Math 184 Notes

Enumerative Combinatorics

Spring 2021
Taught by Professor Daniel Kane

Table of Contents

1 Mathematical Induction

There are two types of mathematical induction: weak induction (also simply known as *induction*) and strong induction.

1.1 Weak Induction

Suppose you have some sequence of statements that you want to prove:

$$S_1, S_2, \ldots, S_n, \ldots$$

The principle of mathematical induction states that it is enough to show:

1. S_1 is true.

2. For all n where S_n is true, so is S_{n+1}.

Together, you can get S_1, which implies S_2, which implies S_3, and so on.

Remark: It should be noted that, instead of proving $S_1, S_2, \ldots, S_n$ for S_{n+1}, you can prove $S_1, S_2, \ldots, S_{n-1}$ for S_n. This just comes down to whichever one is easier to interpret.

1.1.1 Format for Weak Induction

1. Give statement you want to prove by induction.

2. State that you will prove it by induction on (inductive variable).

3. State and prove the base case.

4. Inductive step.

   - State that you are starting inductive step.
   - State inductive hypothesis.
   - Use the inductive hypothesis to prove the next step.

5. Conclude your original claim.

Remark:

- Starting the inductive variable is *necessary*.

- Base case should be the *smallest* n for which you want to prove your statement.

- Induction is usually helpful in situations where statements for n and $n + 1$ is easy to relate.

1.1.2 Example: Weak Induction

Prove $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$.

Proof. We will use mathematical induction on n.

   $\diamondsuit$ <u>Base Case:</u> Let $n = 1$. Then, $\frac{1(1+1)}{2} = \frac{2}{2} = 1$, so this is true.

◇ <u>Inductive Step:</u> Suppose that $a + 2 + \cdots + n = \frac{n(n+1)}{2}$. Then:

$$
\begin{aligned}
1 + 2 + \cdots + n + (n+1) &= \frac{n(n+1)}{2} + (n+1) \\
&= \frac{n(n+1) + 2(n+1)}{2} \\
&= \frac{(n+2)(n+1)}{2} \\
&= \frac{(n+1)(n+2)}{2}
\end{aligned}
$$

And so this completes the inductive steps and proves our results.   □

1.2   Strong Induction

In order to prove S_n for all n, we need:

1. S_1 is true.

2. If S_m is true for all $m < n$, then S_n is true.

This gives us S_1, which implies S_2. Together, they imply S_3, and then S_4, and so on.

1.2.1   Format for Strong Induction

- Give statement you want to prove by *strong* induction.

- State that you will prove it by *strong* induction on an inductive variable.

- State and prove the base case (or skip).

- Inductive steps.

    - State you are starting inductive step.
    - State the inductive hypothesis for all smaller m.
    - Use to prove next step.

- Conclude your original claim.

1.2.2   Example

Prove that $a_n = 2^{n-1}$ by strong induction on $n \geq 1$.

Proof. We will use strong induction on n.

◇ <u>Base Case:</u> Let $n = 1$. Then, $a_1 = 1 = 2^{1-1} = 2^0 = 1$. This is true.

◇ <u>Inductive Step:</u> Suppose that $a_m = 2^{m-1}$ is true for all $1 \leq m < n$. Then:

$$
\begin{aligned}
a_n &= a_{n-1} + a_{n-2} + \cdots + a_2 + a_1 + a_0 \\
&= 2^{n-2} + 2^{n-3} + \cdots + 2 + 1 + 1 \\
&= (2^{n-1} - 1) + 1 \\
&= 2^{n-1}
\end{aligned}
$$

This completes our proof.   □

2 Pigeonhole Principle

There is some pair of people in San Diego with exactly the same number of hairs on their heads.

Suppose we were asked to prove this problem. How would we go about this? A problem like this makes use of the pigeonhole principle.

2.1 The Basic Pigeonhole Principle

> **Theorem 2.1: Pigeonhole Principle**
>
> Given n pigeons each assigned to one of m holes for some $m < n$, there must be some hole with at least two pigeons.

Remark: Consider pigeons as type-A objects and holes as type-A containers. In particular, we have n-many type-A objects and m-many type-A containers where $m < n$. We define some mapping where each type-A object must be in exactly one type-A container. Since there are more type-A objects than containers, there must exist a container with at least two type-A objects.

2.1.1 Example: Hair

Recall the question from above. The proof is as follows.

Proof. There are about one million people in San Diego. Each person has at least 100,000 hairs. Then, using the pigeonhole principles, let the pigeons be the people and the holes be the number of hairs in $\{0, 1, 2, ..., 100,000\}$. So, it follows that there must be at least two people with the same amount of hair on their head. □

2.2 The Generalized Pigeonhole Principle

> **Theorem 2.2: Generalized Pigeonhole Principle**
>
> Given n pigeons each assigned to one of m holes for some m with $(k-1)m < n$, then there must be some hole with at least k pigeons.

Remark: Consider pigeons as type-A objects and holes as type-A containers. In particular, we have n-many type-A objects and m-many type-A containers where $m < \frac{n}{k-1}$. We define some mapping where each type-A object must be in exactly one type-A container. Since there are $(k-1)$-many more type-A objects than containers, there must exist a container where the ratio of type-A objects to containers is $k-1 < \frac{n}{m} \implies k \leq \frac{n}{m}$.

2.2.1 Example: Packing Tetrahedrons

Suppose you placed 1000 points in the unit cube. Show that some 4 of them form a tetrahedron of volume at most $\frac{1}{999}$.

Proof. We could try a similar packing argument, but we really want to find 4 points that are *all* close together. For this, we need to make use of the generalized pigeonhole principle.

Here, we can split the cube into 333 equal slices. Notice that $1000 > 3 \cdot 333$, so by the generalized pigeonhole principle, there is some slice with 4 points. Thus, that tetrahedron has volume at most:

$$\frac{\frac{1}{333}}{3} = \frac{1}{999}$$

The volume of a tetrahedron that you can fit in the box is at most a third of the volume of the box. □

3   Counting Principles

We will now briefly talk about the different counting principles along with their applications. Counting is a major part of combinatorics with a number of applications. For example:

- For the pigeonhole principle.

- To compute probabilities.

- To estimate the length of enumerations.

3.1   Addition Rule

If you have two **disjoint** collections of items, the total number of items is the sum of the number from each collection.

If we have two sets S and T such that $S \cap T = \emptyset$ (i.e. S and T are disjoint), this means:

$$|S \cup T| = |S| + |T|$$

If the sets A_1, A_2, ..., A_n are disjoint, then the addition rule is simply the cardinality of the union of those sets.

$$\left(\bigcap_{i=1}^{n} A_i = \emptyset \right) \implies \left(\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{i=1}^{n} |A_i| \right)$$

3.1.1   Example: Total Items

Consider the question: *You have 4 shirts and 3 pairs of pants. How many items of clothing do you have in total?*

- We are asked to find the total number of clothing that we have. The answer is obviously $4 + 3 = 7$.

3.2   Multiplication Rule

Given two sets S and T, the number of ways to select one element from *each* of the sets is the product of their sizes.

That is, $|S \cdot T| = |S| \cdot |T|$.

3.2.1   Example: Total Combinations

Consider the question: *You have 4 shirts and 3 pair of pants. How many outfits (consisting of a shirt and pair of pants) do you have in total?*

- The answer is $4 \cdot 3 = 12$. For each shirt, you can have up to 3 pairs of pants.

3.3   Exponent Rule

The number of k letter words using an n letter alphabet is simply n^k.

3.3.1   Example: Total Words

Consider the question: *How many possible three letter words are there?*

- The answer is $26^3 = 17576$. One way to see this is:

1st Letter	2nd Letter	3rd Letter
26 Options	26 Options	26 Options

  And so $26 \cdot 26 \cdot 26 = 26^3$.

3.3.2   Example: Subsets

Show that number of subsets of a set $S = \{x_1, x_2, ..., x_n\}$ of size n is 2^n.

Proof. For each i, $x_i \in S$ is either in the subset or not. Because of this, each of these elements have two options. Every combination of these gives us a different subset. These choices can be made independently of each other, so it follows that $2 \cdot 2 \cdot ... \cdot 2 = 2^n$. □

3.4   Generalized Multiplication Rule

If you want to pick pairs of objects and:

- Have n possibilities for the first object.

- For each first object, have m possibilities for the second.

    - Note that the options might depend on which first object you picked so long as the number doesn't.

- The total number of possible combinations is nm.

3.4.1   Example: 2-Letter Words

Consider the question: *How many 2-letter words using the letters A, B, C, D are there that do not repeat letters?*

- The answer is 12. For each 1st letter, there are 3 possibilities for the 2nd letter. For example, if the first letter is A, then your second letter can either be B, C, or D.

3.5   Falling Factorials (They Count Permutations)

Closely associated with the generalized multiplication rule, a falling factorial is written like so:

$$(n)_k = n(n-1)(n-2)...(n-k+1)$$

A special case is if $n = k$; in this case, we get:

$$n! = n(n-1)(n-2)...(2)(1)$$

Something to keep in mind is:

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

3.5.1   Example: Strings Without Repeats

Consider the question: *Suppose you have an n letter alphabet and want to write a word with k letters no two of which are the same. How many ways can you do this?*

- The answer is $n(n-1)(n-2)\ldots(n-k+1)$. To see this, we note the following:

    - n possibilities for the first letter.
    - $n-1$ possibilities for the second letter.
    - $n-2$ possibilities for the third letter.
    - ...
    - $n-k+1$ possibilities for the last letter.

3.5.2   Example: Medal Combinations

Consider the question: *10 people compete in a race. How many possible combinations are there for 1st, 2nd, and 3rd place finishers?*

- The answer is $(10)_3$ because you are picking 3 from 10 without replacement.

3.5.3   Example: Line at the Dentist

Consider the question: *5 people are waiting at the dentist. How many orders might they be dealt with?*

- The answer is 120. We need to order all 5 (without replacement), so $(5)_5 = 5! = 120$.

Example: Bijections

Consider the example: *How many bijections are there for $f : X \to X$?*

- The answer is $n!$. If $|X| = \{x_1, x_2, ..., x_n\}$, then:

    - Have n possibilities for $f(x_1)$.
    - Have $n - 1$ possibilities for $f(x_2)$ (must be different).
    - ...
    - Have 1 possiblity for $f(x_n)$.

    So, the total is $n! = |X|!$.

3.6   Multinomial Coefficients

Suppose we want to put in order:

- a_1 things of type 1.

- a_2 things of type 2.

- a_3 things of type 3.

- $\dots$

- a_m things of type m.

And $a_1 + a_2 + \cdots + a_m = n$. Then, the total number of arrangements is simply:

$$\binom{n}{a_1, a_2, \ldots, a_m} = \frac{n!}{a_1! a_2! \ldots a_m!}$$

Which is known as the multinomial coefficient.

3.6.1   Example: Flower Arranging

Consider the question: *You have 2 red flowers, 2 blue flowers, and 1 yellow flowers. You want to arrange them in a row. How many color platters can you achieve? For example, you could have the patterns RRBBY, RYBRB, BRYRB.*

- The answer is:
$$\frac{5!}{2!2!1!} = 30$$

    The idea is as follows: if we had 5 different colors, then the answer would just be $(5)_5 = 5! = 120$. However, we only have 3 different colors, so this calculation is invalid. How many choices do we have for the second color? Well, it depends.

- If the first flower is yellow, then the second can only be red or blue.
- If the first is red or blue, then the second flower can be anything.

Thus, the generalized multiplication rule doesn't directly apply. So, let's consider the exact flowers. Suppose we have:

- 1 rose (red) - call it R_1.
- 1 tulip (red) - call it R_2.
- 1 iris (blue) - call it B_1.
- 1 morning glory (blue) - call it B_2.
- 1 marigold (yellow) - call it Y_1.

There are 5! ways to arrange these flowers based off the names. However, this is an invalid answer because this assumes that $R_1R_2B_1B_2Y_1$ and $R_2R_1B_1B_2Y_1$ are different sequences even though they lead to the same color pattern: $RRBBY$. This leads to the question: Given a color pattern, how many flower patterns match it? In other words, for a pattern $RRBBY$, how many different combinations are there?

- 2 choices for which red flower is which.
- 2 choices for which blue flowers is which.
- Only 1 yellow flower.

So, there are $2 \cdot 2 = 4$ matching patterns. To confirm this, we consider all possible arrangements that match $RRBBY$. In this case, we can have:

- $R_1R_2B_1B_2Y_1$
- $R_2R_1B_1B_2Y_1$
- $R_1R_2B_2B_1Y_1$
- $R_2R_1B_2B_1Y_1$

This leads to the idea of counting things in two different ways. In other words, consider the question, *how many flower orders are there?* We consider a few things:

- On one hand, we know that the number of ways of ordering 5 things is $5! = 120$.
- On the other hand, there's a second method that we can use that should give the same answer. First, consider one possible color pattern. Then, you can find a flower ordering that corresponds to that pattern. The idea is that for every color pattern, there are exactly 4 flower orderings that give you that flower pattern. So, by the *generalized multiplication rule*, we know that:

$$\text{Num. Color Orders} \cdot \text{Num. Flower Orders per Color Order} = 4 \cdot \text{Num. Color Orders}$$

But, remember that these two things must give us the same number.

And so it follows that $120 = 4 \times$ Num. Color Orders. Thus, the number of color orders is:

$$\frac{120}{4} = \boxed{30}$$

3.6.2   Example: Anagrams

Consider the question: *How many anagrams (letter rearrangements) of "Mississippi" are there?*

- To begin, let's distinguish the letters:

$$m_1 p_1 p_2 i_1 i_2 i_3 i_4 s_1 s_2 s_3 s_4$$

Here, there are 11! ways to order all these letters with the labels. However, we are overcounting because this assumes that each letter is different when, in reality, some of them are actually the same. In other words, this assumes that $m_1 i_1 s_1 s_2 i_2 s_3 s_4 i_3 p_1 p_2 i_4$ and $m_1 i_1 s_1 s_2 i_2 s_3 s_4 i_3 p_2 p_1 i_4$ are two different words when, in reality, they produce the same word.

Now consider that, for each anagram, there are:

- 1! way to label the m's.
- 2! ways to label the p's
- 4! ways to label the i's
- 4! ways to label the s's

And so, it follows that there are $\frac{11!}{1!2!4!4!} = \boxed{34650}$ number of anagrams.

3.6.3   Example: Lottery

Consider the question: *The CA Powerball lottery selects 5 numbers from 1 to 59. How many possible combinations are there?*

- First, we consider that:

  - Each number 1 to 59 is either chosen or not chosen, with 5 chosen and 54 not chosen.
  - Writing them out in order yields a sequence of 5 chosens and 54 not-chosens.

So, the number of combinations is simply:

$$\frac{59!}{5!54!} = \boxed{5006386}$$

3.7   Binomial Coefficients (They Count Combinations)

This is simply a special case of the *multinomial coefficients*.

Definition 3.1: Binomial Coefficient

The **binomial coefficient** n choose k is given by:

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

This is simply the number of ways to pick k items from a set of n items.

Some facts to remember:

$$\binom{n}{0} = \binom{n}{n} = 1$$

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

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

3.7.1   Example: Word Count

Consider the question: *How many possible 5-letter words are there with exactly two vowels?*

- We begin by considering the following:
  - First, we consider $\binom{5}{2} = \frac{5!}{2!(5-2)!} = 10$ (5 vowels but we choose 2) ways to pick which locations have the two vowels.
  - Then, there are $21 \cdot 21 \cdot 21 = 9261$ ways to pick the non-vowels (note that it is *not* $(21)_3$ because we don't care about repetition).
  - And, there are $5^2 = 25$ ways to pick the vowels (note that it is *not* $(5)_2 = 5 \cdot 4 = 20$ because we don't care whether the vowels repeat or not).

So, we have $10 \cdot 9261 \cdot 25 = \boxed{2315250}$ combinations.

4 Balls and Bins

When talking about balls and bins, there are four situations that we need to consider.

- Balls can be distinguishable or indistinguishable.

- Bins can be distinguishable or indistinguishable.

The best way to highlight this is through a simple example. Suppose we have three balls and three bins.

- If the balls and bins are both distinguishable (say we have bin 1, 2, 3 and a red, blue, and green ball), then there are 6 ways to put these balls into the bins:

Num.	Bin 1	Bin 2	Bin 3
1	R	B	G
2	R	G	B
3	B	R	G
4	B	G	R
5	G	B	R
6	G	R	B

- If only the balls are indistinguishable (we only have white balls), then there's only one way to put the balls into the bins:

Num.	Bin 1	Bin 2	Bin 3
1	W	W	W

- If only the bins are indistinguishable (we have unnumbered bins), then there's also only one way to put the balls into the bin:

Num.	Bin	Bin	Bin
1	R	G	B

4.1 Distinguishable Balls into Distinguishable Bins

Each ball b_1, b_2, ..., b_n can go into any of the k bins (possibly empty). There are k options for each, so the total number of possibilities is k^n.

4.2 Indistinguishable Balls into Distinguishable Bins: Compositions

With indistinguishable balls, you can no longer tell where the "first" or "second" ball went. What can you tell, though?
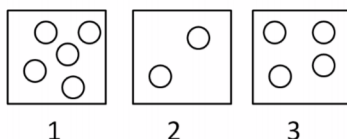
In this example, you are able to count the number of balls in each bin, though. For example, there are 5 balls in bin 1, 2 balls in bin 2, and 4 balls in bin 3.

4.2.1   Definitions

> **Definition 4.1: Weak Composition**
>
> A weak composition of n is a sequence of non-negative integers $a_1, a_2, ..., a_k$ so that $a_1 + a_2 + \cdots + a_n = n$.

> **Definition 4.2: Composition**
>
> A composition of n is a sequence of positive integers $a_1, a_2, \ldots, a_k$ so that $a_1 + a_2 + \cdots + a_n = n$. These are sometimes known as *strict compositions*.

So, if we put n balls into k bins, letting a_i be the number of balls in the ith bin, we get a (weak) composition of n into k parts. In this case:

- If we allow the bins to be empty, we want weak compositions.

- If we do not allow the bins to be empty (i.e. every bin must have at least one ball), we want strict compositions.

4.2.2   Stars and Bars

In order to count compositions, we need a clever way of writing them.

- For each i, write a_i many stars (*).

- Separate these terms with a bar (|). In other words, where does one number start and where does the next one begin.

For example, $5 + 2 + 4$ becomes:

*****|**|****

So, what happens if we write a weak composition of n into k parts in terms of stars and bars? We can get an arrangement with:

- n total stars.

- $k - 1$ total bars.

And any such arrangement gives rise to binomial coefficients.

> **Theorem 4.1: Number of Weak Compositions**
>
> For all positive integers n and k, the number of weak compositions of n into k parts is:
> $$\binom{n + k - 1}{k - 1} = \binom{n + k - 1}{n}$$

What happens if we cannot allow bins to be empty?

- Then, we want compositions rather than weak compositions.

- We can relate the two. We have that $a_1, a_2, \ldots, a_k$ is a composition of n if and only if $a_1 - 1, a_2 - 1, \ldots, a_k - 1$ is also a weak composition of $n - k$.

- So, the number of composition of n into k parts equals the number of weak compositions of $n - k$ into k parts, or $\binom{n-1}{k-1} = \binom{(n-k)+k-1}{k-1}$.

We can also view this using stars and bars. Specifically, we need arrangements of n stars and $k - 1$ bars where no two bars are adjacent and no bars are at the end. For example, these are invalid:

```
|**|
***||**
```

Consider the arrangement:

```
*[ ]*[ ]*[ ]*[ ]*[ ]*[ ]*[ ]*[ ]
```

Where [ ] represents possible locations for a bar. Then, we can select $k-1$ locations from $n-1$ options, so $\binom{n-1}{k-1}$ possiblities.

Corollary 4.1: Number of Compositions

For all positive integers n and k, the number of compositions (not weak compositions) of n into k parts is:
$$\binom{n-1}{k-1}$$

What if we don't restrict the number of bins? Well, we need to consider a few things:

- Must only consider non-empty bins. Otherwise, we can just keep adding empty bins.

- Get the total number of compositions of n.

- From the above argument, we have $n-1$ positions for bars, and we can select any number of them.

- Well, this is simply the number of subsets of a set of size $n-1$, or 2^{n-1}.

Corollary 4.2: Number of All Compositions

For all positive integers n, the number of all compositions (not weak compositions) of n is:
$$\sum_{k=1}^{n}\binom{n-1}{k-1} = 2^{n-1}$$

4.3 Distinguishable Balls into Indistinguishable Bins: Set Partitions

Consider balls labeled $1, 2, \ldots, n$. We can tell which set of balls are in each bin. Here, we make use of set partitions to count these.

Definition 4.3: Set Partition

A **set partition** of a set S is a collection of non-empty subsets $S_1, S_2, \ldots, S_k$ so that each element of S is in exactly one of the S_i.

Definition 4.4: Stirling Numbers

The **stirling numbers of the second kind** are given by:
$$S(n,k) = \text{ Number of partitions of } [n] \text{ into } k \text{ parts.}$$

Note that $[n] = \{1, 2, 3, ..., n\}$.

Theorem 4.2: Stirling Numbers and Recurrence Relations

For all positive integers $k \leq n$ (and $n, k \geq 1$):
$$S(n,k) = S(n-1, k-1) + k \cdot S(n-1, k)$$

Proof. A partition of $[n]$ into k parts has either:

- n is in a set by itself (i.e. n is lone). In this case, removing $\{n\}$ will yield a partition of $[n-1]$ into $k-1$ parts. Thus, the number of such set partitions if $S(n-1, k-1)$.

- n is in a set with other elements. In this case, removing n from whatever part it is in will yield a partition of $[n-1]$ into k parts. Of course, there are $S(n-1, k)$ ways to choose this partition, and then k ways to pick which part to insert n into. Then, the total is simply $k \cdot S(n-1, k)$. □

4.3.1 Example: Stirling Number

What is $S(3, 2)$?

- The answer is **3**. There are 3 possible partitions that partition the set into two subsets:

  - $\{1, 2\}, \{3\}$
  - $\{1, 3\}, \{2\}$
  - $\{2, 3\}, \{1\}$

4.3.2 Facts about Stirling Numbers

- $S(n, n) = 1$ (only partition is $\{1\}, \{2\}, \ldots, \{n\}$).

- $S(n, 1) = 1$ (only partition is $\{1, 2, ..., n\}$).

- $S(0, 0) = 1$ (by definition).

- $S(n, n-1) = \binom{n}{2}$. Any partition of $[n]$ into $n-1$ parts must have $n-2$ parts of size 1 and 1 part of size 2.

- $S(n, 2) = 2^{n-1} - 1$. Splitting $[n]$ into two subsets, one subset contains 1 plus some subset of $\{2, 3, ..., n\}$. The other contains the rest. We can use any subset of $\{2, 3, ..., n\}$ except for the entire thing (as then the other set will be empty).

Corollary 4.3

For all real numbers x and all non-negative integers n:

$$x^n = \sum_{k=0}^{n} S(n, k)(x)_k$$

Proof. We proceed by induction on n.

◇ <u>Base Case:</u> Let $n = 0$. Then, $x^n = x^0 = 1$. Also, $S(0, 0)(x)_0 = 1 \cdot 1 = 1$.

◇ <u>Inductive Step:</u> Assume that $x^{n-1} = \sum_k S(n-1, k)(x)_k$. Then:

$$\sum_{k=1}^{n} S(n, k)(x)_k = \sum_{k=1}^{n} (S(n-1, k-1) + kS(m-1, k))(x)_k$$

$$= \sum_{k=1}^{n} S(n-1, k-1)(x)_k + \sum_{k=1}^{n} kS(n-1, k)(x)_k$$

$$= \sum_{k=1}^{n-1} S(n-1, k)(x)_{k+1} + \sum_{k=1}^{n-1} kS(n-1, k)(x)_k$$

$$= \sum_{k=1}^{n-1} S(n-1, k)((x)_{k+1} + k(x)_k)$$

$$= \sum_{k=1}^{n-1} S(n-1, k)(x)_k((x-k) + k)$$

$$= x \sum_{k=1}^{n-1} S(n-1, k)(x)_k$$

$$= x^n$$

This concludes the proof. □

4.3.3   Bell Numbers

What about the *total* number of set partitions of $[n]$? While there isn't a good formula, we can define one.

> **Definition 4.5: Bell Numbers**
>
> The bell number, $B(n)$, is the number of set partitions of a set of size n.

4.3.4   Example: Bell Numbers

What is $B(3)$?

- The answer is 5. The set partitions are:
  - $\{1\}\{2\}\{3\}$
  - $\{1, 2\}\{3\}$
  - $\{1, 3\}\{2\}$
  - $\{2, 3\}\{1\}$
  - $\{1, 2, 3\}$

4.3.5   Bell Numbers in Terms of Stirling Numbers

Because bell numbers are simply an enumeration of all possible set partitions:

$$B(n) = \sum_{k=0}^{n} S(n, k)$$

But, what about a recursive definition? First, let's ask the question: what is the set containing n?

- If size k, we know that there are $\binom{n-1}{k-1}$ possible sets.

- There are $B(n-k)$ ways to partition remaining elements.

So, the recurrence relation is:

$$B(n) = \sum_{k=1}^{n} \binom{n-1}{k-1} B(n-k)$$

Theorem 4.3

For all non-negative integers n:

$$B(n+1) = \sum_{i=0}^{n} \binom{n}{i} B(i)$$

4.4   Indistinguishable Balls into Indistinguishable Bins: Integer Partitions

Consider unlabeled balls into unlabeled boxes, like so:
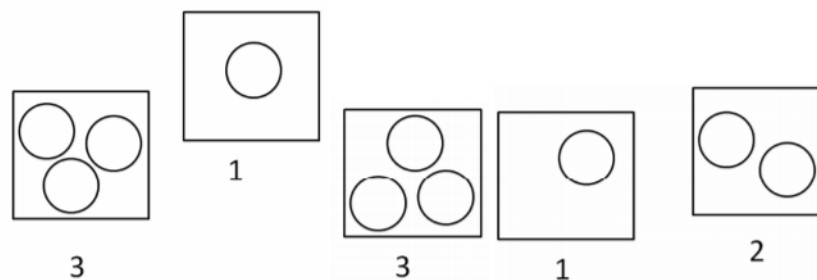
The question, then, is: *what can we keep track of?*

- Number of balls in each box.

Definition 4.6: Integer Partition

An **integer partition** of a positive number n is a sequence of positive integers $a_1 \geq a_2 \geq \cdots \geq a_k$ so that:

$$a_1 + a_2 + \cdots + a_k = n$$

The a_i are called the **parts** of the partition.

Remark:

- Unlike compositions, this is an unordered collection of numbers. So, $3 + 3 + 2 + 1 + 1$ is the same as $3+2+3+1+1$. Because of this, we need to sort the numbers in decreasing order (hence the definition).

- The integer partitions of n correspond to the ways of putting n unlabeled balls into nonempty boxes.

4.4.1   Example: Integer Partitions of 4

How many integer partitions of 4 are there?

- There are 5 partitions.

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

4.4.2   Partition Function

There is no clean formula. However, we can come up with a definition.

> **Definition 4.7: Partition Function**
>
> $p(n)$ is the number of integer partitions of n. $p_k(n)$ is the number of partitions of n into exactly k parts.

> **Theorem 4.4**
>
> $p_k(n)$ is the number of partitions of n with the largest part of size k.

4.4.3   Ferrers Diagram

A Ferrers Diagram is a way to visualize partitions. In particular, we note that the ith row has a_i squares. For instance, consider the Ferrers Diagram:
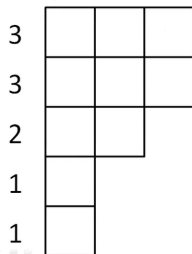
This corresponds to the partition $3 + 3 + 2 + 1 + 1 = 10$.

There are some properties of Ferrers Diagrams to know.

- Total number of squares equals total size of partitions (i.e. it adds up to n).

- No gaps. For any square in the diagram, the square above and the square to the left are also in the diagram (unless this would go over the edge).

- One-to-one correspondence between partitions and Ferrers diagrams.

4.4.4   Conjugates

Given a Ferrers diagram, we can come up with what is known as a **conjugate** Ferrers diagram. This is done by *flipping the rows and columns.*

Consider the example that we used above. Then:
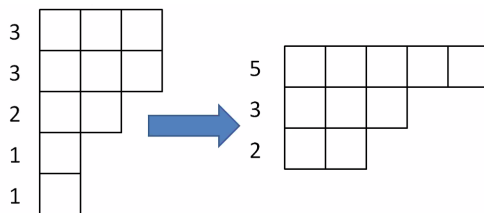
So, the conjugate of $3 + 3 + 2 + 1 + 1$ is $5 + 3 + 2$.

There are some properties of conjugates to keep in mind.

- The size of a partition and its conjugate are the same. In other words, each have the same number of squares.

- The conjugate of the conjugate of a partition is the original partition.

- The number of parts in a partition equals the size of the largeest part of its conjugate.

4.4.5   Self-Conjugate Partitions

Definition 4.8: Self-Conjugate

A partition is called *self-conjugate* if it is the same as its conjugate.

Here is an example of a self-conjugate partition:
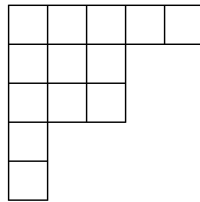
The squares in the Ferrers diagram for any self-conjugate partition can be decomposed into Ls. In the example shown below, we have $9 + 3 + 1$.
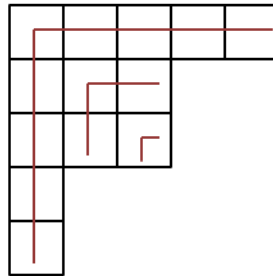
- Each L has an odd number of squares (self-conjugate).

- Each L has fewer squares than the outer ones. In other words, they must nest.

As a result, we have the following theorem:

Theorem 4.5

The number of self-conjugate partitions of n is the same as the number of partitions of n in to distinct, odd parts.

4.5   Summary

n into k could be empty / n into any not empty / n into k not empty	Labeled Balls	Unlabeled Balls
Labeled Bins	n^k / $\sum_{m=1}^{n} m!S(n,m)$ / $k!S(n,k)$	$\binom{n+k-1}{k-1}$ / 2^{n-1} / $\binom{n-1}{k-1}$
Unlabeled Bins	$\sum_{m=1}^{k} S(n,m)$ / $B(n)$ / $S(n,k)$	$\sum_{m=1}^{k} p_m(n)$ / $p(n)$ / $p_k(n)$

5 Binomial Theorem and Related Identities

Recall the identities that we have found already to be useful:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0 \qquad \text{For } n > 0$$

These can all be generalized using the binomial theorem.

5.1 The Binomial Theorem

> **Theorem 5.1: The Binomial Theorem**
>
> For all non-negative integers n:
> $$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Proof. We can prove the binomial theorem by induction on n.

- Base Case: Let $n = 1$. Then:

$$(x+y)^1 = x + y = \binom{1}{0} x^1 y^0 + \binom{1}{1} x^0 y^1$$

- Inductive Step: Suppose that the binomial theorem is true for n. We want to show that this is true for $n+1$. So, by doing some algebra:

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

This concludes the proof. $\square$

5.2 Applications

There are some important theorems from the textbook.

Theorem 5.2

For all positive integers n, the alternating sum of binomial coefficients $\binom{n}{k}$ is zero. In other words:

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

Proof. Let $x = -1$ and $y = 1$. Substitute this into the binomial theorem to get the result. $\square$

Theorem 5.3

For all non-negative integers n, the identity holds:

$$2^n = \sum_{k=0}^{n}\binom{n}{k}$$

Proof. Let $x = y = 1$. Substitute this into the binomial theorem to get the result. $\square$

Theorem 5.4

For all non-negative integers n, the identity holds:

$$n2^{n-1} = \sum_{i=0}^{n}i\binom{n}{i}$$

Proof. Take the derivative of both sides of the binomial theorem with respect to x to get $n(x + y)^{n-1} = \sum_{i=0}^{n}i\binom{n}{i}x^{i-1}y^{n-i}$. Then, let $x = y = 1$. $\square$

Theorem 5.5

For all non-negative integers n and k, the identity holds:

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

Theorem 5.6

For all non-negative integers k and n:

$$\binom{k}{k} + \binom{k+1}{k} + \binom{k+2}{k} + \cdots + \binom{n}{k} = \binom{n+1}{k+1}$$

Theorem 5.7

For positive integers n and k:

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

Theorem 5.8: Vandermonde's Identity

For all positive integers n, m, and k:

$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}$$

Definition 5.1: Multinomial Coefficients

Let $n = \sum_{i=1}^{k} a_i$, where n and $a_1, a_2, \ldots, a_k$ are non-negative integers. We define:

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

Theorem 5.9: Multinomial Theorem

For all non-negative integers n and k, the following equality holds.

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{a_1, a_2, \ldots, a_k} \binom{n}{a_1, a_2, \ldots, a_k} x_1^{a_1} x_2^{a_2} \ldots x_k^{a_k}$$

Here, the sum is taken over all k-tuples of non-negative integers $a_1, a_2, \ldots, a_k$ such that $n = \sum_{i=1}^{k} a_i$.

Theorem 5.10

For all non-negative integers n and $a_1, a_2, \ldots, a_k$ such that $n = \sum_{i=1}^{n} a_i$, the following equality holds:

$$\binom{n}{a_1, a_2, \ldots, a_k} = \binom{n}{a_1} \ldots \binom{n - a_1 - \cdots - a_i}{a_{i+1}} \ldots \binom{n - a_1 - \cdots - a_{k-1}}{a_k}$$

Definition 5.2: Generalized Binomial Coefficient

Let m be any <u>real</u> number and let k be a non-negative integer. Then, $\binom{m}{0} = 1$. Additionally, for $k > 0$:

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

Theorem 5.11: Generalized Binomial Theorem

Let m be any <u>real</u> number. Then:

$$(1+x)^m = \sum_{n \geq 0} \binom{m}{n} x^n = 1 + mx + m(m-1)\frac{x^2}{2} + \cdots + (m)_k \frac{x^k}{k!}$$

Where the sum is taken over all non-negative integers n.

6 Permutations and Cycle Structures

Now, we will talk more about permutations and, in particular, cycle structures.

6.1 Permutations

> **Definition 6.1: Permutation**
>
> A **permutation** of $[n]$ is a list of the numbers from 1 to n with each number appearing exactly once. Looking at the positions of each number, this is equivalent to a bijection $\pi : [n] \to [n]$.

6.1.1 Example: Number of Permutations

How many permutations of $[n]$ are there?

- The answer is **n!**. At the beginning, we have n choices. Once we choose one number, we are left with $n-1$ choices. After removing that number, we have $n-2$ choices. Keep on doing this, we will eventually have $(n)(n-1)(n-2)\ldots(2)(1)$ choices, or $n!$.

- Another way of seeing it is: There are n options for $\pi(1)$, then $n-1$ options for $\pi(2)$, and so on until we have 1 option left for $\pi(n)$.

6.2 Repeated Application

Suppose that you have a permutation π of $[n]$ and some integer $1 \le m \le n$. What happens when you repeatedly apply π to m? That is:
$$m \to \pi(m) \to \pi(\pi(m)) \to \ldots$$

> **Lemma 6.1**
>
> Let $\pi : [n] \to [n]$ be a permutation and let $x \in [n]$. Then, there exists a positive integer $1 \le i \le n$ so that $\pi^i(x) = x$.

6.3 Cycles

So, not only does m repeat itself, it is actually the *first* element of this sequence to repeat itself. In particular, if k is the smallest integer so that $\pi^k(m) = m$, we have that m, $\pi(m)$, $\pi^2(m)$, ..., $\pi^{k-1}(m)$ are all distinct and $\pi^k(m) = m$.

From there, the sequence repeats itself. This sequence m, $\pi(m)$, $\pi^2(m)$, ..., $\pi^{k-1}(m)$ is known as a *cycle*. So, given any permutation π of $[n]$, we can partition $[n]$ into cycles.

For instance, consider `7146352`. This permutation has two cycles.
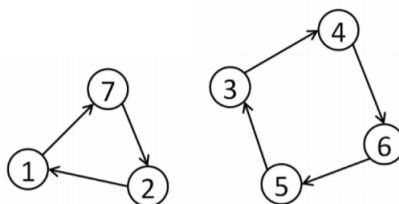
We can also look at the permutation `14253`. This permutation also has two cycles:
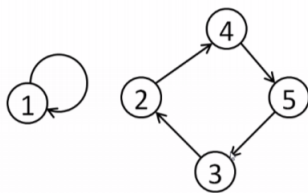
6.4   Cycle Representation

As every element is part of some cycle, if you can specify what each cycle is, you can specify the permutation.

6.4.1   Example: Cycle Representation

7146352 is the permutation with cycles:

- $1 \to 7 \to 2 \to 1$.

- $3 \to 4 \to 6 \to 5 \to 3$.

Any permutation of $[7]$ with these cycles must be 7146352.

6.4.2   Cycle Notation

To write this more conveniently:

- Write each cycle by writing the elements in order contained within parentheses.

- Write the permutation by writing each cycle in it.

6.4.3   Non-Uniqueness

One issue with this representation is that there are multiple ways to write the same permutation. This is because of two things:

- A cycle can start at any position. For example:

$$(172) = (721) = (217)$$

- Cycles can be listed in any order:

$$(172)(3465) = (3465)(172)$$

6.4.4   Canonical Cycle Notation

One way to fix this is to remove this freedom by specifying an order. One way to do this is using *canonical cycle notation*.

> **Definition 6.2: Canonical Cycle Notation**
>
> For any given permutation of $[n]$, we can write it out in Canonical Cycle Notation. In this form, each cycle will be written with its largest element first, and cycles will be written in increasing order of their first elements.

6.4.5   Example: CCN of 58237614

Consider the permutation, 58237614.

- Cycles: (157), (2843), (6).

- In CCN, we write these cycles as: (715), (8432), (6).

- So, in CCN, the whole permutation is written as: $(6)(715)(8432)$.

6.5 Permutations with Particular Cycle Structure

> **Theorem 6.1: Permutations with Particular Cycle Structures**
>
> Let a_1, a_2, ..., a_n be non-negative integers so that the equality $\sum_{i=1}^{n} i \cdot a_i = n$ holds. Then, the number of n-permutations with a_i cycles of length i where $i \in [n]$ is:
>
> $$\frac{n!}{a_1! a_2! \ldots a_n! 1^{a_1} 2^{a_2} \ldots n^{a_n}}$$

6.5.1 Example: Counting

How many permutations of $[9]$ have three cycles of length 1 and two cycles of length 3?

- Idea: Writing in cycle notation and this permutation must be:

$$(x_1)(x_2)(x_3)(x_4 x_5 x_6)(x_7 x_8 x_9)$$

Where x_1, x_2, ..., x_9 are 1, 2, ..., 9 in some order. Then, there are 9! ways to assign each x_i a different elements of $[9]$. Is that the answer? Nope.

This is because we are overcounting. This is not canonical cycle notation because the same permutation can be written in multiple ways. Instead, we consider the following:

- We can rearrange the three 1-cycle in $3! = 6$ ways.
- Each of the *three* 1-cycles can start with any 1 element.
- We can rearrange the two 3-cycle in $2! = 2$ ways.
- Each of the *two* 3-cycles can start with any of their 3 elements.

So, each such permutation can be written in $3! 1^3 2! 3^2$ many ways. Thus, the final answer is:

$$\frac{9!}{3! 1^3 2! 3^2} = 3360$$

6.6 Stirling Numbers

> **Definition 6.3: Stirling Number of the First Kind**
>
> The number of permutations of $[n]$ with k cycles is called a *unsigned Stirling number of the first kind* and is given by:
> $$c(n, k) = \text{Number of Permutations of } [n] \text{ with } k \text{ Cycles.}$$
> The number $s(n, k) = (-1)^{n-k} c(n, k)$ is called a *Stirling number of the first kind*.

Remark: There are some properties of Stirling numbers to keep in mind.

- $c(n, n) = 1$ (Must have n 1-cycle).

- $c(n, 1) = (n - 1)!$ (One n-cycle).

- $c(n, n - 1) = \frac{n(n-1)}{2}$ (One 2-cycle and $n - 2$ 1-cycles).

- $c(n, 2) = \sum_{k=1}^{n} \frac{n!}{2k(n-k)}$

> **Theorem 6.2**
>
> Let n and k be positive integers such that $n \geq k$. Then:
>
> $$c(n, k) = c(n - 1, k - 1) + (n - 1)c(n - 1, k)$$

Proof. This can be broken down into two cases. Suppose we have n people and k tables arranged in a circle. Additionally, suppose $c(n, k)$ is the total number of distinct seating arrangements of n people into k tables.

- <u>Case 1:</u> Suppose the nth person sits at his or her own table. Then, we have $n - 1$ people and $k - 1$ tables left. In this sense, it follows that there are $c(n - 1, k - 1)$ ways to arrange the $n - 1$ remaining people with the $k - 1$ remaining tables.

- <u>Case 2:</u> Now suppose that all but the nth person has been seated at every table; in particular, every table has at least one person. This amounts to $c(n - 1, k)$. Of course, we need to seat the nth person somewhere! So, how many distinct ways can we seat the nth person? Well, we can put the nth person next to one of the other people; since there are $n - 1$ remaining people, we can put the nth person next to someone to get a unique seating arrangement. Thus, there are $(n - 1)c(n - 1, k)$ ways to place our nth person.

All possible arrangements of n people into k tables either have person n seated by themselves or amongst the others, so it follows that:

$$c(n, k) = c(n - 1, k - 1) + (n - 1)c(n - 1, k)$$

Where $n \geq 1$ and $k \geq 1$. $\qquad \square$

Lemma 6.2

Let n be a fixed positive integer. Then:

$$\sum_{k=0}^{n} c(n, k)x^k = x(x + 1) \ldots (x + n - 1)$$

6.7 Study of Individual Cycle Sizes

Lemma 6.3

Let $i \in [n]$. Then, for all $k \in [n]$, there are exactly $(n - 1)!$ permutations of length of length n in which the cycle containing i is of length k.

Proof. The proof by direct counting is intuitive:

- $n - 1$ options for $\pi(n)$ (anything but n).

- $n - 2$ options for $\pi^2(n)$ (anything but n or $\pi(n)$).

- $n - 3$ options for $\pi^3(n)$ (anything but n or $\pi(n)$ or $\pi^2(n)$).

- ...

- $n - k + 1$ options for $\pi^{k-1}(n)$.

- 1 option for $\pi^k(n) = n$.

- $(n - k)!$ options for the permutation of the other $n - k$ elements.

- Multiply together to get $(n - 1)!$ total elements. $\qquad \square$

There is another proof we can use, and this uses canonical cycle notation:

Proof. When you write the permutation in canonical cycle notation, the cycle containing n is exactly everything listed from the appearance of n onward. So, n is in a cycle of length k if and only if n is written kth from last. There are $(n - 1)!$ ways to arrange the other $n - 1$ elements into the other $n - 1$ spots. $\qquad \square$

6.8   EVEN and ODD

Now, we will talk about the notion of *even* and *odd* cycles, or cycles that have only cycles of even and odd lengths, respectively.

Lemma 6.4

For all positive integers m, the equality $ODD(2m) = EVEN(2m)$ holds.

For the above lemma, there is a nice bijection that you can use to prove this. In particular, the algorithm behind this is as follows.

- Work from the last cycle to the first cycle.

- If the last element of the last cycle is *smaller* than the first element of the next-to-last cycle, then move that element from the last cycle to the end of the next-to-last cycle.

- *Otherwise*, move the last element from the last cycle into its own singleton between said cycle and the one before it.

- Move on to the next *untouched* cycle.

Theorem 6.3

For all positive integers m:

$$ODD(2m) = EVEN(2m) = (2m-1)^2(2m-3)^2(2m-5)^2 \ldots 3^2 1^2$$

Theorem 6.4

For all positive integers m:

$$ODD(2m+1) = (2m+1) \cdot ODD(2m) = (2m+1)(2m-1)^2(2m-3)^2(2m-5)^2 \ldots 3^2 1^2$$

Additionally:

$$EVEN(2m+1) = 0$$

---

7 Inclusion-Exclusion

Now, we will briefly discuss inclusion-exclusion.

7.1 Addition Rule

Recall that if S and T are <u>disjoint</u> sets, then:

$$|S \cup T| = |S| + |T|$$

7.1.1 Question: Student Majors

Suppose that registered for this class, we have 100 math majors, 15 computer science majors, 10 economics majors, and no one else. How many students are there total?

- Impossible to determine. What if there are double majors? If there are double majors, they will be counted in more than one category.

7.1.2 Question: Student Majors

Suppose we have 15 computer science majors and 10 economics major, but 2 of them are CS-Economics double major. How many do we have in total?

- There are $15 - 2 = 13$ pure computer science majors, $10 - 2 = 8$ pure economics majors, and 2 double majors. These are all disjoint sets of people. So, there are $13 + 8 + 2 = 23$ people in total.

7.2 Generalized Addition Rule

> **Theorem 7.1**
>
> For any sets S and T we have that:
>
> $$|S \cup T| = |S| + |T| - |S \cap T|$$

Remark: If $|S \cap T| = 0$, then this recovers the usual addition rule.

7.3 The Theorem

> **Theorem 7.2: Generalized Inclusion-Exclusion**
>
> Given any finite sets $S_1, S_2, \ldots, S_n$, we have that:
>
> $$|S_1 \cup S_2 \cup \cdots \cup S_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n} |S_{i_1} \cap S_{i_2} \cap \cdots \cap S_{i_k}|$$
>
> More clearly, we have:
>
> $$\left| \bigcup_{i=1}^{n} S_i \right| = \sum_{i=1}^{n} |S_i| - \sum_{i<j} |S_i \cap S_j| + \sum_{i<j<k} |S_i \cap S_j \cap S_k| - \cdots + (-1)^{n-1} |S_1 \cap \ldots S_n|$$

7.3.1 For Two Sets

Inclusion-exclusion for two sets A and B is simply:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

7.3.2   For Three Sets

Inclusion-exclusion for three sets A, B, and C is simply:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|$$

7.3.3   For Four Sets

Inclusion-exclusion for four sets A, B, C, and D is simply:

$$\begin{aligned}
|A \cup B \cup C \cup D| = {} & |A| + |B| + |C| + |D| \\
& - |A \cap B| - |A \cap C| - |A \cap D| - |B \cap C| - |B \cap D| - |C \cap D| \\
& + |A \cap B \cap C| + |A \cap B \cap D| + |A \cap C \cap D| + |B \cap C \cap D| \\
& - |A \cap B \cap C \cap D|
\end{aligned}$$

7.4   Applications

Here, we discuss some applications of Inclusion-Exclusion.

7.4.1   Stirling Numbers

Recall that the number of way to put n labeled balls into k labeled, non-empty boxes is $k!S(n,k)$ (First, pick the set partition, then pick which part goes into which box). Recall that this is the same as the number of *surjections* from a set of size $[n]$ to a size $[k]$. If we can count this, then we can compute Stirling numbers.

To motivate this, consider the question: *How many surjections are there $[n] \to [3]$?*

- There are 3^n total functions, but not all are surjections. Which are we missing?
  - Functions that don't have 1 in their range.
    * How many functions $[n] \to [3]$ don't have 1 in their image? These are functions $[n] \to \{2,3\}$, or simply 2^n such functions. The similar argument can be made for the other cases.
  - Functions that don't have 2 in their range.
  - Functions that don't have 3 in their range.
- Combining these, we have $3^n - 2^n - 2^n - 2^n$. However, we aren't done. These removed sets overlap, so we need to use inclusion-exclusion. Let's do this more formally.
  - Let A_1, A_2, and A_3 denote the sets of functions missing 1, 2, and 3 respectively. Then, we want $3^n - |A_1 \cup A_2 \cup A_3|$.
  - Using Inclusion-Exclusion:

  $$\boxed{3^n - |A_1| - |A_2| - |A_3| + |A_1 \cap A_2| + |A_2 \cap A_3| + |A_1 \cap A_3| - |A_1 \cap A_2 \cap A_3|}$$

  - What is $|A_1 \cap A_2|$? The functions taking value neither 1 nor 2. In other words, this is only the function that assigns 3 to everything. This applies to the other pairwise intersections.
  - What about $|A_1 \cap A_2 \cap A_3|$? These are the functions not taking values 1, 2, or 3. There are no functions that have this criteria.
  - So, the answer is simply:

  $$3^n - 2^n - 2^n - 2^n + 1 + 1 + 1 = \boxed{3^n - 3 \cdot 2^n + 3}$$

  So, therefore:

  $$\boxed{S(n,3) = \frac{3^{n-1}}{2} + 2^{n-1} + \frac{1}{2}}$$

Generalizing this, $S(n, k)$ represents the number of partitions of $[n]$ into k parts. If we label the parts $1, 2, \ldots, k$ (can be done in $k!$ ways), we get a surjection $[n] \to [k]$. So:

$$\{\text{Number of Surjections } [n] \to [k]\} = k!S(n, k)$$

How many surjections are there? Well, counting all functions is easy. A function fail to be a surjection if and only if some value fails to be in the image. Let S be the set of all functions $[n] \to [k]$. Let A_i be the set of functions $[n] \to [k]$ with i not in the image. So, the number of surjections is simply:

$$\begin{aligned}
\text{Num. Surjections} &= |S| - |A_1 \cup A_2 \cup \cdots \cup A_k| \\
&= k^n - |A_1 \cup A_2 \cup \cdots \cup A_k| \\
&= k^n - \sum_{m=1}^{k}(-1)^{m-1} \sum_{1 \le i_1 < i_2 < \cdots < i_m \le k} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}|
\end{aligned}$$

But since $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}$ is the set of functions from $[n]$ to $[k]$ such that $i_1, i_2, \ldots, i_m$ do not appear in the image, we have that:

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}| = \text{Num. Functions From } [n] \to [k - \{i_1, i_2, \ldots, i_m\}]$$

Thus:

$$\begin{aligned}
\text{Num. Surjective Functions} &= k^n - \sum_{m=1}^{k}(-1)^{m-1} \sum_{1 \le i_1 < i_2 < \cdots < i_m \le k} |A_{i_1} \cap A_{i_2}(k-m)^n \\
&= k^n - \sum_{m=1}^{k}(-1)^{m-1}(k-m)^n \binom{k}{m} \\
&= k^n + \sum_{m=1}^{k}(-1)^{m}(k-m)^n \binom{k}{m} \\
&= \sum_{m=0}^{k}(-1)^{m}(k-m)^n \binom{k}{m}
\end{aligned}$$

Putting this together, we have:

$$k!S(n, k) = \sum_{m=0}^{k}(-1)^{m}(k-m)^n \binom{k}{m} \implies S(n, k) = \sum_{m=0}^{k}(-1)^{m}\frac{(k-m)^n}{k!}$$

Which leads to the following theorem:

Theorem 7.3: Stirling Numbers of the Second Kind

For all positive integers n and k, the following equality holds:

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

7.4.2   Derangements

We want to count the number of permutations $\pi : [n] \to [n]$ satisfying a bunch of conditions:

- $\pi(1) \neq 1$.

- $\pi(2) \neq 2$.

- $\ldots$

- $\pi(n) \neq n$.

Each condition is relatively easy to deal with individually, but many at once is hard. Combining these, the conditions are hard. But, if we negate them, it becomes easier. In other words, we don't want any of the following:

- $\pi(1) = 1$.

- $\pi(2) = 2$.

- $\ldots$

- $\pi(n) = n$.

Letting A_i be the permutations π of $[n]$ with $\pi(i) = i$, we let:

$$D_n = n! - |A_1 \cup A_2 \cup \cdots \cup A_n|$$

Then, using inclusion-exclusion:

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|$$

We first ask: how big is $|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|$? This is the set of permutations of $[n]$ so that:

- $\pi(i_1) = i_1$.

- $\pi(i_2) = i_2$.

- $\ldots$

- $\pi(i_k) = i_k$.

These k values are fixed, and we can use any permutation on the other $n - k$ elements, so the total size of $(n-k)!$. Thus:

$$D_n = n! - \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|$$

$$= n! - \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} (n - k)!$$

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

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

$$= n! \left(1 - \sum_{k=1}^{n} (-1)^{k-1} \frac{1}{k!} \right)$$

This gives us the following definition:

Definition 7.1: Derangements

A derangement is a permutation with no 1-cycles. This is denoted by D_n, or the number of derangements of $[n]$ with no 1-cycles, and can be found using the formula:

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

---

8   Generating Functions

In order to study some interesting sequence of numbers, $a_0, a_1, a_2, \ldots$, we can turn these numbers into a single function $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots$ and study $f(x)$.

The point is that because the elements of the sequence are just the coefficients of x in this power series, $f(x)$ encodes the sequence that we are interested in.

> **Definition 8.1: Ordinary Generating Functions**
>
> Let $\{f_n\}_{n \geq 0}$ be a sequence of real numbers. Then, the formal power series $F(x) = \sum_{n \geq 0}^{\infty} f_n x^n$ is called the *ordinary generating function* of the sequence $\{f_n\}_{n \geq 0}$.

8.1   What is a Generating Function?

We can interpret a generating function $F(x)$ in two ways:

- As an **actual function** of a (real or complex) variable x, this should converge for x in some range.

- A **formal power series**, namely a set of symbols that can be manipulated in the same way real function could, but that can't necessarily be evaluated anywhere.

8.2   Purpose

Why is this a good idea?

1. The function f can encode the sequence $\{a_i\}$ within a single function.

2. Complicated combinatorical properties of the sequence $\{a_i\}$ can often be encoded as algebraic or analytic properties of $f(x)$.

3. This often lets us reduce solving complicated problems to basic algebra.

8.3   Application: Recurrence Relations

A *recurrence relation* is a way of defining a sequence of numbers by giving a formula for each element in terms of the previous ones.

To find the closed form of a recurrence relation, we can make use of generating functions.

8.3.1   Example 1: Basic Recurrence

To better demonstrate this, consider the recurrence relation[1]:

$$a_n = \begin{cases} 2 & n = 0 \\ 4 & n = 1 \\ 3a_{n-1} - 2 & \text{Otherwise} \end{cases}$$

To find the closed form for this recurrence relation:

(1) Define the generating function $F(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \ldots$.

(2) Account for the base case(s). These will usually take up a_0, a_1, and so on. In this case:

$$F(x) = \boxed{a_0 + a_1 x} + a_2 x^2 + a_3 x^3 + \cdots = \boxed{2 + 4x} + a_2 x^2 + \ldots$$

---

[1]The following steps can be applied to exponential generating functions.

(3) For the remaining a_i terms, substitute the generating function in its place. Then, expand it out.

$$F(x) = 2 + 4x + \boxed{\sum_{n=0}^{\infty} a_{n+2}x^{n+2}}$$

$$= 2 + 4x + \boxed{\sum_{n=2}^{\infty} a_n x^n}$$

$$= 2 + 4x + \boxed{\sum_{n=2}^{\infty} (3a_{n-1} - 2)x^n}$$

$$= 2 + 4x + \boxed{\sum_{n=2}^{\infty} 3a_{n-1}x^n - \sum_{n=2}^{\infty} 2x^n}$$

(4) Change all instances of the summation into $F(x)$ defined in step 1 or another known generating function. This will usually involve shifting the summations.

$$F(x) = 2 + 4x + \sum_{n=2}^{\infty} 3a_{n-1}x^n - \sum_{n=2}^{\infty} 2x^n = 2 + 4x + \sum_{n=1}^{\infty} 3a_n x^{n+1} - 2\sum_{n=2}^{\infty} x^n$$

$$= 2 + 4x + 3x\sum_{n=1}^{\infty} a_n x^n - 2\sum_{n=0}^{\infty} x^{n+2} = 2 + 4x + 3x\left(\sum_{n=0}^{\infty} a_n x^n - a_0 x^0\right) - 2x^2\sum_{n=0}^{\infty} x^n$$

$$= 2 + 4x + 3x\left(\boxed{\sum_{n=0}^{\infty} a_n x^n} - 2\right) - \boxed{2x^2\sum_{n=0}^{\infty} x^n}$$

$$= 2 + 4x + 3x\left(\boxed{F(x)} - 2\right) - \boxed{\frac{2x^2}{1-x}}$$

$$= 2 + 4x + 3xF(x) - 6x - \frac{2x^2}{1-x}$$

(5) Isolate $F(x)$ so that you get a rational function.

Since we have:

$$F(x) = 2 + 4x + 3xF(x) - 6x - \frac{2x^2}{1-x}$$

It follows that:

$$F(x) - 3xF(x) = 2 + 4x - 6x - \frac{2x^2}{1-x} \iff F(x) = \frac{2 - 2x - \frac{2x^2}{1-x}}{1-3x} = \boxed{\frac{2 - 4x}{(1-x)(1-3x)}}$$

This gives us the **closed form** of the generating function.

(6) Use partial fraction decomposition on $F(x)$.

In particular, using partial fraction decomposition, we have:

$$F(x) = \frac{A}{(1-x)} + \frac{B}{(1-3x)}$$

So:

$$\frac{A}{(1-x)} + \frac{B}{(1-3x)} = \frac{A(1-3x) + B(1-x)}{(1-x)(1-3x)}$$

We want:
$$2 - 4x = A(1 - 3x) + B(1 - x) = (A + b) - (3A + B)x$$

Solving for A and B, we have:
$$A = B = 1$$

And, as a result, we have that:
$$F(x) = \frac{1}{1 - x} + \frac{1}{1 - 3x}$$

(7) Find a new generating function that encodes this. Then, find the explicit formula from the coefficient of the generating function.

We know how to encode each of the terms of the generating function. That is:
$$\frac{1}{1 - x} = \sum_{n=0}^{\infty} x^n \qquad \frac{1}{1 - 3x} = \sum_{n=0}^{\infty} (3x)^n$$

So, our generating function is:
$$F(x) = \sum_{n=0}^{\infty} x^n + \sum_{n=0}^{\infty} (3x)^n$$

Following up on this, we have:
$$F(x) = \sum_{n=0}^{\infty} x^n + \sum_{n=0}^{\infty} (3x)^n$$
$$= \sum_{n=0}^{\infty} x^n + \sum_{n=0}^{\infty} 3^n x^n$$
$$= \sum_{n=0}^{\infty} \boxed{(1 + 3^n)} x^n$$

So, our explicit formula is:
$$a_n = 1 + 3^n$$

8.3.2 Example 2: Fibonacci

Consider the Fibonacci sequence defined by:
$$a_n = \begin{cases} 1 & n = 0 \\ 1 & n = 1 \\ a_{n-1} + a_{n-2} & \text{Otherwise} \end{cases}$$

(1) Define the generating function $F(x) = \sum_{n=0}^{\infty} a_n x^n$.

(2) Account for the base case(s).

$$F(x) = \boxed{a_0 + a_1 x^1} + a_2 x^2 + a_3 x^3 + \cdots = \boxed{1 + x} + a_2 x^2 + a_3 x^3 + \ldots$$

(3) For the remaining a_i terms, substitute the generating function in its place.

$$F(x) = 1 + x + \boxed{\sum_{n=2}^{\infty} (a_{n-1} + a_{n-2}) x^n}$$
$$= 1 + x + \boxed{\sum_{n=2}^{\infty} a_{n-1} x^n + \sum_{n=2}^{\infty} a_{n-2} x^n}$$

(4) Change all instances of the summation into $F(x)$ defined in step 1 or another known generating function. This will usually involve shifting the summations.

$$
\begin{aligned}
F(x) &= 1 + x + \sum_{n=2}^{\infty} a_{n-1}x^n + \sum_{n=2}^{\infty} x^n \\
&= 1 + x + \sum_{n=0}^{\infty} a_{n+1}x^{n+2} + \sum_{n=0}^{\infty} a_n x^{n+2} \\
&= 1 + x + x\sum_{n=0}^{\infty} a_{n+1}x^{n+1} + x^2 \sum_{n=0}^{\infty} a_n x^n \\
&= 1 + x + x\left(\sum_{n=0}^{\infty} a_n x^n - a_0 x^0\right) + x^2 \sum_{n=0}^{\infty} a_n x^n \\
&= 1 + x + x\left(\sum_{n=0}^{\infty} a_n x^n - 1\right) + x^2 \sum_{n=0}^{\infty} a_n x^n \\
&= 1 + x + x\left(F(x) - 1\right) + x^2 F(x)
\end{aligned}
$$

(5) Isolate $F(x)$ so that you get a rational function.

Since we have:

$$
F(x) = 1 + x + x(F(x) - 1) + x^2 F(x) \iff F(x) = 1 + xF(x) + x^2 F(x)
$$

It follows that:

$$
F(x) - xF(x) - x^2 F(x) = 1 \iff (1 - x - x^2)F(x) = 1 \iff F(x) = \boxed{\dfrac{1}{1 - x - x^2}}
$$

(6) Use partial fraction decomposition on $F(x)$.

This function isn't exactly super nice when it comes to partial fraction decomposition. That being said, we know that:

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

So, it follows that:

$$
\begin{aligned}
F(x) &= \frac{A}{\left(x - \frac{-1+\sqrt{5}}{2}\right)} + \frac{B}{\left(x - \frac{-1-\sqrt{5}}{2}\right)} \\
&= \frac{1}{\sqrt{5}} \frac{1}{\left(x - \frac{-1+\sqrt{5}}{2}\right)} - \frac{1}{5} \frac{1}{\left(x - \frac{-1-\sqrt{5}}{2}\right)} \\
&= \frac{1}{\sqrt{5}}\left(\frac{1}{\left(x - \frac{-1+\sqrt{5}}{2}\right)} - \frac{1}{\left(x - \frac{-1-\sqrt{5}}{2}\right)}\right)
\end{aligned}
$$

(7) Find a new generating function that encodes this. Then, find the explicit formula from the coefficient of the generating function.

Let $\alpha = \frac{-1+\sqrt{5}}{2}$ and $\beta = \frac{-1-\sqrt{5}}{2}$. Then:

$$
\begin{aligned}
F(x) &= \frac{1}{\sqrt{5}} \left(\frac{1}{x-\alpha} - \frac{1}{x-\beta} \right) \\
&= \frac{1}{\sqrt{5}} \left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}x - 1} - \frac{\frac{1}{\beta}}{\frac{1}{\beta}x - 1} \right) \\
&= \frac{1}{\sqrt{5}} \left(\frac{\frac{1}{\beta}}{1 - \frac{1}{\beta}x} - \frac{\frac{1}{\alpha}}{1 - \frac{1}{\alpha}x} \right) \\
&= \frac{1}{\sqrt{5}} \left(\sum_{n=0}^{\infty} \left(\frac{1}{\beta} \right)^{n+1} x^n - \sum_{n=0}^{\infty} \left(\frac{1}{\alpha} \right)^{n+1} x^n \right) \\
&= \sum_{n=0}^{\infty} \frac{1}{\sqrt{5}} \left(\left(\frac{1}{\beta} \right)^{n+1} - \left(\frac{1}{\alpha} \right)^{n+1} \right) x^n
\end{aligned}
$$

So, our explicit formula is:

$$
\begin{aligned}
a_n &= \frac{1}{\sqrt{5}} \left(\left(\frac{1}{\beta} \right)^{n+1} - \left(\frac{1}{\alpha} \right)^{n+1} \right) \\
&= \frac{1}{\sqrt{5}} \left[\left(\frac{2}{-1-\sqrt{5}} \right)^{n+1} - \left(\frac{2}{-1+\sqrt{5}} \right)^{n+1} \right] \\
&= \frac{1}{\sqrt{5}} \left[\left(\frac{1-\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1+\sqrt{5}}{2} \right)^{n+1} \right]
\end{aligned}
$$

8.4   Tools and Operations

Remark: There are several basic tools and operations that we can use to evaluate generating functions.

- If two generating functions are the same, the coefficients are the same. That is, if $\sum_{n\geq 0}^{\infty} a_n x^n = \sum_{n\geq 0}^{\infty} b_n x^n$ for all n, then $a_n = b_n$ for all n.

- Geometric Series:
$$
\frac{1}{1-cx} = \sum_{n\geq 0}^{\infty} (cx)^n = \sum_{n\geq 0}^{\infty} c^n x^n
$$

- Sums of Generating Functions:
$$
\sum_{n\geq 0}^{\infty} a_n x^n + \sum_{n\geq 0}^{\infty} b_n x^n = \sum_{n\geq 0}^{\infty} (a_n + b_n) x^n
$$

- Shifts: If $F(x) = \sum_{n\geq 0}^{\infty} a_n x^n$, then
$$
xF(x) = \sum_{n\geq 0}^{\infty} a_n x^{n+1} = \sum_{n>0}^{\infty} a_{n-1} x^n
$$

- Scaling: We can multiply a generating function by a constant, which scales every term in the associated sequence by the same constant. For example, given the generating function:
$$
F(x) = \sum_{n\geq 0}^{\infty} a_n x^n = a_0 x^0 + a_1 x^1 + a_2 x^2 + \dots
$$

Multiplying this generating function by a constant c gives:

$$cF(x) = c\sum_{n\geq 0}^{\infty} a_n x^n = ca_0 x^0 + ca_1 x^1 + ca_2 x^2 + \ldots$$

- Differentiation: We can also take the derivative of a generating function, which in turn will take the derivative of every term in the associated sequence. For example, given the generating function:

$$F(x) = \sum_{n\geq 0}^{\infty} a_n x^n = a_0 x^0 + a_1 x^1 + a_2 x^2 + \ldots$$

Taking the derivative gives us:

$$\begin{aligned}
\frac{d}{dx}F(x) &= \frac{d}{dx}\sum_{n\geq 0}^{\infty} a_n x^n \\
&= \sum_{n\geq 1}^{\infty} na_n x^{n-1} \\
&= \sum_{n\geq 0}^{\infty} (n+1)a_{n+1} x^n \\
&= \frac{d}{dx}a_0 x^0 + \frac{d}{dx}a_1 x^1 + \frac{d}{dx}a_2 x^2 + \ldots \\
&= a_1 + 2a_2 x + 3a_3 x^2 + \ldots
\end{aligned}$$

8.5 Basic Generating Functions

There are several basic generating functions to know. More are listed in the appendix.

- For the sequence 1, 1, 1, ...

$$F(x) = \sum_{n\geq 0}^{\infty} x^n = \frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \ldots$$

- For the sequence 1, -1, 1, -1, ...

$$F(x) = \sum_{n\geq 0}^{\infty} (-x)^n = \sum_{n\geq 0}^{\infty} (-1)^n x^n = \frac{1}{1-(-x)} = \frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \ldots$$

- For the sequence 1, 2, 3, 4, ...

$$F(x) = \left(\sum_{n\geq 0}^{\infty} x^n\right)^2 = \frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \ldots$$

- For the sequence 1, -2, 3, -4, ...

$$F(x) = \left(\sum_{n\geq 0}^{\infty} (-x)^n\right)^2 = \frac{1}{(1-(-x))^2} = \frac{1}{(1+x)^2} = 1 - 2x + 3x^2 - 4x^3 + 5x^4 - \ldots$$

- For the sequence 1, 0, 1, 0, ...

$$F(x) = \sum_{n\geq 0}^{\infty} (x^2)^n = \sum_{n=0}^{\infty} x^{2n} = \frac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 + \ldots$$

More of these are listed in the appendix.

8.6 Products of Generating Functions

> **Lemma 8.1: Products of Generating Functions**
>
> Let $\{a_n\}_{n\geq 0}$ and $\{b_n\}_{n\geq 0}$ be two sequences and let $A(x) = \sum_{n\geq 0}^{\infty} a_n x^n$ and $B(x) = \sum_{n\geq 0}^{\infty} b_n x^n$ be their respective generating functions. Then, we can define $c_n = \sum_{i=0}^{n} a_i b_{n-i}$ and let $C(x) = \sum_{n\geq 0}^{\infty} c_n x^n$. Then:
> $$A(x)B(x) = C(x)$$
> Here, we can say that the coefficients of x^n in $A(x)B(x)$ is $c_n = \sum_{i=0}^{n} a_i b_{n-i}$.

> **Theorem 8.1: The Product Formula**
>
> Suppose you have objects of type A and objects of type B. Each has a size which is a non-negative integer, and there are a_n objects of type A of size n and b_m objects of type B of size m.
>
> Then, c_k is the number of ways to find a pair of an object of type A and an object of type B where the sum of the sizes is k.

8.7 Compositions of Generating Functions

> **Definition 8.2: Composition of Generating Functions**
>
> Let $F(x) = \sum_{n\geq 0}^{\infty} f_n x^n$ be a formal power series and let G be a formal power series with constant term 0. Then, we can define:
>
> $$F(G(x)) = \sum_{n\geq 0}^{\infty} f_n (G(x))^n = f_0 + f_1 G(x) + f_2 G(x)^2 + \dots$$

> **Theorem 8.2: The Compositional Formula**
>
> Let a_n be the number of ways to build a certain structure on an n-element set and assume that $a_0 = 0$. Let b_n be the number of ways to build a second structure on an n-element set and let $b_0 = 1$. Let g_n be the number of ways to split the set $[n]$ into an unspecified number of non-empty intervals, build a structure of the first kind on each of these intervals, and then build a structure of the second kind on the set of the intervals. Let $g_0 = 1$. Then, let $A(x)$, $B(x)$, and $G(x)$ be the generating functions of the sequences $\{a_n\}$, $\{b_n\}$, $\{g_n\}$. Then:
>
> $$G(x) = B(A(x))$$

8.8 Catalan Numbers

To motivate this, consider the question: *How many ways can you produce n matching pairs of parentheses?*

For $n = 2$, we can have:

- ()()

- (())
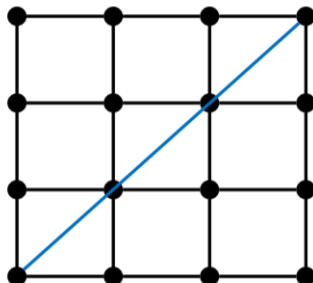
The idea is that we must have n (and n). In this sense, we could say that there are $\binom{2n}{n}$ ways to arrange these parentheses. However, not all of these work. For instance, the following are invalid (for $n = 2$):

-))((

- $)()($

The idea is that we cannot, at any point, have more closing parentheses than opening parentheses.

One way to count this is to image an $n \times n$ graph, something like:

Here, we can make steps of either $(0, 1)$ or $(1, 0)$, and the graph starts at $(0, 0)$ and ends at (n, n). So, for a parenthesis pattern, we can create a path where we can keep track of the number of left parentheses and the number of right parentheses:

$$(\text{Num. Left Paren.}, \text{Num. Right Paren.})$$

So, how many such paths are there?

> **Definition 8.3: Catalan Numbers**
>
> The nth Catalan number C_n is the number of up-left lattice paths from $(0, 0)$ to (n, n) that stay on or above the line $y = x$.

To start counting, we need to consider which parentheses (matches with. In this sense, we can say that it matches with the kth) for some k.

$$\underbrace{(\ldots)}_{k-1 \text{ pairs}} \qquad \underbrace{\ldots}_{n-k \text{ pairs}}$$

So:

- There are C_{k-1} ways to match the first group.

- There are C_{n-k} ways to match the second group.

And, as a result, we have:

$$C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}$$

Using generating functions, we arrive at the following conclusion:

> **Definition 8.4: Catalan Number Function**
>
> Recursively, the nth Catalan number is defined by:
>
> $$C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}$$
>
> Using generating functions, we can define a closed form version by:
>
> $$C_n = \frac{1}{n+1} \binom{2n}{n}$$

9 Exponential Generating Functions

Like ordinary generating functions, this is a way to associate sequence of numbers that we're interested in studying to some power series.

The only difference is that, for expoenntial generating functions, we have:

$$\{a_n\} \longleftrightarrow \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$$

Whereas for ordinary generating functions:

$$\{a_n\} \longleftrightarrow \sum_{n=0}^{\infty} a_n x^n$$

9.1 Comparison Between Ordinary and Exponential Generating Functions

Ordinary	Exponential
$F(x) = \sum_{n=0}^{\infty} a_n x^n$	$F(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$
$a_n = 1$ and $F(x) = \frac{1}{1-x}$	$a_n = 1$ and $F(x) = e^x$
$a_n = n$ and $F(x) = \frac{x}{(1-x)^n}$	$a_n = n$ and $F(x) = xe^x$
$xF(x) = \sum_{n=0}^{\infty} a_n x^{n+1} = \sum_{n>0}^{\infty} a_{n-1} x^n$	$\sum_{n=0}^{\infty} a_{n+1} \frac{x^n}{n!} = F'(x)$
Converges only if a_n grows at most exponentially.	Converges more generally.

9.2 Multiplication of Exponential Generating Function

Multiplication of ordinary generating functions is important, but it is also different compared to ordinary generating function.

Let's suppose we have two generating functions:

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

$$B(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!}$$

$$C(x) = A(x)B(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{n!}$$

Where:

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

9.3 Combinatorial Interpretation of Multiplication

- Define an A-structure to be a thing so that there are a_n A-structures on a set of size n.

- Define a B-structure to be a thing so that there are b_n B-structures on a set of size n.

- Ordinary generating function multiplication talks about the number of ways to find an A-structure and a B-structure of total size n.

- Exponential generating function multiplication has c_n be the number of ways to partition $[n]$ into two sets and put an A-structure on one and a B-structure on the other.

- If the A-structure has size k, then there are $\binom{n}{k}$ ways to partition $[n]$ into a_k A-structures and b_{n-k} B-structures.

9.4   Composition of Generating Functions

Let's suppose we have two generating functions:

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

$$B(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!}$$

$$B(A(x)) = \sum_{n=0}^{\infty} b_n \frac{A(x)^n}{n!}$$

So, the $\frac{x^n}{n!}$ coefficient is:

- b_1 times the number of partitions of $[n]$ into one part with an A-structure; plus

- b_2 times the number of partitions of $[n]$ into two parts with an A-structure on each; plus

- . . .

Summing these all together, the $\frac{x^n}{n!}$ coefficient of $B(A(x))$ counts the number of ways to partition $[n]$ into subsets, put an A-structure on each subset, and put a B-structure on the collection of subsets.

9.5   Applications

Some applications of exponential generating functions are as follows.

9.5.1   Permutations

A permutation of $[n]$ is equivalent to partitioning $[n]$ into subsets and then arranging each subset into a cycle.

The A-structure is a cycle. There are $(k-1)!$ cycles on a k-element subset so $a_k = (k-1)!$. Then:

$$A(x) = \sum_{k \geq 1}^{\infty} (k-1)! \frac{x^k}{k!} = \sum_{k \geq 1}^{\infty} \frac{x^k}{k} = \log\left(\frac{1}{1-x}\right)$$

9.5.2   Stirling Numbers

What if we count permutations of $[n]$ weighted by $y^{\text{Number of Cycles}}$?

- The A-structure is a cycle.

- The B-structure has $b_n = y^n$. Here, $B(x) = \sum_{n=0}^{\infty} x^n \frac{y^n}{n!} = e^{xy}$.

The generating function is:

$$B(A(x)) = e^{y \log\left(\frac{1}{1-x}\right)} = \left(\frac{1}{1-x}\right)^y$$

The $y^k \frac{x^n}{n!}$-coefficient is the number of permutations of $[n]$ with k cycles.

This leads the following theorem.

Theorem 9.1: Unsigned Stirling Numbers of the First Kind

$$\sum_{n,k=0}^{\infty} c(n,k) \frac{x^n}{n!} y^k = \left(\frac{1}{1-x}\right)^y$$

9.5.3   At Least Subsets

How many ways can $[n]$ be partitioned into an even number of subsets of size at least 2?

- The A-structure is the subset of size at least 2. Here, $a_n = 1$ if $n \geq 0$ and 0 otherwise. So, $A(x) = e^x - 1 - x$.

- The B-structure (which is supposed to tell you what you are supposed to do once you have these subsets) is an even number of subsets. Here, $b_n = 1$ is n is even and 0 otherwise. Our generating function is $B(x) = \cosh(x) = \frac{e^x + e^{-x}}{2}$.

The generating function is the composition:

$$\cosh(e^x - 1 - x)$$

9.5.4   Derangements

What about derangements of $[n]$?

- A-structure is a cycle of length **not** 1.

  - $a_n = (n-1)!$ unless $n = 1$, in which case $a_n = 0$. The generating function is:

$$A(x) = \log\left(\frac{1}{1-x}\right) - x$$

  - B-structure is any partition.
$$B(x) = e^x$$

Generating function is:

$$B(A(x)) = e^{\log\left(\frac{1}{1-x}\right) - x} = \frac{e^{-x}}{1-x}$$

If we wanted a permutations without 2-cycles, we would get:

$$\frac{e^{-\frac{x^2}{2}}}{1-x}$$

Here, we got this by setting $A(x)$ equal to the same generating function as above, but with $a_n = 0$ for $n = 2$ instead of $n = 1$.

9.5.5   Permutations with Even/Odd Length Cycles

What about permutations with only cycles of even length or only odd length?

- A-structure is a cycle of appropriate length.

  - $a_n = (n-1)!$ if n is even, 0 otherwise.
  - Generating function is:
$$A(x) = \sum_{n \text{ even}} \frac{x^n}{n}$$

How do you restrict a generating function to only the even terms? Let $f(x) = \sum_n a_n x^n$ be a generating function. Then:

$$f(-x) = \sum_n (-1)^n a_n x^n$$

Then:

$$f(x) + f(-x) = \sum_n (1 + (-1)^n) a_n x^n = 2 \sum_{n \text{ even}} a_n x^n$$

And so:

$$\frac{f(x) + f(-x)}{2} = \sum_{n \text{ even}} a_n x^n$$

Similarly, for odd terms:

$$\frac{f(x) - f(-x)}{2} = \sum_{n \text{ odd}} a_n x^n$$

Thus, the generating function for the A-structure of cycles of even length is:

$$\sum_{n \text{ even}} \frac{x^n}{n} = \frac{\log\left(\frac{1}{1-x}\right) + \log\left(\frac{1}{1+x}\right)}{2} = \log\left(\frac{1}{\sqrt{1-x^2}}\right)$$

It follows that the generating function for permutations with even length cycles is:

$$\boxed{\sum_{n=0}^{\infty} EVEN(n)\frac{x^n}{n!} = e^{\log\left(\frac{1}{\sqrt{1-x^2}}\right)} = \frac{1}{\sqrt{1-x^2}}}$$

The generating function for the A-structure of cycles of odd length is:

$$\sum_{n \text{ odd}} \frac{x^n}{n} = \frac{\log\left(\frac{1}{1-x}\right) - \log\left(\frac{1}{1+x}\right)}{2} = \log\left(\sqrt{\frac{1+x}{1-x}}\right)$$

Thus, the generating function for permutations with odd length cycles is:

$$\boxed{\sum_{n=0}^{\infty} ODD(n)\frac{x^n}{n!} = e^{\log\left(\sqrt{\frac{1+x}{1-x}}\right)} = \sqrt{\frac{1+x}{1-x}}}$$

9.6 Other Generating Functions

- Stirling Numbers of the Second Kind

$$\sum_{n,k=0}^{\infty} S(n,k)\frac{x^n}{n!}y^k = e^{y(e^x - 1)}$$

- Bell Numbers

$$\sum_{n=0}^{\infty} B(n)\frac{x^n}{n!} = e^{e^x - 1}$$

- Derangements

$$\sum_{n=0}^{\infty} D_n \frac{x^n}{n!} = \frac{e^{-x}}{1-x}$$

10   Pattern Avoidance in Permutations

- Patterns in Permutations

- Some Basic Counting Results

- Asymptotic Results

10.1   The Stack

A **stack** is an object that stores numbers. You can push more numbers onto it, or you can pop the most recently pushed number off.

10.1.1   Stack-Sortable Permutations

> **Definition 10.1**
>
> A permutation π of $[n]$ is stack-sortable if there is a series of operations that involve $1, 2, \ldots, n$ being pushed into a stack in π order and popped off the stack in sorted order.

For example, 312 is stack-sortable:

```
Operation           Stack              Popped
-----------------------------------------------
push(3)             [3]                []
push(1)             [3, 1]             []
pop()               [3]                [1]
push(2)             [3, 2]             [1]
pop()               [3]                [1, 2]
pop()               []                 [1, 2, 3]
```

Another example we can look at is 632154, which is stack-sortable:

```
Operation           Stack              Popped
-----------------------------------------------
push(6)             [6]                []
push(3)             [6, 3]             []
push(2)             [6, 3, 2]          []
push(1)             [6, 3, 2, 1]       []
pop()               [6, 3, 2]          [1]
pop()               [6, 3]             [1, 2]
pop()               [6]                [1, 2, 3]
push(5)             [6, 5]             [1, 2, 3]
push(4)             [6, 5, 4]          [1, 2, 3]
pop()               [6, 5]             [1, 2, 3, 4]
pop()               [6]                [1, 2, 3, 4, 5]
pop()               []                 [1, 2, 3, 4, 5, 6]
```

Yet another example we can look at is 231, which is **not** stack-sortable:

```
Operation           Stack              Popped
-----------------------------------------------
push(2)             [2]                []
push(3)             [2, 3]             []
push(1)             [2, 3, 1]          []
pop()               [2, 3]             [1]
pop()               [2]                [1, 3]
```

10.1.2   Sortability

Which permutations are stack-sortable?

- We can never stack larger values on top of smaller values.

- Must pop i before any $j > i$ shows up.

- This is a problem if some $k < i$ shows up after j.

10.1.3   Classification

> **Theorem 10.1**
>
> A permutation is stack-sortable unless there are entries $k < i < j$ so that i occurs before j which occurs before k.

Proof. If such i, j, k exists, it is not stack-sortable. But, if no such i, j, k exist, we can use the following algorithm.

- Pop the stack when the next element to be added is bigger than the one on top or if there are no more elements.

Here, we note that the elements in the stack are always in decreasing sorted order. This is because we only add elements smaller than the current top element.

- Pop i when $j > i$ is about to be pushed.

- No $k < i$ remains to be pushed.

- No $k < i$ is below i on the stack.

Therefore, the elements are popped in sorted order. $\qquad\square$

So, π is stack-sortable unless there are three entries i that comes before j that comes before k with $k < i < j$. Equivalently, unless there are $x < y < z$ with $\pi(z) < \pi(x) < \pi(y)$. Or, three entries where the second smallest comes before the biggest comes before the smallest.

10.2   Patterns

> **Definition 10.2**
>
> Given two permutations π of $[n]$ and ρ of $[m]$, we say that there is a copy of ρ in π if there are $1 \le x_1 < x_2 < \cdots < x_m \le n$ so that $\pi(x_1), \pi(x_2), \ldots, \pi(x_m)$ have the **same relative orders** as $\rho(1), \rho(2), \ldots, \rho(m)$.

For example, a permutation π is stack-sortable if and only if it doesn't have a copy of 231.

Remark: Suppose we have a permutation 42513. Let's say we select three random elements 2, 5 and 3. What we are saying is what the sorted relative order is. We know that 132 means that the smallest number goes first (1), then the biggest (3), and then the next biggest (2). Consider the comparison:

```
2 5 3
1 3 2
```

So, 253 has the same relative order between them as 132 (2 is the "smallest," then 5 is the "biggest," and then 3 is the "next biggest"). This 253 is a *copy* of the pattern 132 showing up within this larger permutation.

Remark: Consider the same permutation 42513. If we look at the elements 4, 5, and 1, then we notice that these elements have the same relative order as 2, 3, and 1. That is:

```
4 5 1
2 3 1
```

Consider 231. This is saying that the "next biggest" element is first, then the "biggest," and then the "smallest." If we look at 451, this has the same relative order (next biggest being 4, then biggest being 5, and finally 1). So, 451 is a *copy* of the pattern 231 showing up within this larger permutation. Thus, this permutation is not stack-sortable.
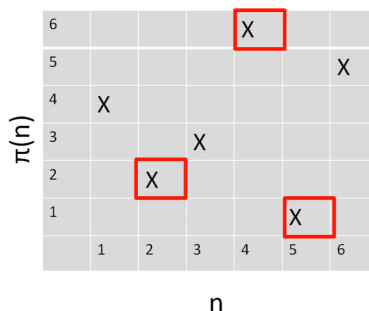
10.2.1   Visualization

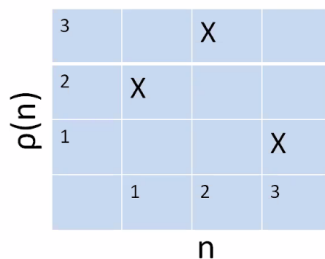It is often useful to consider graphs of the permutations involved.
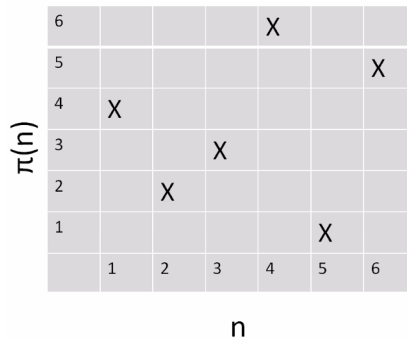
This graph represents the permutation 423615:

- $\pi(1) = 4$.

- $\pi(2) = 2$.

- $\pi(3) = 3$.

- $\pi(4) = 6$.

- $\pi(5) = 1$.

- $\pi(6) = 5$.

Suppose I selected three elements from the above graph (call this graph 1):

And now suppose that we have the following graph (call this graph 2):

Then, if we see the relative positions of the selected elements in graph 1, we notice that they have the same relative positions as the three elements in graph 2. In this sense, we can say that 261 has the same relative order as 231, so 261 is a copy of the pattern 231 showing up within this larger permutation.

10.3   Pattern Avoidance

> **Definition 10.3**
>
> For a permutation ρ, let $S_n(\rho)$ be the set of permutations of $[n]$ that avoid ρ.

Remark: By avoid ρ, we mean the permutation that does not contain a copy of ρ.

10.3.1   Counting

How many stack-sortable permutations are there? In other words, what is $|S_n(231)|$?
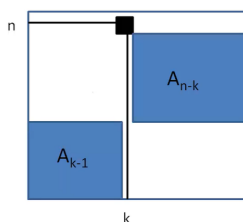
> **Theorem 10.2**
>
> $$|S_n(231)| = C_n$$
>
> Where C_n is the nth Catalan Number.

Proof. Let $A_n = |S_n(231)|$. We will show that A_n satisfies the same recurrence as C_n. Namely:

- $A_0 = 1$ (Counting the empty permutation).

- $A_n = \sum_k A_{k-1} A_{n-k}$.

Consider n in the kth location. We're going to assume that the biggest element n shows up in the kth spot. That is:

Because we don't have any copies of 231 (in particular, we don't have any copies where n fills in this 3 position), the entries on the left must be smaller than entries on the right.

- The entries on the left are $1, 2, \ldots, k-1$.

- The entries on the right are $k+1, \ldots, n$.

The entries on the left and right must be 231-avoiding. So, the number of possibilities is $A_{k-1} A_{n-k}$. Summing over k gives the full value of $A_n = \sum_k A_{k-1} A_{n-k}$. This shows the recursion holds and completes our proof. $\square$

10.4   Rotations

> **Lemma 10.1**
>
> $$|S_n(231)| = |S_n(132)| = |S_n(312)| = |S_n(213)|$$

Proof. Note that the diagrams for these permutations are simply rotations of each other.

For a permutation π, let $R\pi$ be its 90 degree permutation (π contains a copy of ρ if and only if $R\pi$ contains $R\rho$). Therefore, for any ρ, $|S_n(\rho)| = |S_n(R\rho)|$. The permutations that avoid $R\rho$ are just the rotations of those that avoid ρ, so there's a natural bijection. Then, applying this to $\rho = 231$ gives our theorem. □

10.5 Final Pattern

Consider $|S_n(123)| = |S_n(321)|$. We make use of the theorem:

> **Theorem 10.3**
>
> For all n, $|S_{123}| = |S_n(132)| = C_n$.

We want to show a bijection between the two.

10.5.1 Left-to-Right Minima

We will use the following definition:

> **Definition 10.4**
>
> The **left-to-right minima** of a permutation π are all of the indices i so that $\pi(i) < \pi(j)$ for all $j < i$.

An example of the left-to-right minima is:

10.5.2 Validity

We want to classify permutations based on the locations (in the permutation graph) of their left-to-right minima.

> **Definition 10.5**
>
> A set of pairs (i, j) with $1 \le i$ and $j \le n$ is valid is:
>
> 1. They can be ordered (i_1, j_1), (i_2, j_2), ..., (i_k, j_k) so that $i_1 < i_2 < \cdots < i_k$ and $j_1 > j_2 > \cdots > j_k$.[a]
>
> 2. For each $1 \le a \le n$, if i_{last} is the largest of the is less than or equal to a, then $j_{\text{last}} \le n + 1 - a$.
>
> ---
> [a]They have to be decreasing.

> **Lemma 10.2**
>
> A set of pairs can be the set of left-to-right minima of a permutation of $[n]$ only if it is valid.

10.5.3 Minima to Permutations

Proposition. *For every valid set S of pairs, there is exactly one* 123-*avoiding permutation and exactly one* 132-*avoiding permutation with S as its set of left-to-right minima.*

10.5.4 123-Avoiding

> **Lemma 10.3**
>
> The non-left-to-right minima in a 123-avoiding permutation are in decreasing order.

Proof. Suppose that this isn't the case.

- You have (i_1, j_1) and (i_2, j_2) with $i_1 < i_2$ and $j_1 < j_2$.

- Since (i_1, j_1) is not a minimum, there is an (i_0, j_0) with $i_0 < i_1$ and $j_0 < j_1$.

- Then, (i_0, j_0), (i_1, j_1), and (i_2, j_2) is a copy of 123. $\square$

Therefore, for every valid set, there is exactly one 123-avoiding permutation with those left-to-right minimums.

- Only one since the other terms must come in decreasing order.

- Condition (2) implies that none of these will be a new minimum.

- Left-to-right minima decreasing, which means the others are decreasing and threfore, we cannot have a three term increase.

10.5.5 132-Avoiding

We will construct our 132-avoiding permutation from the left-to-right minima set one step at a time.

- We note that $\pi(1)$ will always be one of the left-to-right minima.

- If $\pi(2)$ is not a left-to-right minima, we define it.

- If $\pi(3)$ is not a left-to-right minima, we define it.

10.5.6 Values

> **Lemma 10.4**
>
> When producing a 132-avoiding permutation, each new value must be the smallest not-yet-used value that is larger than the previous left-to-right minimum.

Proof. We know that this:

- Cannot be smaller than previous left-to-right minimum. Otherwise it would be a left-to-right minimum

- Cannot be previously used value (this is a permutation).

- Cannot be a non-minimum value. Otherwise it, the minimum value (whenever it shows up), and the previous left-to-right minimum will be a 132 pattern. $\square$

10.6   Pattern Comparison

> **Theorem 10.4**
>
> $$|S_n(1234) \leq |S_n(1324)|$$

We are going to count the number of permutations with a fixed set of left-to-right minima and a fixed set of right-to-left maxima.

10.6.1   1234-Avoiding

> **Lemma 10.5**
>
> For a 1234-avoiding permutation, the "everything else" (not a minima or maxima) must be sorted in decreasing order.

10.6.2   1324-Avoiding

> **Lemma 10.6**
>
> For every set of left-to-right minima and right-to-left maxima that admits some permutation, there is at least one consistent 1324-avoiding permutation.

10.6.3   Idea

The way we can construct this is through a swapping operation.

- If there is a 1324 in the permutation, take the 3 and 2 and swap them.

- If you keep doing this, eventually there will be a point where there won't be anymore 1324s in the permutation.

- As a result, you will have a permutation that is 1324-avoiding.

10.6.4   Proof of Result

- Each sequence of left-to-right minima and right-to-left maxima that admits a permutation has at most 1 corresponding 1234-avoiding permutation and at least 1 corresponds 1324-avoiding permutation.

Therefore:
$$|S_n(1234)| \leq \#\{\text{Min/Max Patterns}\} \leq |S_n(1324)|$$

Remark: For $n \geq 7$, the inequality is strict. We need to show that two 1324-avoiding permuations with the same min/max sequence.

10.7   Asymptotics

For every permutation ρ, there exists a constant C so that for all n:

$$|S_n(\rho)| \leq (C_\rho)^n$$

10.7.1 Asymptotic Result

What if we take $\rho = 123\ldots k$?

> **Theorem 10.5**
>
> $$|S_n(123\ldots k)| \le (k-1)^{2n}$$

Proof. Split points into levels.

- The first level is the left-to-right minima.

- The kth level is the left-to-right minima after removing the first $(k-1)$th level. □

So, a permutation that avoids $123\ldots k$ has at most $k-1$ levels and the points in each level are sorted.

10.7.2 Erdos-Szekeres

The only way to avoid $123\ldots k$ is to have few levels. These levels must be decreasing.
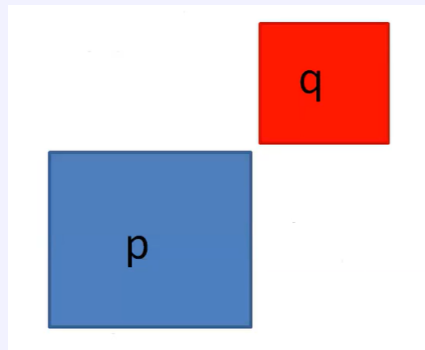
> **Theorem 10.6**
>
> Any permutation of $(n-1)(m-1)+1$ contains a copy of $123\ldots n$ or a copy of $m\ldots 321$.

Proof. If you have n or more levels, then you have $123\ldots n$. Otherwise, pigeonhole implies a level with m, which gives a copy of $m\ldots 321$. □

10.7.3 Permutation Sums

> **Definition 10.6**
>
> Given two permutations, p and q, define their sum $p \oplus q$ to be the permutation whose graphs are given as follows:
>
>

10.7.4 Sum Exponential Growth

> **Theorem 10.7**
>
> Let p and q be permutations with:
> $$|S_n(p \oplus 1)| \le C_p^n$$
> $$|S_n(1 \oplus q)| \le C_q^n$$
>
> Then:
> $$S_n(p \oplus 1 \oplus q)| \le (\sqrt{C_p} + \sqrt{C_q})^{2n}$$

11   Partial Orders

- Basic Definitions.

- Dillworth's Theorem

- Incidence Algebras and Mobius Inversions.

11.1   Partial Orders

> **Definition 11.1**
>
> A **partial order** is a relation $\geq$ on a set satisfying three properties.
>
> - $x \geq x$ for all x (reflexivity).
>
> - If $x \geq y$ and $y \geq z$, then $x \geq z$ (transitivity).
>
> - If $x \geq y$ and $y \geq x$, then $x = y$.

Remark: There may be an x and y that are <u>incomparable</u> (i.e. where neither $x \geq y$ nor $y \geq x$). In this sense, when we say **partial** order, we get some partial information (e.g. some things are bigger and so on).

11.2   Finite Partial Orders

We will study some of the combinatorics of finite partial orders on finite sets.

11.2.1   Subsets

B_n is the collection of **subsets** of $[n]$ ordered by inclusion.

- $A \geq B$ if and only if $A \supseteq B$ (A contains B)

- 2^n total elements.

11.2.2   Set Partitions and Refinements

$\sqcap_n$ is the set of set partitions of $[n]$ ordered by **refinement**.

What is a refinement? A partition P is $\geq$ a partition Q if and only if Q can be obtained from P by breaking up the parts of P into smaller parts. For example:

- $\{1, 3, 4\}\{2, 5, 6\} \geq \{1\}\{3, 4\}\{2, 5, 6\}$ (We can obtain the latter partitions from the former by taking the element 1 out of its set in the former and putting it into its own set in the latter).

11.2.3   Natural Numbers and Divisiblity

The natural numbers ($\mathbb{N}$) ordered by divisiblity. Specifically, $a \geq b$ if and only if a is a multiple of b.

For example, in this partial order:

- $10 \geq 5$ is true because $10 = 5 \cdot 2$.

- $10 \geq 7$ is false because 10 is not a multiple of 7.

11.2.4   Real Numbers and Domination

$\mathbb{R}^n$ ordered by domination. Namely, $(x_1, x_2, \ldots, x_n) \geq (y_1, y_2, \ldots, y_n)$ if and only if $x_1 \geq y_1$, $x_2 \geq y_2$, $\ldots$, $x_n \geq y_n$.

11.2.5 Permutations and Patterns

All permutations ordered by containment of patterns. Namely, $\pi \geq \rho$ if and only if π contains a copy of ρ.

11.3 Min/Max

> **Definition 11.2**
>
> An element x is in a partial order is:
>
> - A **maximal** element if there is no $y > x$.
>
> - A **maximum** element if $x \geq y$ for all y.
>
> - A **minimal** element is there is no $y < x$.
>
> - A **minimum** element if $x \leq y$ for all y.

Suppose we have a bunch of applicants for a job and these applicants are compared based on five different values.

- If there is a **maximum** element among these job applicants (someone that is better than everyone else in every values), then clearly that is the person you want to hire.

- If there isn't a maximum element, there might be several people where each of them are better than someone else at something. In this sense, you have a bunch of pairwise incomparable people. It's not clear *who* you want among that group, but you probably want a **maximal** amount. You don't want to pick somebody where there was somebody else in your applicant pool that was just better than them in all categories that you care about. You at least want somebody that isn't strictly dominated by anyone else – that would be a maximal element.

11.3.1 Example: Minimum Elements

Which of the following have minimum elements?

(a) B_n

(b) $\sqcap_n$

(c) $\mathbb{N}$

(d) $\mathbb{R}^n$

- The answer is `A`, `B`, and `C`.

  - For choice A (B_n), everything contains an empty set $\emptyset$.
  - For choice B ($\sqcap_n$), we can refine the elements so that every element in the set $[n]$ is in its own element.
  - For choice C ($\mathbb{N}$), we can have `1`, which is the minimum element possible that divides every other number.
  - For choice D ($\mathbb{R}^n$), we will always be able to find a smaller number.

11.3.2 Example: Maximum Elements

Which of the following have maximum elements?

(a) B_n

(b) $\sqcap_n$

(c) $\mathbb{N}$

(d) $\mathbb{R}^n$

- The answer is A and B.

  - For choice A (B_n), the biggest element is $\{1, 2, 3, \ldots, n\}$.
  - For choice B ($\sqcap_n$), the biggest element is (again) $\{1, 2, 3, \ldots, n\}$.
  - For choices C and D ($\mathbb{N}$ and $\mathbb{R}^n$, respectively), we can keep finding bigger elements.

11.4 Covering Elements

Definition 11.3

We say that x <u>covers</u> y if $x > y$ but there is no z with $x > z > y$.

Lemma 11.1

In any finite partial order, $x \geq y$ if and only if there is some chain of elements $x = x_0, x_1, \ldots, x_n = y$ with x_i covering x_{i+1}.

This means that if we know which elements cover which other elements, we will be able to recover the partial order.

11.5 Hasse Diagram

The lemma implies that you can specify the partial order by specifying which elements cover which others.

Definition 11.4

The **Hasse Diagram** of a partial order draws the elements of the partial order with lines between x and y if x covers y, and with x positioned vertically above y.

11.5.1 Example: Hasse Diagram for Refinement

Consider the Hasse Diagram for $\sqcap_3$:
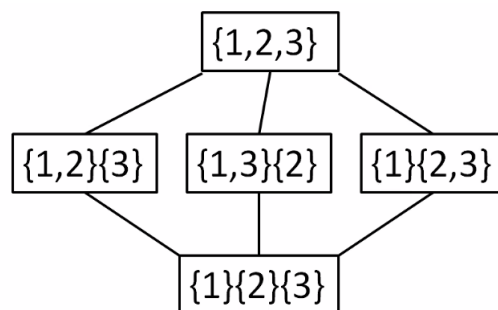
Where the bottom-most element is the minimum element and the top-most element is the maximum element.

11.6   Chains and Anti-Chains

Definition 11.5

A **chain** in a partial order is a sequence of elements

$$x_1 < x_2 < x_3 < \cdots < x_k$$

Definition 11.6

An **anti-chain** in a partial order is a collection of elements x_i so that no two of the x_i are comparable.

Remark:

- A chain is a big sequence of elements pairwise are comparable to each other. Every element in this chain can be compared to every other element in this chain.

- An anti-chain is sort of the opposite. It's essentially a bunch of elements that are pairwise incomparable.
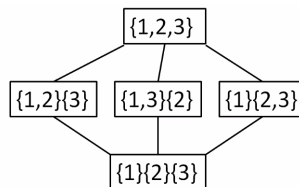
11.6.1   Covers by Chains

Definition 11.7

Given a partial order, a **cover** by chains is a collection of chains whose union is all of the elements of the order.

11.6.2   Example: Covers of $\sqcap_3$
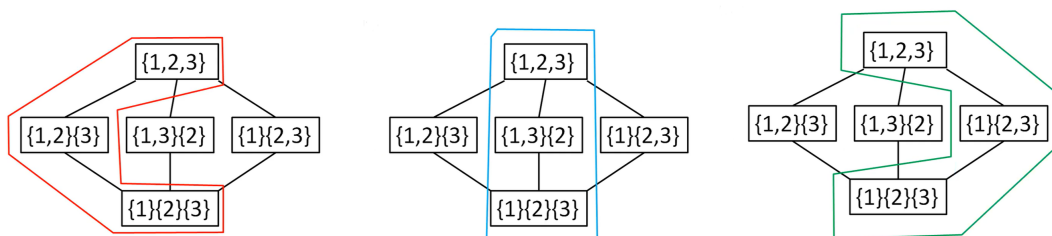
What is the smallest number of chains needed to cover $\sqcap_3$?
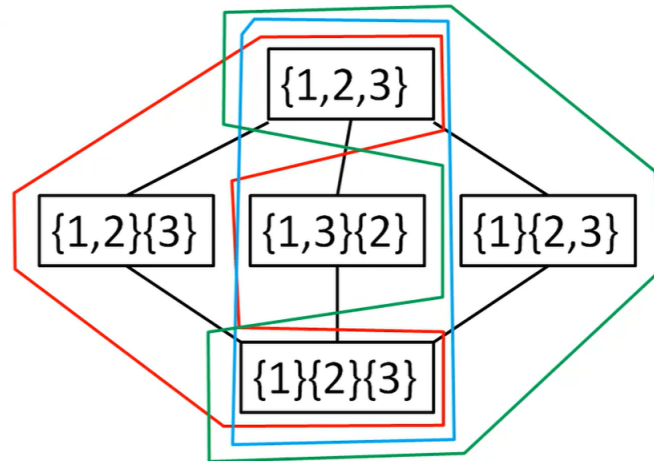

- There are **3** chains. They are shown below:

    - $1|2|3 \to 12|3 \to 123$.
    - $1|2|3 \to 13|2 \to 123$.
    - $1|2|3 \to 1|23 \to 123$.

  Diagrammatically, this looks like:
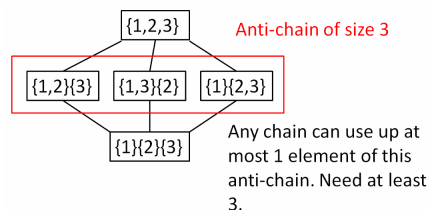

  Combining these together gives:

11.6.3   Dilworth's Theorem

Theorem 11.1: Dilworth's Theorem

The size of any maximum antichain is equal to the number of chains in any smallest chain cover.

With regards to the previous question, Dilworth's Theorem answers the question: *How do we know we cannot use fewer than 3 chains?*

The proof is a bit on the tedious side.

Proof. We know that the minimum chain-cover is greater than or equal to maximum anti-chain. This is because if you have a chain covering an anti-chain, each chain in the cover can intersect at most one element in the anti-chain, and because they have to cover the whole anti-chain, the number of chains in the cover needs to be at least the number of points in the anti-chain.

Now, we need to prove by strong induction on the size of the partial order P. This is significantly more difficult.

◇   <u>Base Case:</u> If $|P| = 1$, the size of the maximum chain is equal to 1, which is equal to the size of the minimum chain.

◇   <u>Inductive Step:</u> Suppose Dilworth's Theorem holds for all partial orders with fewer than $|P|$ elements. Let A be an maximal anti-chain. Then, every other element x in P must be comparable to some element of A. Otherwise, $A \cup \{x\}$ would be a bigger anto-chain.

Each x can be bigger than some element of A or smaller, but not both. For example, if $x > a_1$ and $x < a_2$, then by transitivity we have $a_1 < x < a_2$, which means that these two elements are comparable and thus A is not an anti-chain.

We can partition the elements of P into 3 sets.

    – A

- U, the set of elements bigger than some element in A.
- L, the set of elements less than some elements in A.

This looks like:
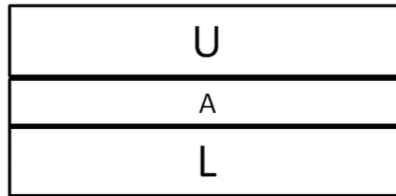
Now, consider these cases.

- <u>Case 1</u>: Suppose that there is some maximal anti-chain A so that both U and L are non-empty. Then, applying the inductive hypothesis to $A \cup L$ and $A \cup U$, we know that A is still a maximal anti-chain for each of these partial orders and, therefore, by the inductive hypothesis, there are chain covers of size $|A|$ for each chain. Now, the only way we can have chain covers of sizes exactly the size of A is if each element in A is in exactly one chain in each cover. So, gluing the upper and lower chain covers together:
  * We can combine the chain from $A \cup U$ with $x \in A$ with the chain in $A \cup L$ with x to get the new chain.
  * These chains cover $L \cup A \cup U = P$.
- <u>Case 2</u>: Suppose that every maximal anti-chain A either has $L = \emptyset$ or $U = \emptyset$. Then, every maximal anti-chain either contains all minimal elements of P or all maximal elements of P. Otherwise, a minimal element will be in L and a maximal element will be in U. Then, suppose we find a maximal element x and a minimal element y with $x \geq y$. Then, every element is bigger than some minimal element. Additionally, from our assumptions, every maximal anti-chain contains either x or y. This means that $P - \{x, y\}$ has a smaller maximum anti-chain. In particular, letting k be the size of the maximum anti-chain of P. Then, $P - \{x, y\}$ has the maximum anti-chain of size at most $k-1$. By the inductive hypothesis, we can cover $P - \{x, y\}$ with the chains $C_1, C_2, \ldots, C_{k-1}$. This can cover P with chains $C_1, C_2, \ldots, C_{k-1}, \{x, y\}$.

This completes our inductive proof.

Hence, this completes this proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

11.7   Incidence Algebra

> **Definition 11.8**
>
> For a partial order P, let[a]:
> $$\text{Int}(P) = \{(x, y) \mid x, y \in P, x \leq y\}$$
> Here, Int means the interval. So, the above means the set of intervals of P.
>
> ---
> [a]The set of all comparable things.

The **incidence algebra** of P is the set of functions:

$$f : \text{Int}(P) \to \mathbb{R}$$

This has a multiplication operation:

$$(f \cdot g)(x, y) := \sum_{x \leq w \leq y} f(x, w)g(w, y)$$

Remark. Similar to recursion for f_m.

11.7.1   Identity

The incidence algebra has an identity:

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

Essentially:

$$(f \cdot \delta)(x, y) = \sum_{x \le w \le y} f(x, w)\delta(w, y) = f(x, y)$$

Remark. The only non-zero term in the sum is when $w = y$.

11.7.2   Zeta

Another useful member of the incidence algebra is the following:

$$\zeta(x, y) = 1 \qquad \forall x \le y$$

Claim. The number of chains of length k from x to y is:

$$\underbrace{(\zeta - \delta) \cdot (\zeta - \delta) \cdot \ldots (\zeta - \delta)(x, y)}_{k-2 \text{ Times}}$$

Proof. Let $D = \zeta - \delta$. Then:

$$D(x, y) = \begin{cases} 1 & x < y \\ 0 & \text{Otherwise} \end{cases}$$

Then:

$$
\begin{aligned}
(D \cdot D \cdot \cdots \cdot D)(x, y) &= \sum_{x \le w_1 \le w_2 \le \cdots \le w_{k-2} \le y} D(x, w_1)D(w_1, w_2) \ldots D(w_{k-3}, w_{k-2})D(w_{k-2}, y) \\
&= \sum_{x < w_1 < w_2 < \cdots < w_{k-2} < y} 1 \\
&= \text{Number of length } k \text{ chains from } x \text{ to } y
\end{aligned}
$$

$\square$

11.7.3   Inverses

We have that δ acts as an identity element for the incidence algebra. What about inverses?

Lemma 11.2

For any function f with $f(x, x)$ being non-zero for all x in P, there is a g so that $(f \cdot g) = \delta$.

11.8   Mobius Function

Of particular interests is the inverse of ζ, commonly denoted by μ.

For $\mu(x, x) = 1$ and for $x < y$, we have:

$$0 = \sum_{x \le w \le y} \zeta(x, w)\mu(w, y) = \sum_{x \le w \le y} \mu(w, y)$$

In particular:

$$\mu(x, y) = - \sum_{x \le w < y} \mu(w, y)$$

11.8.1 Example 1: B_n

For two sets $S \subseteq T$, we have:

$$\mu(S,T) = (-1)^{|T|-|S|}$$

To verify that this works, suppose that $|T| - |S| = k > 0$. There are $\binom{k}{m}$ subsets $S \subseteq R \subseteq T$ with $|T| - |R| = m$ (we take a subset of $T - S$ of size m). Thus:

$$\sum_{S \subseteq R \subseteq T} \mu(R,T) = \sum_{m=0}^{k} \binom{k}{m}(-1)^m = 0$$

11.8.2 Example 2: $\mathbb{N}$

To compute $\mu(n,m)$, write $\frac{m}{n}$ as a product of primes $p_1 p_2 \ldots p_k$. Then:

$$\mu(n,m) = \begin{cases} (-1)^k & \pi \text{ are distinct} \\ 0 & \text{Otherwise} \end{cases}$$

To show that this works, we need to show that the sum of $\mu(d,m)$ over all $n|d|m$ is 0. We note that if $\frac{m}{n}$ is divisible by k distinct primes, there will be $\binom{k}{t}$ such ds where $\frac{m}{d}$ is a product of t distinct primes (choose which primes from the k). The computation, then, is similar to the previous example.

11.8.3 Example 3: Partial Order

An example can be found at the end.

11.9 Mobius Inversion

> **Theorem 11.2**
>
> Given two functions f and g on P with:
>
> $$g(x) = \sum_{x \leq y} f(y)$$
>
> We have that:
>
> $$f(y) = \sum_{y \leq x} g(x)\mu(y,x)$$

11.10 Products of Partial Orders

> **Definition 11.9**
>
> If P and Q are partial orders, the product partial order is defined in the set $P \times Q$ where (x_1, x_2) is $\leq (y_1, y_2)$ if and only if $x_1 \leq y_1$ and $x_2 \leq y_2$.

11.10.1 Example: $\mathbb{R}$ and B_n

$$\mathbb{R}^n = \underbrace{\mathbb{R}^1 \times \mathbb{R}^1 \times \cdots \times \mathbb{R}^1}_{n \text{ Times}}$$

$$B_n = \underbrace{B_1 \times B_1 \times \cdots \times B_1}_{n \text{ Times}}$$

Each copy of B_1 asks if a single element is in the set.

$$\mathbb{N} \text{ is a product over primes.}$$

For each p, we have a partial order on $\{1, p, p^2, p^3, \dots\}$. $\mathbb{N}$ is equivalent to the partial order on the product of these.

11.10.2   Mobius on Products

> **Theorem 11.3**
>
> We have for partial orders P and Q:
>
> $$\mu_{P \times Q}((x, x'), (y, y')) = \mu_P(x, y)\mu_Q(x', y')$$

Proof. It is enough to show that this satisfies the definition.

$$\sum_{(x,x') \leq (w,w') \leq (y,y')} \mu_P(w, y)\mu_Q(w', y') = \left(\sum_{x \leq w \leq y} \mu_P(w, y) \right) \left(\sum_{x' \leq w' \leq y'} \mu_Q(w', y') \right) = \delta((x, x'), (y, y'))$$

$\square$

11.11   Lattices

Another special case:

> **Definition 11.10**
>
> A partial order is a lattice if for every two elements x, y we have:
>
> - A minimum upper bound: $a = x \vee y$
>
> - A maximum lower bound: $b = x \wedge y$.
>
> In particular:
>
> - $z \geq x$ and $z \geq y$ if and only if $z \geq a$.
>
> - $z \leq x$ and $z \leq y$ if and only if $z \leq b$.

From the book:

> **Definition 11.11**
>
> A partial order is called a lattice if any two elements x and y of said partial order have a minimum common upper bound a, and a maximum common lower bound b.
>
> Here, a is called the *join* of x and y, and b is called the *meet* of x and y. We denote these relations by $x \vee y = a$ and $x \wedge y = b$.

11.11.1   Example: Lattices

Which of the following partial orders are lattices?

(a) B_n

(b) $\mathbb{N}$

(c) $\mathbb{R}^n$

- The answer is all of them.

  – B_n: Union/intersection.

  – $\mathbb{N}$: LCM/GCD.

  – $\mathbb{R}^n$: Coordinate-wise max/min.

11.11.2   Minimums and Maximums

> **Lemma 11.3**
>
> Any finite lattice has minimum and maximum elements.

Proof. Take the minimum upper bound and maximum lower bound of all elements. $\square$

11.11.3   Another Criteria

> **Lemma 11.4**
>
> If a finite partial order P has a maximum element and has maximum lower bounds, it is a lattice.

Proof. Need to show that for any x and y, we have a minimum upper bound. Take the maximum lower bound of all upper bounds of x and y (exist because maximum). $\square$

11.11.4   An Application: Refinement

> **Corollary 11.1**
>
> $\sqcap_n$ is a lattice.

Proof. $\sqcap_n$ is finite, has a maximum element ($[n]$), and has maximum lower bounds (common refinement). $\square$

11.12   Weisner's Theorem

> **Theorem 11.4**
>
> Let P be a lattice with maximum element 1 and minimum element 0. Then, for any $a \neq 1$:
> $$\mu(0,1) = -\sum_{\substack{x \wedge a = 0 \\ x \neq 0}} \mu(x,1)$$

Proof. We need to show that:
$$\sum_{x:x \wedge a = 0} \mu(x,1) = 0$$

Define the function f on P by:

$$f(x) := \begin{cases} 1 & x \wedge a = 0 \\ 0 & \text{Otherwise} \end{cases} = \sum_{\substack{y \leq a \\ y \leq x}} \mu(0,y)$$

We have that:

$$\sum_{x:x \wedge a = 0} \mu(x,1) = \sum_x \mu(x,1) f(x) = \sum_x \sum_{\substack{y \leq x \\ y \leq a}} \mu(x,1)\mu(0,y) = \sum_{y \leq a} \mu(0,y) \underbrace{\sum_{x \geq y} \mu(x,1)}_{\text{Always } 0} = 0$$
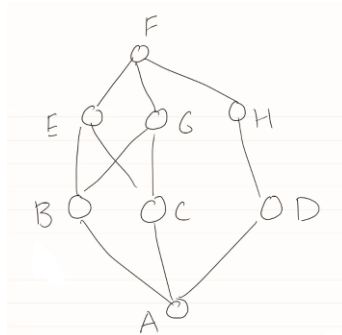
$\square$

11.13   Practice Problem

Here are some random practice problems.

11.13.1   Example: Hasse Diagram

Consider the following Hasse Diagram[2]:

- Is this a lattice?

- Find $\mu(A, F)$.

- What is the largest anti-chain? What is the minimal chain cover?

The solutions are as follows:

- Let's begin by considering two cases.

   - Let's consider the two elements B and C. Recall that an upper-bound for two elements x and y is an element u such that $x \leq u$ and $y \leq u$.

      * First, what elements are above B? We have elements G, E, and F.
      * Now, what elements are above C? We have elements E, G, and F.

      Further, recall that two elements x and y have a minimum common upper-bound if there is **one** element u that is the "closet" to both x and y such that $x \leq u$ and $y \leq u$. While it is true that both B and C share the same upper-bounds, B and C do **not** have a minimum common upper-bound (because there are two different "minimum" upper-bounds).

   - Let's now consider the two elements E and G. Recall that a lower-bound for two elements x and y is an element u such that $u \leq x$ and $u \leq y$.

      * First, what elements are below E? We have elements B, C, and A.
      * Now, what elements are below G? We have elements B, C, and A.

      By similar reasoning, while both E and G share the same lower-bounds, E and G do **not** have a maximum common lower-bound (because there are two different "maximum" lower-bounds).

   By this, it follows that the above Hasse Diagram does not represent a lattice.

- We know that:
$$\mu(A, F) = - \sum_{A \leq Z < F} \mu(A, Z)$$

   We know that (this is given):
$$\mu(A, A) = 1$$

   Consider the elements that connect directly to A. This is simply:
$$\mu(A, B) = \mu(A, C) = \mu(A, D) = -\mu(A, A) = -1$$

---

[2]Thanks to TA Jesse Kim for this drawing.

Now, we know that:
$$\mu(A, H) = -(\mu(A, A) + \mu(A, D)) = -(1 - 1) = 0$$
$$\mu(A, E) = -(\mu(A, A) + \mu(A, B) + \mu(A, C)) = -(1 + -1 + -1) = 1$$
$$\mu(A, G) = -(\mu(A, A) + \mu(A, B) + \mu(A, C)) = -(1 - 1 - 1) = 1$$
$$\mu(A, F) = -(\mu(A, A) + \mu(A, B) + \mu(A, C) + \mu(A, D) + \mu(A, E) + \mu(A, G) + \mu(A, H))$$
$$= -(1 + (-1) + (-1) + (-1) + 1 + 1 + 0)$$
$$= -(3 - 3)$$
$$= 0$$

- `BCD`, `EGH`, `BCH`, and `EGD` are the largest anti-chains. By Dilworth's Theorem, the size of the largest anti-chain is 3, so there are 3 chains in the smallest chain cover. In this case, the 3 chains are `ABEF`, `ACGF`, and `ADHF`.

12   Appendix

Here are some more information that may be helpful.

12.1   Corresponding Lectures

The following are the corresponding lectures. These may be useful if the notes don't cover everything. If you click on a link, make sure to right-click and open in a new tab.

L	Topic	Video	Slides
1	Weak Induction	Link	Link
2	Strong Induction, Pigeonhole	Link	Link
3	Generalized Pigeonhole, Basic Counting Techniques	Link	Link
4	Multinomial Coefficients, Balls and Bins	Link	Link
5	Stars and Bars, Set Partitions, Stirling Numbers	Link	Link
6	Stirling Numbers, Bell Numbers, Integer Partitions	Link	Link
7	Integer Partitions	Link	Link
8	Permutations, Cycle Structures, Counting	Link	Link
9	Counting Permutations, Stirling, Cycles w/ Specific Elements, Odd/Even	Link	Link
10	Sizes of Cycles w/ Individual Elements, Odd/Even Cycles	Link	Link
11	Computation of Even/Odd Cycles, Inclusion-Exclusion	Link	Link
12	Inclusion-Exclusion, Applications	Link	Link
13	Derangements, Binomial Theorem	Link	Link
14	Generalizations of Binomial Theorem, Generating Functions	Link	Link
15	Solving Recurrence Relations, Products of Generating Functions	Link	Link
16	Applications of Generating Functions, Catalan Numbers	Link	Link
17	Composition of Generating Functions	Link	Link
18	Basics of Exponential Generating Functions	Link	Link
19	Compositions of Generating Functions	Link	Link
20	Applications of Compositions of EGF	Link	Link
21	Patterns and Pattern Avoidance in Permutation	Link	Link
22	Computing $|S_n(123)|$, Left-to-Right Minima	Link	Link
23	Pattern Avoidance Asymptotics	Link	Link
24	Partial Orders, Chains, Examples	Link	Link
25	Incidence Algebras, Inverses	Link	Link
26	Computing μ	Link	Link
27	Professor Kane's Research	Link	Link

12.2   Common Generating Functions

- For the sequence of numbers: $1, 1, 1, 1, 1, \ldots$

$$x^0 + x^1 + x^2 + x^3 + x^4 + x^5 + \cdots = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

$$\frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

- For the sequence of numbers: $c^0, c^1, c^2, c^3, c^4, \ldots$

$$(cx)^0 + (cx)^1 + (cx)^2 + (cx)^3 + (cx)^4 + \cdots = \sum_{n=0}^{\infty} (cx)^n = \frac{1}{1-cx}$$

$$\frac{(cx)^0}{0!} + \frac{(cx)^1}{1!} + \frac{(cx)^2}{2!} + \frac{(cx)^3}{3!} + \frac{(cx)^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{(cx)^n}{n!} = e^{cx}$$

- For the sequence of numbers: $1, -1, 1, -1, \ldots$

$$x^0 - x^1 + x^2 - x^3 + x^4 - \cdots = \sum_{n=0}^{\infty}(-1)^n x^n = \frac{1}{1-(-x)} = \frac{1}{1+x}$$

- For the sequence of numbers: $1, 2, 3, 4, \ldots$

$$x^0 + 2x^1 + 3x^2 + 4x^3 + \cdots = \sum_{n=0}^{\infty}\frac{1}{(1-x)^2}$$

- For the sequence of numbers: $1, 0, 1, 0, \ldots$

$$x^0 + x^2 + x^4 + x^6 + \cdots = \sum_{n=0}^{\infty} x^{2n} = \frac{1}{1-x^2}$$

- For the sequence of numbers: $0, 1, 0, -1, 0, 1, \ldots$

$$0 + \frac{x^1}{1!} + 0 - \frac{x^3}{3!} + 0 + \frac{x^5}{5!} + 0 - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{(2n+1)!} = \sin(x)$$

- For the sequence of numbers: $1, 0, -1, 0, 1, 0, \ldots$

$$\frac{x^0}{0!} + 0 - \frac{x^2}{2!} + 0 + \frac{x^4}{4!} + 0 - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} = \cos(x)$$

- For the sequence of numbers: $0, 1, 0, -2, 0, 24, 0, -720, \ldots$

$$0 + \frac{x^1}{1!} + 0 - 2\frac{x^3}{3!} + 0 + 24\frac{x^5}{5!} + 0 - 720\frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{2n+1} = \arctan(x)$$

- For the sequence of numbers: $0, 1, 0, 1, 0, 9, 0, 225, \ldots$

$$0 + \frac{x^1}{1!} + 0 + \frac{x^3}{3!} + 0 + 9\frac{x^5}{5!} + 0 + 225\frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty}\frac{(2n)! x^{2n+1}}{(2^n n!)^2 (2n+1)} = \arcsin(x)$$

- For the sequence of numbers: $0, 1, 1, 2, 6, 24, \ldots$

$$0 + \frac{x^1}{1!} + \frac{x^2}{2!} + 2\frac{x^3}{3!} + 6\frac{x^4}{4!} + 24\frac{x^5}{5!} + \cdots = \sum_{n=0}^{\infty}\frac{x^n}{n} = \log\left(\frac{1}{1-x}\right)$$

- For the sequence of numbers: $0, 1, 0, 1, 0, 1, 0, \ldots$

$$0 + \frac{x^1}{1!} + 0 + \frac{x^3}{3!} + 0 + \frac{x^5}{5!} + 0 + \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!} = \sinh(x)$$

- For the sequence of numbers: $1, 0, 1, 0, 1, 0, 1, \ldots$

$$\frac{x^0}{0!} + 0 + \frac{x^2}{2!} + 0 + \frac{x^4}{4!} + 0 + \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!} = \cosh(x)$$

12.3   Applying Generating Functions

Here are some general applications of generating functions.

12.3.1   Integer Partitions

We will briefly discuss some generating functions for integer partitions.[3]

- The generating function for integer partitions is:

$$\prod_{i=1}^{\infty} \frac{1}{1-x^i} = \frac{1}{(1-x)(1-x^2)(1-x^3)\dots}$$

Where each use of i contributes i to the total size n. We can think of an integer partition as an explicit number of ones, twos, threes, and so on. Hence, $n = a_1 + 2a_2 + 3\dots a_3 + \dots$. Considering that $\frac{1}{1-x^k} = \sum_{i\geq 0} x^{ki} = 1 + x^k + x^{2k} + \dots$, the product of all such generating functions would, as a sum, express the number of ways to partition n into a_1 many ones, a_2 many twos, and so on.

If we expand this out, we get something like:

$$\prod_{i=1}^{\infty} \frac{1}{1-x^i} = 1 + x + 2x^2 + 3x^3 + 5x^4 + 7x^5 + 11x^6 + 15x^7 + 22x^8 + \dots$$

So, for instance, $p(7) = 15$ (there are 15 integer partitions for $n = 7$).

- The generating function for integer partitions where every part is odd is:

$$\prod_{i \text{ odd}}^{\infty} \frac{1}{1-x^i} = \frac{1}{(1-x)(1-x^3)(1-x^5)\dots}$$

If we expand this out, we get something like:

$$\prod_{i \text{ odd}}^{\infty} \frac{1}{1-x^i} = 1 + x + x^2 + 2x^3 + 2x^4 + 3x^5 + 4x^6 + 5x^7 + \dots$$

So, for instance, $p_o(7) = 5$ (there are 5 integer partitions for $n = 7$ such that all parts are odd).

- The generating function for integer partitions with distinct parts is:

$$\prod_{i=1}^{\infty}(1 + x^i)$$

The idea is that, for each i, you choose whether to use i once or not at all.

If we expand this out, we get something like:

$$\prod_{i=1}^{\infty}(1 + x^i) = 1 + x + x^2 + 2x^3 + 2x^4 + 3x^5 + 4x^6 + \dots$$

So, for instance, $q(5) = 3$ (there are 3 integer partitions for $n = 5$ such that all parts are unique).

---

[3]Taken from: https://math.berkeley.edu/ mhaiman/math172-spring10/partitions.pdf

12.3.2 Weak and Strong Compositions

We will now briefly discuss some generating functions for weak and strong compositions.

- Recall that the strong composition (often just denoted as the "composition" or "strict composition") of n into k parts is given by:

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

The corresponding generating function is:

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

For instance, suppose we wanted to find the number of compositions of 6 into 4 parts. The corresponding generating function is:

$$\left(\frac{x}{1-x}\right)^4 = x^4 + 4x^5 + \boxed{10x^6} + 20x^7 + 35x^8 + 56x^9 + \ldots$$

So, the number of compositions of 6 into 4 parts is **10**. Using the binomial coefficient above, we know that:

$$\binom{6-1}{4-1} = 10$$

- Recall that the weak compositions of n into k parts is given by:

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

The corresponding generating function is:

$$\left(\sum_{n=0}^{\infty} x^n\right)^k = \left(\frac{1}{1-x}\right)^k$$