 \cdot 2 + \alpha_2 \cdot (-1) \end{cases}$$

- Solving these two equations gives $\alpha_1 = 3$ and $\alpha_2 = -1$.
- Hence, the *solution* to the recurrence equation with given initial conditions is the sequence (a_n) with

$$a_n = 3 \cdot 2^n - (-1)^n$$

Fibonacci Numbers

Example: The recurrence $F_n = F_{n-1} + F_{n-2}$ has characteristic equation $r^2 - r - 1 = 0$ with roots:

$$\varphi = \frac{1+\sqrt{5}}{2} \approx 1.618, \quad \psi = \frac{1-\sqrt{5}}{2} \approx -0.618.$$

The general solution is $F_n = \alpha_1 \varphi^n + \alpha_2 \psi^n$ for constants α_1, α_2.

Using initial conditions $F_0 = 0$ and $F_1 = 1$:

$$\begin{cases} \alpha_1 + \alpha_2 = 0 \\ \alpha_1 \varphi + \alpha_2 \psi = 1 \end{cases}$$

Binet's Formula

Solving for α_1, α_2 yields $\alpha_1 = 1/\sqrt{5}$ and $\alpha_2 = -1/\sqrt{5}$, giving *Binet's formula*:

$$F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}$$

Since $|\psi| \approx 0.618 < 1$, the term $\psi^n \to 0$, so F_n is the nearest integer to $\varphi^n/\sqrt{5}$.

In particular, $F_{n+1}/F_n \to \varphi$ as $n \to \infty$.

n	1	2	3	4	5	6	7	8
F_n	1	1	2	3	5	8	13	21
F_{n+1}/F_n	1.000	2.000	1.500	1.667	1.600	1.625	1.615	1.619

Eigenvalues and the Companion Matrix

Definition 27: Let A be an $m \times m$ matrix. A scalar λ is an *eigenvalue* of A if there exists a nonzero vector $\vec{v}$ (called an *eigenvector*) such that

$$A\vec{v} = \lambda\vec{v}$$

The eigenvalues are the roots of the *characteristic polynomial* $p(\lambda) = \det(A - \lambda I)$.

Why does this matter for recurrences? Define the *state vector* $\vec{v}_n = (a_n, a_{n-1}, ..., a_{n-k+1})^T$. Then the recurrence $a_n = c_1 a_{n-1} + ... + c_k a_{n-k}$ is equivalent to $\vec{v}_n = C\vec{v}_{n-1}$, where C is the *companion matrix*:

$$C = \begin{pmatrix} c_1 & c_2 & c_3 & ... & c_k \\ 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 \end{pmatrix}$$

Eigenvalues and the Companion Matrix [2]

Iterating gives $\vec{v}_n = C^n \vec{v}_0$, so computing a_n reduces to *matrix exponentiation*. The characteristic polynomial of C is exactly $p(\lambda) = \lambda^k - c_1\lambda^{k-1} - \ldots - c_k$, *i.e.* the same characteristic equation we solve for the recurrence.

For $k = 2$: $C = \begin{pmatrix} c_1 & c_2 \\ 1 & 0 \end{pmatrix}$, and $\det(C - \lambda I) = \lambda^2 - c_1\lambda - c_2$.

Example: Fibonacci: $C = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, eigenvalues φ, ψ. Diagonalizing $C = PDP^{-1}$ yields $C^n = PD^nP^{-1}$, recovering Binet's formula. Matrix exponentiation computes F_n in $O(\log n)$ multiplications (binary exponentiation), vs. $O(n)$ for naive iteration.

Single (Repeated) Root Case

Theorem 22: Let c_1 and c_2 be real numbers. Suppose that $r^2 - c_1 r - c_2 = 0$ has a *single (repeated)* root r_0. A sequence (a_n) is a solution of the recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ if and only if $a_n = (\alpha_1 + \alpha_2 n)\, r_0^n$ for $n = 0, 1, 2, \ldots$, where α_1 and α_2 are constants.

Why nr_0^n?: When the root is repeated, r_0^n alone only provides one linearly independent solution. The second solution nr_0^n arises by taking the *derivative* of r^n with respect to r and then evaluating at $r = r_0$:

$$\frac{\mathrm{d}}{\mathrm{d}r}(r^n) = nr^{n-1} \quad \Longrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}r}(r^n) \cdot r = nr^n$$

This is analogous to the resonance term $te^{\lambda t}$ in differential equations.

Single (Repeated) Root Case [2]

Example: Solve $a_n = 6a_{n-1} - 9a_{n-2}$ with $a_0 = 1$ and $a_1 = 6$.

The characteristic equation is $r^2 - 6r + 9 = 0$ with a single (repeated) root $r_0 = 3$.

Hence, the solution is of the form $a_n = (\alpha_1 + \alpha_2 n)\, 3^n$.

$$\begin{cases} a_0 = 1 = \alpha_1 \\ a_1 = 6 = (\alpha_1 + \alpha_2) \cdot 3 \end{cases} \implies \begin{cases} \alpha_1 = 1 \\ \alpha_2 = 1 \end{cases}$$

Thus, the *solution* is $a_n = (1 + n)\, 3^n$.

General Case

Theorem 23: Let $c_1, ..., c_k$ be real numbers. Suppose that the recurrence $a_n = c_1 a_{n-1} + ... + c_k a_{n-k}$ has t distinct characteristic roots $r_1, ..., r_t$ with multiplicities $m_1, ..., m_t$, respectively, so that $m_i \geq 1$ for $i = 1, ..., t$ and $m_1 + ... + m_t = k$.

Then the general solution of the recurrence relation is given by

$$a_n = \sum_{i=1}^{t} \left(r_i^n \cdot \sum_{j=0}^{m_i - 1} \alpha_{i,j} n^j \right)$$

where $\alpha_{i,j}$ are constants for $1 \leq i \leq t$ and $0 \leq j \leq m_i - 1$.

Example: Find generic solution for $a_n = 7a_{n-1} - 16a_{n-2} + 12a_{n-3}$.

The characteristic equation $r^3 - 7r^2 + 16r - 12 = 0$ has $t = 2$ distinct roots: $r_1 = 2$ repeated $m_1 = 2$ times, and $r_2 = 3$ with multiplicity $m_2 = 1$. Hence, the solution is of the form

$$a_n = \left(\alpha_{1,0} + \alpha_{1,1} n\right) 2^n + \alpha_{2,0}\, 3^n$$

Complex Roots

When the characteristic equation has complex roots, they come in *conjugate pairs* $\rho e^{\pm i\theta}$ (for recurrences with real coefficients).

Definition 28: For conjugate roots $\rho e^{\pm i\theta}$, the real general solution is:

$$a_n = \rho^n(\alpha \cos(n\theta) + \beta \sin(n\theta))$$

This follows from Euler's formula: $\rho^n e^{in\theta} = \rho^n(\cos(n\theta) + i\sin(n\theta))$.

Complex roots produce *oscillating exponential* behavior: magnitude ρ^n grows (or decays), while $\cos(n\theta)$ and $\sin(n\theta)$ produce periodic oscillation with period $2\pi/\theta$.

Complex Roots Example

Example: Solve $a_n = 2a_{n-1} - 2a_{n-2}$ with $a_0 = 1$ and $a_1 = 1$.

Characteristic equation $r^2 - 2r + 2 = 0$, roots $r = 1 \pm i$.

Polar form: $1 + i = \sqrt{2}e^{i\pi/4}$, so $\rho = \sqrt{2}$ and $\theta = \pi/4$.

General solution: $a_n = \left(\sqrt{2}\right)^n (\alpha \cos(n\pi/4) + \beta \sin(n\pi/4))$.

Initial conditions: $a_0 = 1 \to \alpha = 1$ and $a_1 = 1 \to \beta = 0$.

Therefore $a_n = \left(\sqrt{2}\right)^n \cos(n\pi/4)$.

Root Types Summary

For $a_n = c_1 a_{n-1} + c_2 a_{n-2}$, the solution method is always the same four steps. Only the *form* depends on the roots.

Step	Distinct Real	Repeated	Complex
1. Char. eq.	$r^2 - c_1 r - c_2 = 0$	$r^2 - c_1 r - c_2 = 0$	$r^2 - c_1 r - c_2 = 0$
2. Roots	$r_1 \neq r_2$	r_0 (multiplicity 2)	$\rho e^{\pm i\theta}$
3. Form	$\alpha_1 r_1^n + \alpha_2 r_2^n$	$(\alpha_1 + \alpha_2 n) r_0^n$	$\rho^n(\alpha \cos n\theta + \beta \sin n\theta)$
4. Fit ICs	2 equations	2 equations	2 equations

Recurrences and Differential Equations

Linear recurrences are *discrete analogues* of linear differential equations. The entire theory mirrors the continuous case.

Aspect	Differential Equation	Recurrence
1st order	$y' = cy \to y_0 e^{ct}$	$a_n = ca_{n-1} \to a_0 c^n$
2nd order	$y'' + py' + qy = 0$	$a_n + pa_{n-1} + qa_{n-2} = 0$
Char. eq.	$\lambda^2 + p\lambda + q = 0$	$r^2 + pr + q = 0$
Double root	$(c_1 + c_2 t)e^{\lambda t}$	$(\alpha_1 + \alpha_2 n) r_0^n$
Particular sol.	Undetermined coefficients	Same method
Resonance	$te^{\lambda t}$ factors	nr_0^n factors

Solving Linear Recurrences

Step	Homogeneous	Non-homogeneous
1. Classify	No $F(n)$ term?	Has $F(n)$ term?
2. Char. eq.	$r^k - c_1 r^{k-1} - \ldots - c_k = 0$	Same (ignore $F(n)$)
3. Roots	Distinct / repeated / complex	Same
4. Homog. sol.	$a_n^{(\mathrm{h})}$ from root type	Same
5. Partic. sol.	—	Guess form matching $F(n)$
6. General	$a_n = a_n^{(\mathrm{h})}$	$a_n = a_n^{(\mathrm{h})} + a_n^{(\mathrm{p})}$
7. Fit ICs	Solve linear system	Solve linear system

Particular solution cheat sheet:

- $F(n) =$ polynomial of degree $d \Longrightarrow$ try $c_0 + c_1 n + \ldots + c_d n^d$
- $F(n) = c \cdot r^n$, r not a root $\Longrightarrow$ try βr^n

Solving Linear Recurrences [2]

- $F(n) = c \cdot r^n$, r is a root (mult. m) $\Longrightarrow$ try $\beta n^m r^n$ (resonance)

Linear Non-Homogeneous Recurrence Relations

Definition 29: A *linear non-homogeneous* recurrence relation *of degree k* with constant coefficients is a recurrence relation of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + ... + c_k a_{n-k} + F(n),$$

where $c_1, c_2, ..., c_k$ are constants and $F(n) \neq 0$ is a function of n (the *forcing term*).

The *associated homogeneous recurrence* is the relation with $F(n)$ removed:

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + ... + c_k a_{n-k}$$

Example:

- $a_n = 3a_{n-1} + 2n$ — non-homogeneous, $F(n) = 2n$ (polynomial).
- $a_n = 2a_{n-1} + 3^n$ — non-homogeneous, $F(n) = 3^n$ (exponential).
- $H_n = 2H_{n-1} + 1$ — non-homogeneous, $F(n) = 1$ (constant).

Solving Non-Homogeneous Recurrence Relations

Theorem 24: If $(a_n^{(\mathrm{p})})$ is a *particular* solution of the non-homogeneous recurrence $a_n = c_1 a_{n-1} + \ldots + c_k a_{n-k} + F(n)$, then *every solution* is of the form $a_n = a_n^{(\mathrm{p})} + a_n^{(\mathrm{h})}$, where $a_n^{(\mathrm{h})}$ is a solution of the associated homogeneous recurrence.

The recipe:

1. Solve the homogeneous part.
2. Find a particular solution by guessing a form matching $F(n)$.
3. Add them.
4. Fit initial conditions.

Example: Find all solutions of $a_n = 3a_{n-1} + 2n$ with $a_1 = 3$.

Homogeneous part: $a_n^{(\mathrm{h})} = \alpha 3^n$.

Particular solution: $F(n) = 2n$ is degree-1 polynomial, so try $p_n = cn + d$.
Substituting: $cn + d = 3(c(n-1) + d) + 2n$.
Equating coefficients: $c = -1$, $d = -\frac{3}{2}$. So $a_n^{(\mathrm{p})} = -n - \frac{3}{2}$.

Solving Non-Homogeneous Recurrence Relations [2]

General solution: $a_n = a_n^{(\mathrm{p})} + a_n^{(\mathrm{h})} = -n - \frac{3}{2} + \alpha 3^n$.

With $a_1 = 3$: $\alpha = \frac{11}{6}$, giving $a_n = -n - \frac{3}{2} + \left(\frac{11}{6}\right)3^n$.

Exponential Forcing Term

When $F(n) = c \cdot r^n$, try $a_n^{(\mathrm{p})} = \beta r^n$.

Warning: resonance. If r is already a characteristic root, the trial βr^n fails. Instead, multiply by n: try $a_n^{(\mathrm{p})} = \beta n r^n$ (or n^m if r has multiplicity m).

Example: Solve $a_n = 2a_{n-1} + 3^n$ with $a_0 = 1$.

- *Homogeneous solution:* $a_n^{(\mathrm{h})} = \alpha 2^n$.
- *Particular solution.* Since 3 is *not* a characteristic root, try $a_n^{(\mathrm{p})} = \beta \cdot 3^n$:

$$\beta 3^n = 2\beta 3^{n-1} + 3^n \quad \Longrightarrow \quad \beta = 2\frac{\beta}{3} + 1 \quad \Longrightarrow \quad \beta = 3$$

  So $a_n^{(\mathrm{p})} = 3^{n+1}$.
- *General solution:* $a_n = \alpha 2^n + 3^{n+1}$.
- Using $a_0 = 1$: $1 = \alpha + 3$, so $\alpha = -2$.
- The *solution* is $a_n = 3^{n+1} - 2^{n+1}$.

Recurrences and Generating Functions

Both characteristic equations and generating functions solve the same recurrences, but differently.

Key idea. A generating function encodes the *entire sequence* (a_n) as coefficients of a power series $A(x) = \sum a_n x^n$. The recurrence turns into an *algebraic equation* for $A(x)$.

Example: *Derive* the generating function for Fibonacci: $F_n = F_{n-1} + F_{n-2}$, $F_0 = 0$, $F_1 = 1$.

Write $G(x) = F_0 + F_1 x + F_2 x^2 + F_3 x^3 + \ldots = x + x^2 + 2x^3 + 3x^4 + \ldots$.

Multiply by x and x^2 to shift indices:

$$\begin{aligned} xG(x) &= F_0 x + F_1 x^2 + F_2 x^3 + \ldots \\ x^2 G(x) &= \qquad F_0 x^2 + F_1 x^3 + \ldots \end{aligned}$$

Since $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$, each coefficient of $G - xG - x^2 G$ vanishes except $F_1 x = x$:

$$G(x) - xG(x) - x^2 G(x) = x \quad \Longrightarrow \quad G(x) = \frac{x}{1 - x - x^2}$$

Recurrences and Generating Functions [2]

Now factor the denominator. The roots of $1 - x - x^2 = 0$ are:

$$x = \frac{1 \pm \sqrt{5}}{-2} \quad \Longrightarrow \quad x_1 = -\frac{1}{\varphi}, \quad x_2 = -\frac{1}{\psi}$$

where $\varphi = \frac{1+\sqrt{5}}{2}$, $\psi = \frac{1-\sqrt{5}}{2}$.

The denominator $1 - x - x^2$ and the characteristic polynomial $r^2 - r - 1$ differ only by the substitution $x \to \frac{1}{r}$. Setting the denominator to zero gives values of x where the series *diverges* — these are the *poles* of $G(x)$. The characteristic roots r_1, r_2 are the *reciprocals* of these poles.

Recurrences and Generating Functions [3]

Partial fraction decomposition of $G(x)$ then produces exactly Binet's formula:

$$G(x) = \frac{x}{1 - x - x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1 - \varphi x} - \frac{1}{1 - \psi x}\right) = \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty} (\varphi^n - \psi^n) x^n$$

Why two methods?

- **Characteristic equations** are mechanical and fast for constant-coefficient recurrences.
- **Generating functions** reveal combinatorial structure, handle variable coefficients, produce closed forms via partial fractions — at the cost of more algebra.

Summary: Recurrence Relations

Solution methods

1. *Telescoping* — unwind first-order recurrences
2. *Characteristic equations* — for constant-coefficient linear recurrences
3. *Generating functions* — encode the whole sequence algebraically

Recipe: characteristic equation

Polynomial ⇝ roots ⇝ general solution (distinct / repeated / complex) ⇝ fit initial conditions.

General solution forms

- Distinct: $\alpha_1 r_1^n + \alpha_2 r_2^n$
- Repeated: $(\alpha_1 + \alpha_2 n) r_0^n$
- Complex: $\rho^n(\alpha \cos n\theta + \beta \sin n\theta)$
- Non-homogeneous: $a_n = a_n^{(\mathrm{h})} + a_n^{(\mathrm{p})}$

Connections

- Companion matrix, eigenvalues ⇝ linear algebra
- Same structure as linear differential equations
- Characteristic roots = reciprocals of GF poles

Annihilators

Operators

Definition 30: An *operator* takes a function and produces another function.

The three operators we need are:

- *Sum*: $(f+g)(n) \coloneqq f(n)+g(n)$
- *Scale*: $(\alpha \cdot f)(n) \coloneqq \alpha \cdot f(n)$
- *Shift*: $(\mathbf{E}\, f)(n) \coloneqq f(n+1)$

Examples:

- *Linearity*: $\mathbf{E}(f+g) = \mathbf{E}\, f + \mathbf{E}\, g$, and $\mathbf{E}(3f) = 3 \cdot \mathbf{E}\, f$.
- *Composition*: $(\mathbf{E}-2)f \coloneqq \mathbf{E}\, f - 2f$.
- *Powers*: $\mathbf{E}^2\, f(n) = f(n+2)$, $\mathbf{E}^k\, f(n) = f(n+k)$.
- *Products*: $(\mathbf{E}-2)^2 = \mathbf{E}^2 - 4\,\mathbf{E} + 4$, $(\mathbf{E}-1)(\mathbf{E}-2) = \mathbf{E}^2 - 3\,\mathbf{E} + 2$.

Applying Operators

Examples: Below are the results of applying different operators to $f(n) = 2^n$:

$$
\begin{aligned}
2f(n) &= 2 \cdot 2^n = 2^{n+1} \\
3f(n) &= 3 \cdot 2^n \\
\mathbf{E}\, f(n) &= 2^{n+1} \\
\mathbf{E}^2\, f(n) &= 2^{n+2} \\
(\mathbf{E} - 2) f(n) &= \mathbf{E}\, f(n) - 2f(n) = 2^{n+1} - 2^{n+1} = 0 \\
(\mathbf{E}^2 - 1) f(n) &= \mathbf{E}^2\, f(n) - f(n) = 2^{n+2} - 2^n = 3 \cdot 2^n
\end{aligned}
$$

Compound Operators

The compound operators can be seen as polynomials in "variable" $\mathbf{E}$.

Example: The compound operators $\mathbf{E}^2 - 3\,\mathbf{E} + 2$ and $(\mathbf{E}-1)(\mathbf{E}-2)$ are equivalent:

$$
\begin{aligned}
\text{Let } g(n) &:= (\mathbf{E}-2)f(n) = f(n+1) - 2f(n) \\
\text{Then } (\mathbf{E}-1)(\mathbf{E}-2)f(n) &= (\mathbf{E}-1)g(n) \\
&= g(n+1) - g(n) \\
&= [f(n+2) - 2f(n+1)] - [f(n+1) - 2f(n)] \\
&= f(n+2) - 3f(n+1) + 2f(n) \\
&= (\mathbf{E}^2 - 3\,\mathbf{E} + 2)f(n) \quad \checkmark
\end{aligned}
$$

Operators Summary

Operator	Definition
addition	$(f+g)(n) \coloneqq f(n)+g(n)$
subtraction	$(f-g)(n) \coloneqq f(n)-g(n)$
multiplication	$(\alpha \cdot f)(n) \coloneqq \alpha \cdot f(n)$
shift	$\mathbf{E}\, f(n) \coloneqq f(n+1)$
k-fold shift	$\mathbf{E}^k\, f(n) \coloneqq f(n+k)$
composition	$(\boldsymbol{X}+\boldsymbol{Y})f \coloneqq \boldsymbol{X}\, f+\boldsymbol{Y}\, f$
	$(\boldsymbol{X}-\boldsymbol{Y})f \coloneqq \boldsymbol{X}\, f-\boldsymbol{Y}\, f$
	$\boldsymbol{X}\,\boldsymbol{Y}\, f \coloneqq \boldsymbol{X}(\boldsymbol{Y}\, f)=\boldsymbol{Y}(\boldsymbol{X}\, f)$
distribution	$\boldsymbol{X}(f+g)=\boldsymbol{X}\, f+\boldsymbol{X}\, g$

Annihilators

Definition 31: An *annihilator* of a function f is any non-trivial operator X such that $X\,f = 0$.

Intuition: An annihilator "kills" the function — after applying the operator, nothing remains. This is the same idea as finding roots of a polynomial: if $(x-2)$ is a factor, then plugging in $x=2$ makes the polynomial zero. Here, $(\mathbf{E}-a)$ is a factor, and applying it to a^n makes the result zero.

Examples:

- $(\mathbf{E}-1)$ annihilates any constant $f(n)=c$: $(\mathbf{E}-1)c = c-c = 0$.
- $(\mathbf{E}-2)$ annihilates $f(n)=\alpha\cdot 2^n$: $(\mathbf{E}-2)(\alpha 2^n) = \alpha 2^{n+1} - 2\alpha 2^n = 0$.
- $(\mathbf{E}-1)^2$ annihilates any linear $f(n)=\alpha n+\beta$: $(\mathbf{E}-1)(\alpha n+\beta) = \alpha$ (a constant), and $(\mathbf{E}-1)\alpha = 0$.

Annihilators Summary

Operator	Functions annihilated
$\mathbf{E}-1$	α
$\mathbf{E}-a$	αa^n
$(\mathbf{E}-a)(\mathbf{E}-b)$	$\alpha a^n + \beta b^n \quad [\text{if } a \neq b]$
$(\mathbf{E}-a_0)(\mathbf{E}-a_1)...(\mathbf{E}-a_k)$	$\sum_{i=0}^{k} \alpha_i a_i^n \quad [\text{if } a_i \text{ are distinct}]$
$(\mathbf{E}-1)^2$	$\alpha n + \beta$
$(\mathbf{E}-a)^2$	$(\alpha n + \beta)a^n$
$(\mathbf{E}-a)^2(\mathbf{E}-b)$	$(\alpha n + \beta)a^n + \gamma b^n \quad [\text{if } a \neq b]$
$(\mathbf{E}-a)^d$	$\left(\sum_{i=0}^{d-1} \alpha_i n^i\right) a^n$

Properties of Annihilators

Theorem 25: If $\boldsymbol{X}$ annihilates f, then $\boldsymbol{X}$ also annihilates αf for any constant α.

Proof: Since operators are linear (they distribute over scaling), $\boldsymbol{X}(\alpha f) = \alpha \cdot \boldsymbol{X}\, f = \alpha \cdot 0 = 0$. □

Theorem 26: If $\boldsymbol{X}$ annihilates both f and g, then $\boldsymbol{X}$ also annihilates $f \pm g$.

Proof: By linearity of operators: $\boldsymbol{X}(f \pm g) = \boldsymbol{X}\, f \pm \boldsymbol{X}\, g = 0 \pm 0 = 0$. □

Theorem 27: If $\boldsymbol{X}$ annihilates f, then $\boldsymbol{X}$ also annihilates $\mathbf{E}\, f$.

Proof: Any operator $\boldsymbol{X}$ that is a polynomial in $\mathbf{E}$ *commutes* with $\mathbf{E}$ itself.

Therefore $\boldsymbol{X}(\mathbf{E}\, f) = \mathbf{E}(\boldsymbol{X}\, f) = \mathbf{E}(0) = 0$. □

Properties of Annihilators [2]

Theorem 28: If $\boldsymbol{X}$ annihilates f and $\boldsymbol{Y}$ annihilates g, then $\boldsymbol{X}\,\boldsymbol{Y}$ annihilates $f \pm g$.

Proof: Since $\boldsymbol{X}$ annihilates f, and by the previous theorem $\boldsymbol{X}$ also annihilates $\mathbf{E}\,f$, $\mathbf{E}^2\,f$, *etc.*, the operator $\boldsymbol{X}$ annihilates *any* function obtained by applying shift-operators to f. In particular, $\boldsymbol{X}$ annihilates $\boldsymbol{Y}\,f$. Now apply $\boldsymbol{X}\,\boldsymbol{Y}$ to $f \pm g$:

$$\boldsymbol{X}\,\boldsymbol{Y}(f \pm g) = \boldsymbol{X}(\boldsymbol{Y}\,f \pm \boldsymbol{Y}\,g) = \boldsymbol{X}(\boldsymbol{Y}\,f) \pm \boldsymbol{X}(\boldsymbol{Y}\,g)$$

The first term $\boldsymbol{X}(\boldsymbol{Y}\,f)$: since $\boldsymbol{X}$ annihilates f and commutes with $\boldsymbol{Y}$, this is $\boldsymbol{Y}(\boldsymbol{X}\,f) = \boldsymbol{Y}(0) = 0$. The second term $\boldsymbol{X}(\boldsymbol{Y}\,g)$: since $\boldsymbol{Y}$ annihilates g, we get $\boldsymbol{X}(0) = 0$. Both terms are zero, so the result is zero. $\square$

Annihilating Recurrences

The annihilator method is a systematic five-step procedure for solving recurrences.

1. *Rewrite* the recurrence in operator form (replace a_n with $\mathbf{E}$-operators).
2. *Identify* the annihilator of both sides and multiply additional factors to kill the residue.
3. *Factor* the annihilator into linear factors $(\mathbf{E} - a)$.
4. *Read off* the generic solution from the factorization (one term per factor, using the summary table).
5. *Fit* the initial conditions to determine the free constants.

Example: $r(n) = 5r(n-1)$ with $r(0) = 3$.

Step 1. $r(n+1) - 5r(n) = 0$

$(\mathbf{E} - 5)r(n) = 0$

Step 2. $(\mathbf{E} - 5)$ annihilates $r(n)$.

Step 3. $(\mathbf{E} - 5)$ is already factored.

Step 4. $r(n) = \alpha 5^n$ is a generic solution.

Step 5. $r(0) = \alpha = 3 \implies \alpha = 3$

Thus, $r(n) = 3 \cdot 5^n$.

Annihilating Recurrences [2]

Example: $T(n) = 2T(n-1) + 1$ with $T(0) = 0$

1. $T(n+1) - 2T(n) = 1$
   $(\mathbf{E} - 2)T(n) = 1$
2. $(\mathbf{E} - 2)$ does *not* annihilate $T(n)$: the residue is 1.
   $(\mathbf{E} - 1)$ annihilates the residue 1.
   Thus, $(\mathbf{E} - 1)(\mathbf{E} - 2)$ annihilates $T(n)$.
3. $(\mathbf{E} - 1)(\mathbf{E} - 2)$ is already factored.
4. $T(n) = \alpha 2^n + \beta$ is a generic solution.
5. Find the coefficients α, β using $T(0) = 0$ and $T(1) = 2T(0) + 1 = 1$:

$$\left.\begin{aligned} T(0) = 0 &= \alpha \cdot 2^0 + \beta \\ T(1) = 1 &= \alpha \cdot 2^1 + \beta \end{aligned}\right\} \implies \begin{cases} \alpha = 1 \\ \beta = -1 \end{cases}$$

Thus, $T(n) = 2^n - 1$.

Annihilating Recurrences [3]

Example: $T(n) = T(n-1) + 2T(n-2) + 2^n - n^2$

1. Operator form:
   $(\mathbf{E}^2 - \mathbf{E} - 2)T(n) = \mathbf{E}^2(2^n - n^2)$
2. Annihilator:
   $(\mathbf{E}^2 - \mathbf{E} - 2)(\mathbf{E} - 2)(\mathbf{E} - 1)^3$
3. Factorization:
   $(\mathbf{E} + 1)(\mathbf{E} - 2)^2(\mathbf{E} - 1)^3$
4. Generic solution:
   $T(n) = \alpha(-1)^n + (\beta n + \gamma)2^n + \delta n^2 + \varepsilon n + \zeta$
5. There are no initial conditions. We can only provide an asymptotic bound.

Thus, $T(n) \in \Theta(n2^n)$

Annihilators and Characteristic Equations

Note: The characteristic equation method and the annihilator method are *the same algebra* in different notation. Replacing r with $\mathbf{E}$ in the characteristic polynomial gives the annihilator, and vice versa. Understanding both gives you flexibility: use whichever is more convenient.

Theorem 29: For the linear homogeneous recurrence $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}$:

- **Characteristic polynomial:** $p(r) = r^k - c_1 r^{k-1} - \ldots - c_k$
- **Annihilator polynomial:** $p(\mathbf{E}) = \mathbf{E}^k - c_1\,\mathbf{E}^{k-1} - \ldots - c_k$

If $r_1, \ldots, r_t$ are the distinct characteristic roots with multiplicities $m_1, \ldots, m_t$, then

$$p(r) = (r - r_1)^{m_1} \ldots (r - r_t)^{m_t} = p(\mathbf{E})$$

The factorization is identical; only the interpretation differs.

Proof: The recurrence $a_n - c_1 a_{n-1} - \ldots - c_k a_{n-k} = 0$ can be written as $\left(\mathbf{E}^k - c_1\,\mathbf{E}^{k-1} - \ldots - c_k\right) a_n = 0$. This is the *annihilator polynomial* applied to a_n.

Annihilators and Characteristic Equations [2]

The characteristic equation is $r^k - c_1 r^{k-1} - \ldots - c_k = 0$.

Both polynomials have the same coefficients — the only difference is the symbol (r vs. $\mathbf{E}$). Therefore they have the same roots and the same factorization. □

Example: $a_n = 5a_{n-1} - 6a_{n-2}$.

- **Characteristic:** $r^2 - 5r + 6 = (r-2)(r-3) = 0$.
- **Annihilator:** $(\mathbf{E}-2)(\mathbf{E}-3)a_n = 0$.
- Both give: $a_n = \alpha 2^n + \beta 3^n$.

Common Mistakes

Example: Solve $a_n = 2a_{n-1} - 2^n$ with $a_0 = 1$.

Wrong answer: "Characteristic equation: $r - 2 = 0$, so $r = 2$. General solution: $a_n = c \cdot 2^n$."

Where's the error?

- This is *non-homogeneous* (has extra term "-2^n"), but the student *ignored it*!
- We need a *particular* solution in addition to the homogeneous one.
- Since 2 is *also* the characteristic root (resonance!), the simple trial $\beta 2^n$ fails. We must try $a_n^{(\mathrm{p})} = \beta n \cdot 2^n$.

Correct: $a_n = (\alpha + n)2^n$. Using $a_0 = 1$: $\alpha = 1$, so $a_n = (1 + n)2^n$.

Summary: Annihilators

Core idea. An operator X *annihilates* f if $X\,f = 0$.

The factor $(\mathbf{E} - a)$ kills a^n. Higher powers $(\mathbf{E} - a)^d$ kill $\left(\alpha_0 + \alpha_1 n + \ldots + \alpha_{d-1} n^{d-1}\right) a^n$.

Finding the annihilator of a recurrence *automatically* reveals the form of the solution.

Algorithm

1. Rewrite recurrence in operator form
2. Find the annihilator of both sides
3. Factor into $(\mathbf{E} - a_i)$ terms
4. Read off generic solution from factors
5. Fit initial conditions

Key properties

- If X annihilates f,
  then it also annihilates αf, $f \pm g$, $\mathbf{E}\,f$
- If X annihilates f and Y annihilates g,
  then $X\,Y$ annihilates $f \pm g$
- Same polynomial as characteristic equation, with $r \Rightarrow \mathbf{E}$

Asymptotic Analysis

"A mathematician who is not also something of a poet will never be a complete mathematician."

— Karl Weierstrass

Asymptotic Notation

Definition 32 (*Big-O notation*): $f(n) \in O(g(n))$ means f is *asymptotically bounded from above* by g, up to a constant factor:

$$f(n) \in O(g(n)) \iff \exists c > 0.\, \exists n_0.\, \forall n > n_0 : |f(n)| \leq c \cdot g(n)$$

Definition 33 (*Small-o notation*): $f(n) \in o(g(n))$ means f is *strictly dominated* by g: no constant makes f keep up with g.

$$f(n) \in o(g(n)) \iff \forall c > 0.\, \exists n_0.\, \forall n > n_0 : |f(n)| \leq c \cdot g(n)$$

Note: The only difference is $\exists c$ (Big-O: *some* constant works) vs. $\forall c$ (Small-o: *every* constant works, no matter how small).

Note: Flip $\leq$ to $\geq$ to get the dual notations $f \in \Omega(g)$ (bounded below) and $f \in \omega(g)$ (strictly dominates).

Asymptotic Notation [2]

Definition 34 (*Theta notation*): $f \in \Theta(g)$ iff $f \in O(g)$ and $f \in \Omega(g)$ — f and g grow at the *same rate*, up to constant factors.

Example: Prove that $3n^2 + 5n \in \Theta(n^2)$.

Upper bound: $3n^2 + 5n \leq 3n^2 + 5n^2 = 8n^2$ for $n \geq 1$, so $c = 8$, $n_0 = 1$.

Lower bound: $3n^2 + 5n \geq 3n^2$ for $n \geq 0$, so $c = 3$, $n_0 = 0$.

Both bounds hold, so $3n^2 + 5n \in \Theta(n^2)$.

Limits

Notation	Name	Description	Limit definition
$f \in o(g)$	Small Oh	f is dominated by g	$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$
$f \in O(g)$	Big Oh	f is bounded above by g	$\limsup_{n\to\infty} \frac{\lvert f(n)\rvert}{g(n)} < \infty$
$f \sim g$	Equivalence	f is asympotically equal to g	$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$
$f \in \Omega(g)$	Big Omega	f is bounded below by g	$\liminf_{n\to\infty} \frac{f(n)}{g(n)} > 0$
$f \in \omega(g)$	Small Omega	f dominates g	$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty$

Asymptotic Equivalence

Definition 35: $f \sim g$ means f and g are *asymptotically equivalent* — their ratio tends to 1:

$$f \sim g \iff \lim_{n \to \infty} \frac{f(n)}{g(n)} = 1$$

Example: $n^2 + 3n \sim n^2$, since $\frac{n^2+3n}{n^2} = 1 + \frac{3}{n} \to 1$.

Note: $f \sim g$ is *stronger* than $f \in \Theta(g)$: $2n \in \Theta(n)$ but $2n \sim n$ is false (the ratio is 2, not 1).

Properties of Asymptotics

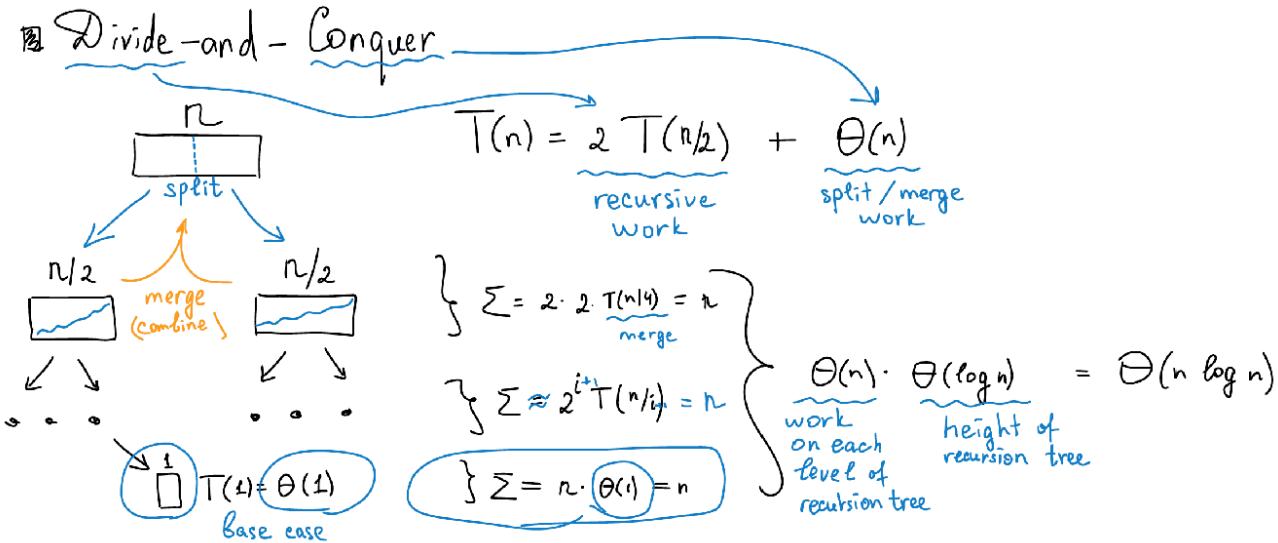
Theorem 30: The asymptotic notations satisfy:

$$
\begin{aligned}
f \in O(g) \text{ and } f \in \Omega(g) &\iff f \in \Theta(g) \\
f \in O(g) &\iff g \in \Omega(f) \\
f \in o(g) &\iff g \in \omega(f) \\
f \in o(g) &\to f \in O(g) \\
f \in \omega(g) &\to f \in \Omega(g) \\
f \sim g &\to f \in \Theta(g)
\end{aligned}
$$

Note: **Growth rate hierarchy:** $1 \subset o(\ln n) \subset o(n) \subset o(n \ln n) \subset o(n^2) \subset o(2^n) \subset o(n!)$.

Each class is strictly contained in the next: $n^k \in o(c^n)$ for any k and $c > 1$.

Divide-and-Conquer Algorithms Analysis



Divide-and-Conquer Recurrence

$$T(n) = a \cdot T\left(\frac{n}{b}\right) + f(n)$$

- $T(n)$ is the *cost* of the recursive algorithm
- a is the number of *parts* (*sub-problems*)
- n/b is the *size* of each part
- $T\left(\frac{n}{b}\right)$ is the cost of each *sub-problem*
- $f(n)$ is the cost of *splitting* and *merging* the solutions of the subproblems

Hereinafter, $c_{\text{crit}} = \log_b a$ is a *critical constant*.

Master Theorem

The master theorem [2] gives a direct asymptotic bound for recurrences of the form

$$T(n) = a \cdot T\left(\frac{n}{b}\right) + f(n)$$

where $c_{\text{crit}} = \log_b a$ is the *critical exponent*.

Intuition: The recurrence splits into a sub-problems of size n/b. The total work at each level of the recursion tree is a copies of the sub-problem cost, plus the merge cost $f(n)$.

- If $f(n)$ is small relative to $n^{c_{\text{crit}}}$, the recursion *leaves* dominate — total cost is $\Theta(n^{c_{\text{crit}}})$.
- If $f(n)$ is comparable to $n^{c_{\text{crit}}}$, the levels contribute equally — a log factor appears.
- If $f(n)$ is large, the *root level* dominates — total cost is $\Theta(f(n))$.

Master Theorem [2]

Case	Description	Condition	Bound
Case I	"merge" $\ll$ "recursion"	$f(n) \in O(n^c)$, $c < c_{\text{crit}}$	$T(n) \in \Theta(n^{c_{\text{crit}}})$
Case II	"merge" $\approx$ "recursion"	$f(n) \in \Theta\big(n^{c_{\text{crit}}} \log^k n\big)$, $k \geq 0$	$T(n) \in \Theta\big(n^{c_{\text{crit}}} \log^{k+1} n\big)$
Case III	"merge" $\gg$ "recursion"	$f(n) \in \Omega(n^c)$, $c > c_{\text{crit}}$	$T(n) \in \Theta(f(n))$

Note: Case III also requires the *regularity condition* to hold: $af(n/b) \leq kf(n)$ for some constant $k < 1$ and all sufficiently large n.

Examples of Master Theorem Application

Example: $T(n) = 2T(n/4) + n^{0.51}$

$a = 2, b = 4, c_{\text{crit}} = \log_4 2 = \frac{1}{2}$. $f(n) = n^{0.51} \in \Omega(n^c)$ with $c = 0.51 > 0.5 = c_{\text{crit}}$.

Case III: $T(n) \in \Theta(n^{0.51})$.

Example: $T(n) = 5T(n/25) + n^{0.49}$

$a = 5, b = 25, c_{\text{crit}} = \log_{25} 5 = \frac{1}{2}$. $f(n) = n^{0.49} \in O(n^c)$ with $c = 0.49 < 0.5 = c_{\text{crit}}$.

Case I: $T(n) \in \Theta\left(n^{\frac{1}{2}}\right)$.

Example: $T(n) = 3T(n/9) + \sqrt{n}$

$a = 3, b = 9, c_{\text{crit}} = \log_9 3 = \frac{1}{2}$. $f(n) = n^{\frac{1}{2}} = n^{c_{\text{crit}}}$, so $f(n) \in \Theta(n^{c_{\text{crit}}} \log^0 n)$.

Case II ($k = 0$): $T(n) \in \Theta\left(n^{\frac{1}{2}} \log n\right)$.

Akra–Bazzi Method

The Akra–Bazzi method [3] is a *generalization* of the master theorem to recurrences of the form

$$T(n) = f(n) + \sum_{i=1}^{k} a_i T\left(b_i n + \underbrace{h_i(n)}_{*}\right)$$

- k is a constant
- $a_i > 0$
- $0 < b_i < 1$
- $h_i(n) \in O\left(\frac{n}{\log^2 n}\right)$ is a *small perturbation*

Bound of $T(n)$ by Akra–Bazzi method:

$$T(n) \in \Theta\left(n^p \cdot \left(1 + \int_1^n \frac{f(x)}{x^{p+1}}\,\mathrm{d}x\right)\right)$$

where p is the solution for the equation $\sum_{i=1}^{k} a_i b_i^p = 1$

Akra–Bazzi Example

Example: Suppose the runtime of an algorithm is expressed by the following recurrence relation:

$$T(n) = \begin{cases} 1 \text{ for } 0 \le n \le 3 \\ n^2 + \frac{7}{4}T\left(\left\lfloor \frac{1}{2}n \right\rfloor\right) + T\left(\left\lceil \frac{3}{4}n \right\rceil\right) \text{ for } n > 3 \end{cases}$$

- Note that the Master Theorem *is not* applicable here, since there are *two* different recursive terms.
- Let's apply the Akra–Bazzi method. First, solve the equation $\frac{7}{4}\left(\frac{1}{2}\right)^p + \left(\frac{3}{4}\right)^p = 1$. This gives us $p = 2$.
- Next, use the formula from AB-method to obtain the bound:

$$\begin{aligned} T(x) \in \Theta\left(x^p\left(1 + \int_1^x \frac{f(u)}{u^{p+1}}du\right)\right) = \\ = \Theta\left(x^2\left(1 + \int_1^x \frac{u^2}{u^3}du\right)\right) = \\ = \Theta(x^2(1 + \ln x)) = \\ = \Theta(x^2 \log x) \end{aligned}$$

Summary: Asymptotic Analysis

Notations

- $O(g)$: bounded above by g (some constant c)
- $\Omega(g)$: bounded below by g
- $\Theta(g)$: bounded above *and* below — tight
- $o(g)$: strictly dominated (no constant saves you)
- $\omega(g)$: strictly dominates
- $f \sim g$: ratio $\to 1$ (stronger than Θ)

Growth hierarchy: $1 \subset o(\ln n) \subset o(n) \subset o(n \ln n) \subset o(n^2) \subset o(2^n) \subset o(n!)$

Master Theorem

$T(n) = aT(n/b) + f(n)$, where $c_{\text{crit}} = \log_b a$.

Compare $f(n)$ with $n^{c_{\text{crit}}}$:

- Case I: f is smaller $\to \Theta(n^{c_{\text{crit}}})$
- Case II: f is comparable $\to \Theta(n^{c_{\text{crit}}} \log n)$
- Case III: f dominates $\to \Theta(f(n))$

Akra–Bazzi

Generalizes Master Theorem to unequal subproblem sizes. Solve $\sum a_i b_i^p = 1$ for p, then integrate.

Advanced Topics

Why Extend the Factorial?

The factorial $n! = 1 \cdot 2 \cdot \cdots \cdot n$ is defined only for non-negative integers. But many formulas *need* factorial-like values at non-integer points:

- **Generalized binomial coefficients:** $\binom{r}{k}$ for real r requires a continuous version of $r!$.
- **Probability:** the Gaussian integral $\int e^{-t^2}\,dt$ evaluates to a non-integer "factorial" value.
- **Asymptotics:** Stirling's formula $n! \approx \sqrt{2\pi n}(n/e)^n$ comes from the analytic properties of Γ.

How should we *interpolate* $n!$ between the integers? There are infinitely many smooth extensions; we need a principled way to choose.

Euler Integral Definition

Definition 36: The *Gamma function* is defined for $z \in \mathbb{C}$ with $\Re(z) > 0$ by the *Euler integral*:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

The integral converges for all z with $\Re(z) > 0$ because the exponential e^{-t} decays faster than any power of t grows.

Example: $\Gamma(1) = \int_0^\infty e^{-t} dt = 1$. A change of variable $t = s^2$ gives $\Gamma(1/2) = \int_0^\infty e^{-s^2} ds = \sqrt{\pi}$.

The Functional Equation

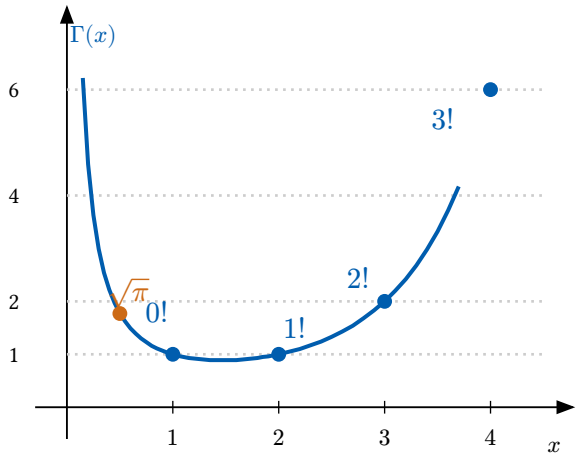
Theorem 31: For all z with $\Re(z) > 0$: $\Gamma(z+1) = z \cdot \Gamma(z)$. In particular, $\Gamma(n+1) = n!$ for every positive integer n.

Proof: Start from $\Gamma(z+1) = \int_0^\infty t^z e^{-t}\,dt$. Integrate by parts with $u = t^z$, $dv = e^{-t}dt$:

$$\Gamma(z+1) = \left[-t^z e^{-t}\right]_0^\infty + z\int_0^\infty t^{z-1}e^{-t}\,dt$$

The boundary term vanishes: at $t = 0$ we get $t^z e^{-t} = 0$ (since $z > 0$), and at $t = \infty$ the exponential decay dominates. Therefore $\Gamma(z+1) = z \cdot \Gamma(z)$. Iterating gives $\Gamma(n+1) = n \cdot \cdots \cdot 1 \cdot \Gamma(1) = n!$, since $\Gamma(1) = \int_0^\infty e^{-t}\,dt = 1$. $\square$

Plot of $\Gamma(x)$



Key values

1. $\Gamma(1) = 1 = 0!$
2. $\Gamma(2) = 1 = 1!$
3. $\Gamma(3) = 2 = 2!$
4. $\Gamma(4) = 6 = 3!$
5. $\Gamma(1/2) = \sqrt{\pi} \approx 1.77$

Behavior

- Pole at $x = 0^+$: $\Gamma(x) \to +\infty$
- Minimum near $x \approx 1.46$: $\Gamma \approx 0.89$
- Poles at all $x = 0, -1, -2, \ldots$

Bohr–Mollerup Theorem

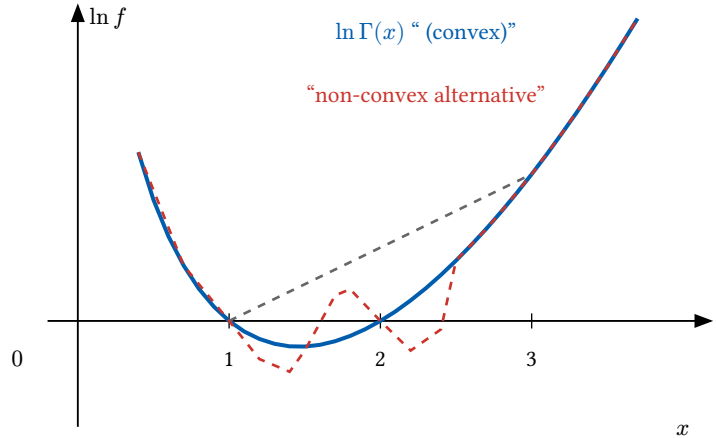
Theorem 32: There is a *unique* function $f : (0, \infty) \to (0, \infty)$ satisfying all three conditions:

1. $f(1) = 1$
2. $f(x + 1) = x \cdot f(x)$ for all $x > 0$
3. $\ln f(x)$ is convex

That unique function is $f(x) = \Gamma(x)$.

The first two conditions alone are not enough: there are infinitely many extensions of $n!$ satisfying the functional equation. Convexity of $\ln \Gamma$ is the natural smoothness requirement, ruling out functions that wiggle wildly between integer points. By Bohr–Mollerup, any function satisfying these three conditions *must be* Γ. This will guarantee that the Gauss limit and Weierstrass product (introduced next) define the same function as the Euler integral.

Convexity of $\ln \Gamma$



The chord (gray dashed) lies *above* $\ln \Gamma$, confirming convexity. The red dashed curve is an alternative extension that satisfies the first two Bohr–Mollerup conditions but is *not* convex.

Gauss's Limit Form

Definition 37: Gauss's *limit form* defines $\Gamma(x)$ for $x > 0$ as:

$$\Gamma(x) := \lim_{n\to\infty} \frac{n! \cdot n^x}{\prod_{k=0}^{n}(x+k)}$$

Theorem 33: The Gauss limit satisfies the same functional equation: $\Gamma(x+1) = x \cdot \Gamma(x)$.

Proof: Write $\Gamma(x+1)$ using the limit:

$$\Gamma(x+1) = \lim_{n\to\infty} \frac{n! \cdot n^{x+1}}{\prod_{k=0}^{n}(x+1+k)}$$

Shift the index: $\prod_{k=0}^{n}(x+1+k) = \prod_{k=1}^{n+1}(x+k)$. Factor out $\Gamma(x)$ and the extra terms:

Gauss's Limit Form [2]

$$= \lim_{n\to\infty} x \cdot \frac{n! \cdot n^x}{\prod_{k=0}^{n}(x+k)} \cdot \frac{n}{x+n+1}$$

As $n \to \infty$, the factor $n/(x+n+1) \to 1$, and the remaining limit is $x \cdot \Gamma(x)$ by definition. $\square$

Weierstrass Product

Definition 38: The *Weierstrass product* defines $1/\Gamma(x)$ as an infinite product:

$$\frac{1}{\Gamma(x)} = x \cdot e^{\gamma x} \cdot \prod_{n=1}^{\infty} \left((1 + x/n) e^{-x/n} \right)$$

where $\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} 1/k - \ln n \right) \approx 0.5772$ is the *Euler–Mascheroni constant*.

This form defines $1/\Gamma(x)$ as an entire function (no poles), so the poles of $\Gamma(x)$ at $x = 0, -1, -2, \ldots$ come from the simple zeros of the product. Since all three definitions satisfy the Bohr–Mollerup conditions, they define the same function.

Equivalence of Definitions

Theorem 34: The Euler integral and Gauss limit definitions agree: for all $x > 0$,

$$\int_0^\infty t^{x-1} e^{-t} dt = \lim_{n \to \infty} \frac{n! \cdot n^x}{\prod_{k=0}^{n} (x+k)}$$

Proof: Define $I_n := \int_0^n t^{x-1}(1 - t/n)^n dt$. As $n \to \infty$, $(1 - t/n)^n \to e^{-t}$, so $I_n \to \Gamma(x)$.

Substitute $u = t/n$:

$$I_n = n^x \int_0^1 u^{x-1}(1-u)^n du = n^x \cdot \mathrm{B}(x, n+1)$$

where $\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt$.

Equivalence of Definitions [2]

Using the Beta–Gamma relation $\mathrm{B}(x,y) = (\Gamma(x)\cdot\Gamma(y))/\Gamma(x+y)$ (itself provable from the Euler integral):

$$I_n = n^x \cdot \frac{\Gamma(x)\cdot\Gamma(n+1)}{\Gamma(x+n+1)} = \frac{n!\cdot n^x}{x\cdot(x+1)\cdot\cdots(x+n)}$$

since $\Gamma(n+1) = n!$ and $\Gamma(x+n+1) = (x+n)(x+n-1)\cdots x\cdot\Gamma(x)$.

Taking $n\to\infty$ and matching both sides completes the proof. $\square$

Euler's Reflection Formula

Theorem 35: For all non-integer x:

$$\Gamma(x) \cdot \Gamma(1-x) = \frac{\pi}{\sin(\pi x)}$$

Example: $\Gamma(1/2) \cdot \Gamma(1/2) = \frac{\pi}{\sin(\pi/2)} = \pi$, so $\Gamma(1/2) = \sqrt{\pi}$. This confirms that the factorial extends to half-integers: $(1/2)! = \Gamma(3/2) = \sqrt{\pi}/2$.

Stirling's Approximation

Theorem 36: As $n \to \infty$:

$$n! \sim \sqrt{2\pi n}(n/e)^n$$

More precisely, $n! = \sqrt{2\pi n}(n/e)^n e^{\theta_n}$ where $1/(12n+1) < \theta_n < 1/(12n)$.

Stirling's approximation lets us estimate factorials and binomial coefficients without computing huge products.

For example, $\binom{2n}{n} \approx 4^n/\sqrt{\pi n}$ follows directly.

Example: Estimate 10! using Stirling: $\sqrt{20\pi}(10/e)^{10} \approx 3.599 \times 10^6$. The exact value is $10! = 3628800 \approx 3.629 \times 10^6$, error less than 1%.

Gamma Function Applications

The Gamma function appears throughout combinatorics and analysis.

- **Generalized binomial coefficients** (for arbitrary $r \in \mathbb{R}$): $\binom{r}{k} = \frac{\Gamma(r+1)}{\Gamma(k+1)\cdot\Gamma(r-k+1)}$, the definition underlying Newton's Binomial Theorem for real exponents.
- **Beta function:** $\mathrm{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^1 t^{x-1}(1-t)^{y-1}dt$.
- **Volume of the n-ball:** $V_n = \pi^{n/2}/\Gamma(n/2+1)$, a natural application of Γ at half-integer arguments.
- **Asymptotics of partitions and Stirling numbers:** leading terms involve Γ through saddle-point methods.

Summary: Gamma Function

Why Γ is unique (Bohr–Mollerup)

1. $\Gamma(1) = 1$
2. $\Gamma(x+1) = x \cdot \Gamma(x)$
3. $\ln \Gamma(x)$ is convex

Three equivalent definitions

- Euler integral: $\int_0^\infty t^{z-1} e^{-t} dt$
- Gauss limit: $\lim(n! n^x)/(\prod(x+k))$
- Weierstrass: $1/\Gamma = x e^{\gamma x} \prod\left((1 + x/n) e^{-x/n}\right)$

Key properties

- Functional equation: $\Gamma(x+1) = x \cdot \Gamma(x)$
- Integer values: $\Gamma(n) = (n-1)!$
- Reflection: $\Gamma(x)\Gamma(1-x) = \pi / \sin(\pi x)$
- $\Gamma(1/2) = \sqrt{\pi}$

Applications

- Extends $n!$ to reals and complexes
- Generalizes $\binom{n}{k}$ to real n
- Stirling: $n! \approx \sqrt{2\pi n}(n/e)^n$

Möbius Inversion

"The opposite of a correct statement is a false statement.
But the opposite of a profound truth may well be another profound truth."

— Niels Bohr

Gian-Carlo Rota

August Ferdinand Möbius

Motivation

Throughout this course we have derived many counting formulas:

- Multiplication principle: k^n
- Permutations: $n!/(n-k)!$
- Combinations: $\binom{n}{k}$
- Stars and bars: $\binom{n+k-1}{n}$

- Stirling numbers: $s_k^{\text{II}}(n)$
- Bell numbers: B_n
- PIE: $\sum (-1)^i \binom{n}{i} (n-i)^k$
- Derangements: D_n

They all seem unrelated.

But many of them are unified by *one* deep pattern: **Möbius inversion on partially ordered sets.**

Posets

Definition 39: A *partially ordered set* (poset) is a pair $\mathcal{P} = (P, \leq)$ where $\leq$ is a reflexive, antisymmetric, and transitive binary relation.

We will work with three key examples:

Boolean lattice

$\left(2^{[n]}, \subseteq\right)$

Elements: subsets of $[n]$.

$A \leq B$ iff $A \subseteq B$.

Divisor lattice

$(\mathbb{Z}^+, |)$

Elements: positive integers.

$a \leq b$ iff $a \mid b$.

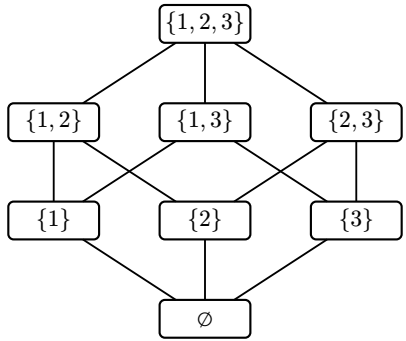
Partition lattice

Π_n

Elements: partitions of $[n]$.

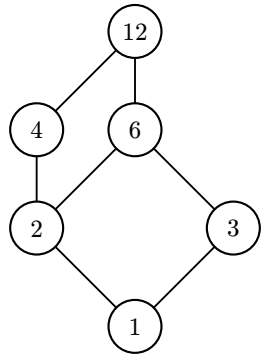
$\pi \leq \sigma$ iff π refines σ.

The Boolean Lattice B_3



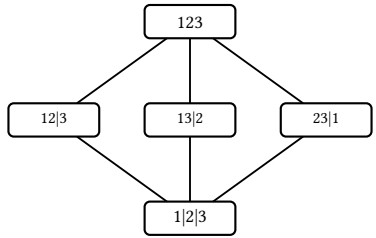
The Boolean lattice $B_3 = \left(2^{[3]}, \subseteq\right)$: elements are subsets of $[3] = \{1, 2, 3\}$, ordered by inclusion.

The Divisor Lattice of 12



The divisor lattice of 12: elements are divisors of $12 = 2^2 \cdot 3$, ordered by divisibility.

The Partition Lattice Π_3



The partition lattice Π_3: elements are partitions of $[3]$, ordered by refinement. $\pi \leq \sigma$ means every block of π is contained in a block of σ.

Stirling number $s_k^{\Pi}(n)$ counts the elements of Π_n with exactly k blocks.

Intervals

Definition 40: Let $\mathcal{P} = (P, \leq)$ be a poset and $x \leq y$. The *interval* $[x, y]$ is the set:

$$[x, y] = \{z \in P \mid x \leq z \leq y\}$$

A poset is *locally finite* if every interval is finite.

Example: In the Boolean lattice $\left(2^{[3]}, \subseteq\right)$:

$$[\{1\}, \{1, 2, 3\}] = \{\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$$

In the divisor lattice of 12: $[2, 12] = \{2, 4, 6, 12\}$

The Incidence Algebra

Definition 41: Let $\mathcal{P} = (P, \leq)$ be a locally finite poset. The *incidence algebra* $I(\mathcal{P}; R)$ consists of functions $f : \{(x, y) \in P^2 \mid x \leq y\} \to R$ with the *convolution product*:

$$(f * g)(x, y) = \sum_{x \leq z \leq y} f(x, z) \cdot g(z, y)$$

Matrix intuition.

1. Index elements $x_1, x_2, \ldots$ so that $x_i \leq x_j$ implies $i \leq j$.
2. Each function becomes an *upper-triangular matrix*.
3. Convolution is just *matrix multiplication*.

Three Key Functions

The incidence algebra has three elements that make everything work.

Definition 42: The *Kronecker delta* is the identity for convolution:

$$\delta(x, y) = \begin{cases} 1 \text{ if } x = y \\ 0 \text{ otherwise} \end{cases}$$

Definition 43: The *zeta function* encodes the order relation:

$$\zeta(x, y) = \begin{cases} 1 \text{ if } x \leq y \\ 0 \text{ otherwise} \end{cases}$$

Definition 44: The *Möbius function* is the inverse of ζ in the incidence algebra:

$$\mu * \zeta = \zeta * \mu = \delta$$

Three Key Functions [2]

Convolving with ζ *sums over everything below*: for any f, $(f * \zeta)(y) = \sum_{x \leq y} f(x)$.

Convolving with μ *undoes* that sum.

Analogy:

- ζ = "integrate" (accumulate)
- μ = "differentiate" (recover original)
- δ = "do nothing"

Recursive Formula

The algebraic definition $\mu * \zeta = \delta$ yields an explicit recursion:

$$\mu(x, x) = 1 \text{ for all } x \in P$$

For $x < y$:

$$\mu(x, y) = -\sum_{x \leq z < y} \mu(x, z)$$

Example (Möbius function on the chain [3]): For the chain $(\{1, 2, 3\}, \leq)$:

- $\mu(1, 1) = 1$, $\mu(2, 2) = 1$, $\mu(3, 3) = 1$
- $\mu(1, 2) = -\mu(1, 1) = -1$
- $\mu(2, 3) = -\mu(2, 2) = -1$
- $\mu(1, 3) = -(\mu(1, 1) + \mu(1, 2)) = -(1 + (-1)) = 0$

Computing μ on the Boolean Lattice B_3

B	$\lvert B\rvert$	$\mu(\emptyset, B)$	Computation
$\emptyset$	0	1	base case
$\{1\}$	1	-1	$-\mu(\emptyset,\emptyset)$
$\{2\}$	1	-1	$-\mu(\emptyset,\emptyset)$
$\{3\}$	1	-1	$-\mu(\emptyset,\emptyset)$
$\{1,2\}$	2	1	$-(1+(-1)+(-1))$
$\{1,3\}$	2	1	$-(1+(-1)+(-1))$
$\{2,3\}$	2	1	$-(1+(-1)+(-1))$
$\{1,2,3\}$	3	-1	$-(1-1-1-1+1+1+1)$

Non-bottom intervals confirm the same pattern:

- $\mu(\{1\},\{1,2\}) = -1 = (-1)^1$ (distance 1)
- $\mu(\{1\},\{1,2,3\}) = -(1+(-1)+(-1)) = 1 = (-1)^2$ (distance 2)
- $\mu(\{1,2\},\{1,2,3\}) = -1 = (-1)^1$ (distance 1)

The pattern is clear: $\mu(A,B) = (-1)^{|B|-|A|}$.

Möbius Inversion Formula

The Möbius function satisfies the *cancellation identity*: for any $x \leq y$,

$$\sum_{x \leq z \leq y} \mu(x, z) = \begin{cases} 1 \text{ if } x = y \\ 0 \text{ if } x < y \end{cases}$$

Theorem 37: Let $\mathcal{P} = (P, \leq)$ be a finite poset and $f, g : P \to R$. Then:

$$f(x) = \sum_{y \leq x} g(y) \quad \Longleftrightarrow \quad g(x) = \sum_{y \leq x} \mu(y, x) \cdot f(y)$$

In words:

1. f counts "everything below" g.
2. μ tells you how to *recover* g from f.

Proof of the Inversion Formula

Proof: *(⇒)* Suppose $f(x) = \sum_{y \le x} g(y) = (g * \zeta)(x)$. Convolve both sides with μ:

$$\begin{aligned} f * \mu &= (g * \zeta) * \mu \\ &= g * (\zeta * \mu) \\ &= g * \delta = g \end{aligned}$$

So $g(x) = (f * \mu)(x) = \sum_{y \le x} f(y) \cdot \mu(y, x)$.

(⇐) Conversely, if $g = f * \mu$, then $g * \zeta = f * \mu * \zeta = f * \delta = f$. □

Möbius Function on the Boolean Lattice

Consider the Boolean lattice $\mathcal{P} = \left(2^{[n]}, \subseteq\right)$.

Theorem 38: In the Boolean lattice, the Möbius function is:

$$\mu(A, B) = (-1)^{|B|-|A|}$$

for $A \subseteq B$.

Proof: By strong induction on $d = |B| - |A|$.

Base ($d = 0$): $A = B$, so $\mu(A, A) = 1 = (-1)^0$. ✓

Inductive step: For $A \subset B$ with $|B| - |A| = d \geq 1$:

$$\mu(A, B) = -\sum_{A \subseteq C \subset B} \mu(A, C)$$

Möbius Function on the Boolean Lattice [2]

There are $\binom{d}{i}$ sets C with $|C| - |A| = i$ (choose which i elements to add).

By the inductive hypothesis, each contributes $(-1)^i$:

$$= -\sum_{i=0}^{d-1} \binom{d}{i} (-1)^i$$

Adding the missing $i = d$ term and subtracting it back:

$$= -\left(\sum_{i=0}^{d} \binom{d}{i} (-1)^i - (-1)^d \right)$$
$$= -\big(0 - (-1)^d\big) = (-1)^d$$

since $(1-1)^d = 0$ for $d \geq 1$. ✓ □

PIE as Möbius Inversion

Plugging $\mu(A, B) = (-1)^{|B|-|A|}$ into the inversion formula:

If $f(S) = \sum_{T \subseteq S} g(T)$ for all S, then:

$$g(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} f(T)$$

PIE IS Möbius inversion on the Boolean lattice!

The alternating signs $(-1)^{|S|}$ come directly from $\mu(A, B) = (-1)^{|B|-|A|}$.

This is the *deep reason* PIE works: it is a special case of a general algebraic identity.

The Chain: Finite Differences

Consider the simplest poset: $\mathbb{N}$ with the usual order $\leq$.

The Möbius function here is:

$$\mu(x, y) = \begin{cases} 1 & \text{if } y = x \\ -1 & \text{if } y = x + 1 \\ 0 & \text{otherwise} \end{cases}$$

Plugging into the inversion formula:

1. If $G(n) = \sum_{k=0}^{n} F(k)$ (accumulated sum),
2. then $F(n) = G(n) - G(n-1)$ (finite difference).

Möbius inversion on the chain is the *fundamental theorem of discrete calculus*.

Summation and differencing are inverse operations — just like integration and differentiation are in continuous calculus.

Zhegalkin Polynomials

The *Zhegalkin polynomial* (algebraic normal form) of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is a representation as an XOR-sum of conjunctions:

$$f(x_1, ..., x_n) = \bigoplus_{S \subseteq [n]} a_S \cdot \prod_{i \in S} x_i,$$

where $a_S \in \{0,1\}$ and $\oplus$ is XOR (addition mod 2).

The value of f on input set B is:

$$f(B) = \bigoplus_{A \subseteq B} a_A.$$

This is convolution with ζ on the Boolean lattice, but over the field $GF(2)$.

By Möbius inversion over $GF(2)$, subtraction equals addition, so $(-1)^{|B|-|A|} \equiv 1$:

$$a_B = \bigoplus_{A \subseteq B} f(A).$$

Zhegalkin Polynomials [2]

Example: Let $f(x, y) = x$ and y. Truth table: $f = (0, 0, 0, 1)$.

Compute coefficients:

- $a_1 = f(00) = 0$,
- $a_x = f(00) \oplus f(10) = 0$,
- $a_y = f(00) \oplus f(01) = 0$,
- $a_{xy} = f(00) \oplus f(10) \oplus f(01) \oplus f(11) = 0 \oplus 0 \oplus 0 \oplus 1 = 1$.

Result: $f(x, y) = xy$.

Takeaway. The standard algorithm for converting a truth table to Zhegalkin form (the "triangle method") is exactly Möbius inversion on the Boolean cube over $GF(2)$.

Counting Surjections

Example: How many *surjections* $h : [k] \to [n]$ are there?

Define:

- $g(S) = |\{h : [k] \to [n] \mid \text{image of } h \text{ equals } S\}|$ for $S \subseteq [n]$.
- $F(S) = |\{h : [k] \to [n] \mid \text{image of } h \subseteq S\}| = |S|^k$.

Then $F(S) = \sum_{T \subseteq S} g(T)$, *i.e.*, $F = g * \zeta$.

By Möbius inversion on the Boolean lattice:

$$\begin{aligned} g([n]) &= \sum_{T \subseteq [n]} (-1)^{n-|T|} F(T) \\ &= \sum_{T \subseteq [n]} (-1)^{n-|T|} |T|^k \\ &= \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} i^k \end{aligned}$$

This is exactly the PIE formula for surjections that we derived earlier!

Computing μ on the Divisor Lattice

Before stating the general formula, let us compute μ on the divisor lattice of 12.

Example: Divisors of $12 = 2^2 \cdot 3$ are: $\{1, 2, 3, 4, 6, 12\}$.

Using the recursion $\mu(1, d) = -\sum_{\substack{c \mid d, \\ c<d}} \mu(1, c)$:

d	$\mu(1, d)$	Computation
1	1	base case
2	-1	$-\mu(1,1)$
3	-1	$-\mu(1,1)$
4	0	$-(1 + (-1))$
6	1	$-(1 + (-1) + (-1))$
12	0	$-(1 - 1 - 1 + 0 + 1)$

Notice: $\mu(d) = 0$ when d has a squared prime factor ($4 = 2^2$, $12 = 2^2 \cdot 3$), and $\mu(d) = (-1)^k$ when d is a product of k distinct primes.

The Classical Möbius Function

Consider the divisor lattice $\mathcal{P} = (\mathbb{Z}^+, |)$.

Theorem 39: The Möbius function on $(\mathbb{Z}^+, |)$ is the *classical Möbius function*:

$$\mu(d) = \begin{cases} 1 & \text{if } d = 1 \\ (-1)^k & \text{if } d = p_1 \cdots p_k \text{ (distinct primes)} \\ 0 & \text{if } d \text{ has a squared prime factor} \end{cases}$$

where $\mu(a, b) = \mu(b/a)$ for $a \mid b$.

Proof: We need $\sum_{d \mid n} \mu(d) = \begin{cases} 1 \text{ if } n=1 \\ 0 \text{ if } n>1 \end{cases}$.

Case $n = 1$: $\sum_{d \mid 1} \mu(d) = \mu(1) = 1$. ✓

Case $n > 1$: Let $n = p_1^{a_1} \ldots p_r^{a_r}$. Divisors with a squared factor have $\mu(d) = 0$, so only *square-free* divisors contribute:

The Classical Möbius Function [2]

$$\begin{aligned}\sum_{d \mid n} \mu(d) &= \sum_{S \subseteq [r]} \mu\left(\prod_{i \in S} p_i\right) \\ &= \sum_{S \subseteq [r]} (-1)^{|S|} \\ &= \sum_{j=0}^{r} \binom{r}{j} (-1)^j \\ &= (1-1)^r = 0\end{aligned}$$

□

The identity $\mu(a,b) = \mu(b/a)$ holds because the interval $[a,b]$ in the divisor lattice is isomorphic to the divisor lattice of b/a.

Example:

- $\mu(1) = 1$, $\mu(2) = -1$, $\mu(3) = -1$
- $\mu(4) = 0$ (since $4 = 2^2$)
- $\mu(6) = \mu(2 \cdot 3) = 1$, $\mu(30) = \mu(2 \cdot 3 \cdot 5) = -1$

Number-Theoretic Möbius Inversion

Theorem 40: For arithmetic functions $f, g : \mathbb{Z}^+ \to R$:

$$f(n) = \sum_{d \mid n} g(d) \quad \Longleftrightarrow \quad g(n) = \sum_{d \mid n} \mu(d) \cdot f(n/d)$$

This is exactly Möbius inversion on $(\mathbb{Z}^+, |)$, where the order relation is divisibility.

The convolution here is the classical *Dirichlet convolution* of arithmetic functions.

Proof: Suppose $f(n) = \sum_{d \mid n} g(d)$. This is the Dirichlet convolution $f = g * \zeta$. Convolving both sides with μ: $f * \mu = g * \zeta * \mu = g * \delta = g$. Writing out explicitly: $g(n) = \sum_{d \mid n} \mu(d) f(n/d)$. □ □

Euler's Totient Function

Example: Define $\varphi(n) = |\{k \in [n] \mid \gcd(k, n) = 1\}|$.

Every $k \in [n]$ has $\gcd(k, n) = d$ for a unique $d \mid n$, so $n = \sum_{d \mid n} \varphi(d)$.

By Möbius inversion: $\varphi(n) = n \cdot \sum_{d \mid n} \frac{\mu(d)}{d}$.

Only square-free divisors survive, so the sum simplifies:

$$\varphi(n) = n \cdot \prod_{\substack{p \mid n, \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right)$$

Example: For $n = 12 = 2^2 \cdot 3$, we have:

$$\varphi(12) = 12 \cdot \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) = 12 \cdot \frac{1}{2} \cdot \frac{2}{3} = 4$$

The four numbers coprime to 12 are $1, 5, 7, 11$. ✓

Summary: Möbius Inversion

We saw four posets, each with its own Möbius function:

Poset	Möbius function	What it gives
Chain $(\mathbb{N}, \leq)$	$\mu(n, n) = 1$, $\mu(n-1, n) = -1$	Finite differences: $F(n) = G(n) - G(n-1)$
Boolean $\left(2^{[n]}, \subseteq\right)$	$\mu(A, B) = (-1)^{\lvert B\rvert - \lvert A\rvert}$	Inclusion–exclusion
Boolean over $GF(2)$	$\mu(A, B) = 1$ always	Zhegalkin polynomials
Divisors $(\mathbb{Z}^+, \lvert)$	$\mu(d)$: classical	Number-theoretic inversion

Common pattern: ζ sums over everything below, μ undoes the sum.

The poset determines the form of μ.

The Twelvefold Way

"The purpose of computing is insight, not numbers."

— Richard Hamming

Balls and Boxes

Consider the most basic combinatorial question:

How many ways can we place n balls into k boxes?

There are three independent dimensions of variation:

Balls

1. Distinct (labeled $1, \ldots, n$)
2. Indistinguishable (identical)

Boxes

1. Distinct (labeled $1, \ldots, k$)
2. Indistinguishable (unlabeled)

Placement

1. Unrestricted
2. Injective (≤ 1 per box)
3. Surjective (≥ 1 per box)

Total: $2 \times 2 \times 3 = 12$ cases.

Small Examples

Example: Place $n = 2$ distinct balls into $k = 2$ distinct boxes, surjectively.

Placement	Description
Ball 1 $\to$ Box 1, Ball 2 $\to$ Box 2	each ball in its own box
Ball 1 $\to$ Box 2, Ball 2 $\to$ Box 1	balls swapped

Answer: $k! s_k^{\mathrm{II}}(n) = 2! \cdot s_2^{\mathrm{II}}(2) = 2 \cdot 1 = 2$. ✓

Example: Place $n = 3$ indistinguishable balls into $k = 2$ distinct boxes, surjectively.

Each box gets ≥ 1, so we need compositions of 3 into 2 positive parts: $3 = 1 + 2 = 2 + 1$.

Answer: $\binom{n-1}{k-1} = \binom{2}{1} = 2$. ✓

Distinct Balls, Distinct Boxes

This is the theory of *functions* $f : [n] \to [k]$.

Unrestricted

Any function: k^n.

Each of n balls has k choices. Multiplication principle.

Injective

At most one ball per box: $(k)_n = k!/(k-n)!$.

The number of injective functions $[n] \to [k]$. Requires $k \geq n$.

Surjective

Every box gets ≥ 1 ball: $k! s_k^{\text{II}}(n)$.

Partition $[n]$ into k labeled blocks (Stirling number $s_k^{\text{II}}(n)$), then assign the k labels to blocks in $k!$ ways.

The surjection count uses PIE (equivalently, Möbius inversion on the Boolean lattice):

$k! s_k^{\text{II}}(n) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$.

Indistinguishable Balls, Distinct Boxes

This is the theory of *multisets* and *compositions*.

Unrestricted

Multiset of size n from $[k]$: $\binom{n+k-1}{n}$.

Stars and bars: arrange n stars and $k-1$ bars.

Injective

At most one per box: $\binom{k}{n}$.

Choose which n of k boxes are occupied. Requires $k \geq n$.

Surjective

Every box gets ≥ 1: $\binom{n-1}{k-1}$.

Compositions of n into k positive parts (give each box 1 first, then distribute $n-k$ freely). Requires $n \geq k$.

Distinct Balls, Indistinguishable Boxes

This is the theory of *set partitions*.

Unrestricted

Into $\leq k$ non-empty blocks: $\sum_{j=1}^{k} s_j^{\text{II}}(n)$.

Sum of Stirling numbers. If $k \geq n$, this equals B_n.

Injective

At most one ball per box: $[n \leq k]$.

If $n \leq k$: each ball in its own box, one way. If $n > k$: impossible (0).

Surjective

Into exactly k non-empty blocks: $s_k^{\text{II}}(n)$.

Stirling number of the second kind. Requires $n \geq k$.

Here $[P]$ denotes the *Iverson bracket*:

$[P] = 1$ if P is true, 0 otherwise.

Indistinguishable Balls, Indistinguishable Boxes

This is the theory of *integer partitions*.

Unrestricted

Into $\leq k$ parts: $\sum_{j=1}^{k} p_{j(n)}$.

Sum of partition counts into exactly j parts.

Injective

At most one ball per box: $[n \leq k]$.

Same as Row 3: if $n \leq k$, place one per box. If $n > k$: impossible.

Surjective

Into exactly k parts: $p_{k(n)}$.

Integer partitions of n into exactly k parts. Requires $n \geq k$.

Here $p_{j(n)}$ denotes the number of integer partitions of n into exactly j parts.

Note that Rows 3 and 4 are *unlabeled*, so the formulas differ from Rows 1 and 2 only in "forgetting" the box labels.

From Unrestricted to Surjective

The surjective column is obtained from the unrestricted column by Möbius inversion on the *Boolean lattice*.

Distinct balls, distinct boxes:

Unrestricted: k^n (all functions).

Surjective: $k! s_k^{\mathrm{II}}(n) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$.

This is PIE — Möbius inversion on $2^{[k]}$!

Indistinguishable balls, distinct boxes:

Unrestricted: $\binom{n+k-1}{n}$.

Surjective: $\binom{n-1}{k-1}$.

The same PIE argument, applied to "is box i empty?" properties.

Forgetting Labels

The transition from distinct to indistinguishable boxes is "forgetting the labels." For *surjective* arrangements, every placement becomes one of $k!$ equivalent arrangements (one for each permutation of the labels), so dividing by $k!$ gives the count for indistinguishable boxes.

Key identity: dividing Row 1 surjective by $k!$ gives Row 3 surjective:

$$s_k^{\Pi}(n) = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$$

More generally, this "forget labels" operation is governed by the Möbius function of the *partition lattice* Π_k.

The Twelvefold Way

Balls	Boxes	Unrestricted	Injective	Surjective
Distinct	**Distinct**	k^n	$(k)_n$	$k! s_k^{\text{II}}(n)$
Indist.	**Distinct**	$\binom{n+k-1}{n}$	$\binom{k}{n}$	$\binom{n-1}{k-1}$
Distinct	**Indist.**	$\sum_{j=1}^{k} s_j^{\text{II}}(n)$	$[n \leq k]$	$s_k^{\text{II}}(n)$
Indist.	**Indist.**	$\sum_{j=1}^{k} p_{j(n)}$	$[n \leq k]$	$p_{k(n)}$

All formulas hold for the standard non-degenerate cases $n, k \geq 1$.

Course Wrap-Up

Counting fundamentals
- Addition, multiplication, bijection principles
- Permutations and combinations
- Multisets and compositions
- Binomial and multinomial theorems

Analytic methods
- Generating functions
- Characteristic equations
- Annihilators
- Master theorem, Akra–Bazzi method

Structures
- Set partitions (Stirling, Bell)
- Integer partitions (Ferrers, Young tableaux)
- Inclusion–exclusion (PIE, derangements)

Advanced topics
- Gamma function
- Möbius inversion
- The Twelvefold Way

Bibliography

[1] G. Pólya, *Mathematics and Plausible Reasoning, Volume 1: Induction and Analogy in Mathematics*. 1954.

[2] J. L. Bentley, D. Haken, and J. B. Saxe, "A General Method for Solving Divide-and-Conquer Recurrences," *ACM SIGACT News*, vol. 12, no. 3, pp. 36–44, Sept. 1980, doi: 10.1145/1008861.1008865.

[3] M. Akra and L. Bazzi, "On the Solution of Linear Recurrence Equations," *Computational Optimization and Applications*, vol. 10, no. 2, pp. 195–210, May 1998, doi: 10.1023/A:1018373005182.