NOTES FOR A6800: COMBINATORIAL ANALYSIS

ZAJJ DAUGHERTY

Contents

Date: December 9, 2015.

1. Introduction to enumerative combinatorics

1.1. Basic counting.

Warmup. Count the following.

(a) The number of possible outcomes of flipping a coin 3 times.
(b) The number of possible outcomes of flipping a coin 10 times.
(c) The number of possible outcomes of flipping a coin 10 times so that the first flip comes up tails, but the last flip doesn't.
(d) The number of possible outcomes of flipping a coin 10 times so that either the first flip comes up tails or the last flip comes up tails, but not both.
(e) The number of ways to pick a president, a vice president, and a secretary for a club with 10 members.
(f) The number of ways to separate four distinct playing cards into two sets of two.

Basic rules of counting. For basic counting, we like to devise a procedure or algorithm for counting the number of ways to do something in an organized fashion.

(1) Product rule. Suppose a procedure for counting can be broken into two steps, where there are the same number of ways to do the second step no matter what you do in the first step. If there are n_1 ways to do the first task, and for each of these ways of doing the first task, there are n_2 ways to doing the second task, then there are $n_1 n_2$ ways in total to do the procedure.
Example: If I flip a coin three times, there are 2 possible outcomes for the first flip; no matter what the first flip, there are 2 possible outcomes for the second flip; and no matter what the first two flips, there are 2 possible outcomes for the third flip. SO there are $2*2*2 = 8$ possible outcomes.
Example: Suppose I flip a coin ten times, but I know that the first flip comes up tails and the last flip doesn't. There is one possibility for the first flip, 2 possibilities for the second flip, 2 for the third, $\ldots$, 2 for the ninth, and 1 for the last (it must be heads). So in total, there are $1 * 2 * 2 * \cdots * 2 * 1 = 2^8$ possible outcomes. **Example:** Suppose there are four people heading to stand in a line. Then there are 4 possible people who could end up at the front; once that person has stood at the front, there are three remaining people to stand second (no matter who is in front); then there are two left to stand third; and one left to stand last. So there are 4*3*2*1 = 4! ways for those people to line up.

Aside: The *permutations* of a set of n distinguishable objects are the set of *linear orders* (or *rearrangements*) of those objects. For example, the permutations of the set $\{A, B, C\}$ are

$$ABC, ACB, BAC, BCA, CAB, \text{ and } CBA.$$

By the product rule, there are $n! = n(n-1)(n-2)\cdots 1$ permutations of n distinguishable objects.

(2) Sum rule. Suppose a procedure for counting can be broken into two disjoint cases. If case 1 can be done n_1 ways and case 2 can be done in n_2 ways, where there is no overlap in the n_1 and n_2 ways, then there are $n_1 + n_2$ ways in total.
Example: Suppose I flip a coin ten times, and I want to know how many outcomes have either the first flip comes up tails or the last flip comes up tails, but not both. In case 1, the first flip comes up tails and the last doesn't, which can happen any of 2^8 ways (as above). In case 2, the last flip comes up tails and the first doesn't, which can happen any of 2^8 ways (product rule again). Since an outcome can't both fall into case 1 and case 2, there is a total of $2^8 + 2^8$ possibilities.

(3) **Complement rule.** Suppose a counting problem has a condition, and you're able to count both the number of outcomes n_1 without the condition, and the number of outcomes n_2 *not* satisfying the condition. Then the number of outcomes satisfying the condition is $n_1 - n_2$.
Example: Suppose I'm picking outfits again from my 4 pants and 5 shirts, but one of my shirts is red and two of my pairs of pants are green, and I'm not willing to wear those colors together. The condition is that the outfit doesn't have both the red shirt and the green pants in it. There are two such outfits. Since there are 20 total outfits, I now have $20 - 2 = 18$ outfits I'm willing to wear.

(4) **Division rule.** Suppose I devise a procedure to count outcomes, but for every actual outcome, there are k results of my procedure that produce that same outcome. Then if there are n procedural outcomes, there are n/k distinct outcomes.
Example: Suppose I have have 4 cards, and I want to pick a pair from them. I could pick the cards one at a time—where there are 4 choices for the first card, and then no matter what, there are 3 choices for the second card. Product rule tells me that this gives $4 * 3 = 12$ outcomes. However, my procedure doesn't quite line up with my desired outcomes. If I pick card A, followed by card B, this is the same as first picking card B and then card A. So for every pair, there were 2 procedural outcomes that produced that pair. Therefore, there are $12/2 = 6$ possible pairs.
Example: Now suppose that I want to just divide my four cards into two piles of 2. I could do this by choosing two card from the 4, of which there are 6 ways. But If I choose A and B, that separates my 4 cards into the same two pairs as if I had *not* chosen A and B. In this way, no matter what pair I pick, there is a corresponding pair that results in the same division. Therefore, there are $6/2 = 3$ possibilities.

1.2. **Set notation and first combinatorial identities.**

A brief note on mathematical writing. Mathematical writing is *writing.* There are nouns, and verbs, and adjectives, and everything else that we are used to language being based upon. The goal of mathematical writing, like most writing, is to convey ideas; the goal is to educate and convince your reader. We write in complete sentences. We mind punctuation. The tricky part, perhaps the art, is that we also use notation to consolidate lots of information into bite-sized packages. If you go back and read Euclid's manuscripts, it can be hard to figure out what he was saying, because he was doing geometry in prose! Notation is great. Notation is useful. But we all have to learn how to balance brevity, clarity, and readability.

Some example of notation that exists purely for brevity:

$$\begin{aligned} \forall &\text{ means “for all”}, \\ \exists &\text{ means “there exists”}, \\ \text{iff} &\text{ means “if and only if”}, \\ \text{s.t.} &\text{ means “such that”}, \end{aligned}$$

and so on. I mostly use this kind of notation either on the board, or *occasionally* in writing when it is useful to package things for visual reasons. But it's delicate. For example, the definition of the *limit* of a function $f(x)$, as x approaches ∞, can be written as,

$$\lim_{x\to\infty} f(x) = L \quad \text{iff} \quad \forall \epsilon > 0 \quad \exists a \quad \text{ s.t.} \quad \forall x > a, \quad |f(x) - L| < \epsilon.$$

Sure, it's brief; it fits on a T-shirt. But it is also terribly opaque and requires a long time to unpack its meaning. It would be less brief, but more readable, to say the following.

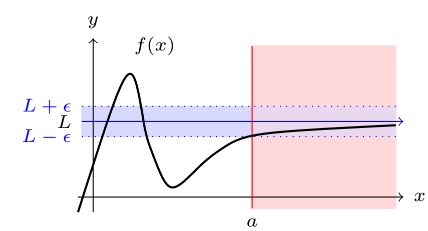
> The function $f(x)$ has *limit* L *as* x *approaches* ∞, written
> $$\lim_{x\to\infty} f(x) = L,$$
> if, for any $\epsilon > 0$, there is some corresponding point a such that so long as $x > a$, we have that the distance of $f(x)$ from L is less than ϵ.

There, "$\lim_{x\to\infty} f(x) = L$" is an abbreviation, but it serves to give the reader something visual to latch on to. Same thing with $\epsilon > 0$, where I could have written "epsilon, a number greater than 0".

When I am deciding whether to use notation, or whether to write things out, I ask myself a few questions.

(1) What will be easiest to read?
(2) What will be the most accurate and concise?
(3) Do I want to use notation because I'm lazy?
(4) Will drawing a picture effect my decisions?

In fact, in this class, we will see that drawing pictures is actually extremely effective in many cases. For example, some version of the following picture often accompanies the above definition of a limit.



Mathematical writing is an art, just like any writing. You learn through reading and through practice.

That all being said, here's some more notation.

Set notation:

$$\{\ \text{elements}\ \underbrace{\ \mid\ }_{\substack{\text{“such that”}\\ \text{“satisfying”}}}\ \text{conditions}\ \}$$

For example, the unit circle is the set of real-valued points

$$\{(x, y) \mid x^2 + y^2 = 1\}.$$

Note that sometimes, as in our book, people use ":" instead of "|".

For sets S and T,

$S \subseteq T$ means that every element of S is also an element of T;

$S \subset T,\ S \subsetneq T$ both mean $S \subseteq T$, but $S \neq T$;

$x \in S$ means that x is an element of S;

$S \cap T$ means the *intersection* of S and T, i.e. the set of elements in both S and T;

$S \cup T$ means the *union* of S and T, i.e. the set of elements in either S or T or both;

$S \sqcup T,\ S \,\dot{\cup}\, T$ both mean the *disjoint union* of S and T, i.e. $S \cup T$ where $S \cap T = \emptyset$;

$S - T,\ S \setminus T$ both mean the set of elements of S that are not in T;

$|S|,\ \#S$ both mean the size of S, i.e. the number of elements in S.

Like a lot of things in this introduction, I expect you to be familiar with most of these already. But it's always good to start on the same footing, and get our conventions straight.

Important sets of numbers:

$$\begin{aligned}
\mathbb{Z} &= \{0, \pm 1, \pm 2, \dots\} && \text{Integers}\\
\mathbb{N} = \mathbb{Z}_{\geq 0} &= \{0, 1, 2, \dots\} && \text{Natural numbers}\\
& && \text{(non-negative integers)}\\
\mathbb{Z}_{>0} &= \{1, 2, \dots\} && \text{Whole numbers}\\
\mathbb{Q} &= \{a/b \mid a, b \in \mathbb{Z}, b \neq 0\} && \text{Rational numbers}\\
\mathbb{R} &= \left\{ \sum_{n=N}^{\infty} \epsilon a_n 10^{-n} \;\middle|\; N \in \mathbb{Z}, a_n \in \{0, 1, \dots, 9\}, \epsilon \in \{-1, 1\} \right\} && \text{Real numbers}\\
\mathbb{C} &= \{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\} && \text{Complex numbers}\\
\emptyset &= \{\} && \text{Empty set}\\
[n] &= \{1, 2, \dots, n\}, \text{ for any } n \in \mathbb{Z}_{>0}; [0] = \emptyset\\
[m, n] &= \{m, m+1, \dots, n-1, n\}, \text{ for any } m \leq n \in \mathbb{Z}
\end{aligned}$$

If S is some finite set, define

$$\begin{aligned}
2^S &= \{A \mid A \subseteq S\}, && \text{the } \textit{power set of } S\text{, and}\\
S^n &= \{(s_1, s_2, \dots, s_n) \mid s_i \in S\}, && n\text{-tuples of elements of } S.
\end{aligned}$$

For example, if $S = [2] = \{1, 2\}$, then

$$\begin{aligned}
2^S &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \qquad \text{and}\\
S^3 &= \{(1,1,1), (1,1,2), (1,2,1), (1,2,2), (2,1,1), (2,1,2), (2,2,1), (2,2,2)\}.
\end{aligned}$$

Exercise 1. Extended warmup.

(a) Write out the following sets explicitly.

(i) $[4]^2$ (ii) $2^{[4]}$ (iii) $S = \{(a, b) \mid a \in [2, 4], b \in [-4, 7]\}$ (iv) $[4]^2 \cap 2^{[4]}$ (v) $[4]^2 \cap S$ (vi) $2^{[3]} \cup 2^{[4]}$

(b) For sets A and B, decide whether the following identities are **true or false**, and why.

(i) $A \cap B = B \cap A$ (ii) $A \cup B = B \cup A$ (iii) $A - B = B - A$ (iv) $|A - B| = |A| - |B|$

(c) Answer the following counting problems, leaving your numerical answer unsimplified.
- (i) A particular kind of shirt comes in two different cuts–male and female, each in three color choices and five sizes. How many different choices are made available?
- (ii) On a 10-question true-or-false quiz, how many different ways can a student fill out the quiz if they answer all of the questions? if they might leave questions blank?
- (iii) How many 3-letter words (not "real" words, just strings of letters) are there?
- (iv) How many 3-letter words are there that have no repeated characters?
- (v) How many 3-letter words are there that have the property that if they start in a vowel then they don't end in a vowel?

Counting the elements of a set. Again, let S be a finite set, now with n elements. Notice that

$$|S^m| = |S|^m,$$

since there are

$$\begin{aligned}&|S| \text{ choices for the first coordinate, times}\\&|S| \text{ choices for the second coordinate, times}\\&|S| \text{ choices for the third coordinate, and so on.}\end{aligned}$$

Maybe this starts to make us feel good about our notation.

Now what about $|2^S|$? As a combinatorial exercise, we will actually construct a different set that we already know how to measure, and show that it is the same size as 2^S. We do this by constructing a *bijection* (a function that is one-to-one and onto) between 2^S and that other set.

Fix an order for the elements of S,

$$S = \{x_1, x_2, \dots, x_n\}.$$

Consider

$$\{0,1\}^n = \{\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \mid \varepsilon_i = 0 \text{ or } 1 \text{ for } i = 1, 2, \dots, n\}.$$

Since $|\{0,1\}| = 2$, we have $|\{0,1\}^n| = 2^n$. Define a map

$$\begin{aligned}\theta : 2^S &\to \{0,1\}^n,\\ A &\mapsto (\varepsilon_1, \dots, \varepsilon_n),\end{aligned}$$

where

$$\varepsilon_i = \begin{cases} 0, & \text{if } x_i \notin A,\\ 1 & \text{if } x_i \in A.\end{cases}$$

For example, suppose $n = 4$. Then

$$\theta(\{x_1, x_4\}) = (1,0,0,1), \quad \theta(\{x_2\}) = (0,1,0,0), \quad \theta(\emptyset) = (0,0,0,0), \quad \text{and} \quad \theta(S) = (1,1,1,1).$$

We call $\theta(A)$ the *characteristic vector* of A.

We would like to show that θ is

(1) *injective*, or *one-to-one*, meaning that every subset $A \subseteq S$ maps to a unique element of $\{0,1\}^n$; and
(2) *surjective*, or *onto*, meaning that every element of $\{0,1\}^n$ is the image of some subset $A \subseteq S$.

By definition, a function is *bijective* if it is both injective and surjective. As a theorem, a function is bijective if and only if it is invertible. In this case, the easiest way to see that θ is bijective actually is to build an *inverse function*, i.e.

$$\varphi : \{0,1\}^n \to 2^S \qquad \text{satisfying} \qquad \varphi(\theta(A)) = A \text{ and } \theta(\varphi(\varepsilon)) = \varepsilon.$$

In other words, you want a map that takes in a series of 0's and 1's, and puts out a subset of S, such that $\theta(S)$ is the series of 0's and 1's you start with. More qualitatively, φ should take in a sequence of "yes, this element is in the set" (represented by a 1) and "no, this element is not in the set" (represented by a 0) , and tell you what the set is. So define,

$$\begin{aligned}\varphi : \{0,1\}^n &\to 2^S, &&\text{by}\\ (\varepsilon_1, \dots, \varepsilon_n) &\mapsto A = \{x_i \in S \mid \varepsilon_i = 1\}.\end{aligned}$$

For example, if $n = 4$ again, then

$$\varphi((1,0,0,1)) = \{x_1, x_4\}, \quad \theta((0,1,0,0)) = \{x_2\}, \quad \theta((0,0,0,0)) = \emptyset, \quad \text{and} \quad \theta((1,1,1,1)) = S.$$

Then

$$\varphi(\theta(A)) = A \text{ for } A \in 2^S, \quad \text{and} \quad \theta(\varphi(\varepsilon)) = \varepsilon \text{ for } \varepsilon \in \{0,1\}^n.$$

Thus

$$\varphi = \theta^{-1}, \qquad \text{the } \textit{inverse} \text{ of } \theta.$$

Formalizing what I described above, for finite sets, we have the following principle of counting.

For finite sets S and T, we have

$$|S| = |T| \qquad \textit{if and only if} \qquad \textit{there exists a bijective map } S \overset{\theta}{\leftrightarrow} T.$$

So since $|\{0,1\}^n| = 2^n$, and there is a bijective function between $\{0,1\}^n$ and 2^S, we have

$$|2^S| = |\{0,1\}^n| = 2^n = 2^{|S|}. \tag{1.1}$$

Again, this makes us feel good about our notation.

A quick note on counting, and combinatorics. What we did just here, measuring the size of the power set of S by comparing it to the size of another measured set, is pretty much what we do in combinatorics. We count the size of sets mostly in one of two ways: (1) we find an organized way of counting the elements in a straightforward way (like we did for $|S^n|$), or (2) we find another set whose elements are easier to count, and then line up the elements of the two sets using a bijective function (like we did for $|2^S|$). We'll also learn how to use tools from other areas of mathematics to set up and manipulate bijections, but at the heart of everything we do is organized counting and bijective functions.

Back to sets. For a set S and an integer $0 \leq k \leq |S|$, define the set $\binom{S}{k}$, read S *choose* k, by

$$\binom{S}{k} = \{A \subseteq S \mid |A| = k\} \subseteq 2^S,$$

i.e. the set of subsets of S of size k. For example, if $S = [3]$, then

$$\begin{aligned} \binom{[3]}{0} &= \{\emptyset\}, \\ \binom{[3]}{1} &= \{\{1\},\{2\},\{3\}\}, \\ \binom{[3]}{2} &= \{\{1,2\},\{1,3\},\{2,3\}\}, \text{ and} \\ \binom{[3]}{3} &= \{\{1,2,3\}\}. \end{aligned}$$

Notice that the size of $\binom{S}{k}$ only depends on k and the size of S, not the elements themselves. So for integers $0 \leq k \leq n$, define the number $\binom{n}{k}$ by

$$\binom{n}{k} = \left|\binom{S}{k}\right| \qquad \text{for any set } S \text{ of size } n.$$

Again, $\binom{n}{k}$ is read n *choose* k. Our goal is to prove that

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)\cdots(2)(1)} = \frac{n!}{(n-k)!k!}. \tag{1.2}$$

Note that the second equality is just the identity

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{(n(n-1)\cdots(n-k+1))\,((n-k)(n-k-1)\cdots(2)(1))}{((n-k)(n-k-1)\cdots(2)(1))} = \frac{n!}{(n-k)!}.$$

Denote

$$(n)_k = \frac{n!}{(n-k)!}. \tag{1.3}$$

To prove (1.2), we are going to give a "combinatorial proof", which means that we are going to count the elements of a set in two ways, one that gives one formula and another that gives a different formula, and conclude that the two formulas must give the same number. However, we can avoid using the division rule by employing the technique of *clearing the denominator*, and prove the equivalent identity

$$\binom{n}{k} k! = (n)_k. \tag{1.4}$$

Proof. We count in two ways the number $N(n,k)$ of ways of choosing a size k subset $A \subseteq S$, and then linearly ordering the elements of A.

Case 1: Show $N(n,k) = \binom{n}{k}k!$.
By definition, there are $\binom{n}{k}$ ways to pick A. Then for each choice of A, since $|A| = k$, there are $k!$ permutations (linear orders) of the elements of A. So $N(n,k) = \binom{n}{k}k!$.

Case 2: Show $N(n,k) = (n)_k$.
On the other hand, we could have chosen the elements of A one-by-one, in order. This way, there are

$$\begin{gathered} n \text{ choices for the first element,} \\ n-1 \text{ choices for the second element,} \\ n-2 \text{ choices for the third element,} \\ \vdots \\ n-(k-1) \text{ choices for the } k\text{th and last element.} \end{gathered}$$

So

$$N(n,k) = n(n-1)(n-2)\cdots(n-(k-1)) = (n)_k.$$

Therefore, we can conclude

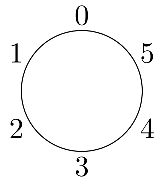
$$\binom{n}{k}k! = N(n,k) = (n)_k,$$

thus proving (1.4). Since $0! = 1$, $k!$ is never 0, so (1.4) is equivalent to (1.2). □

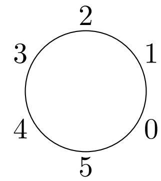
Exercise 2 (EC 1.2)**.** Give as simple a solution as possible. Justify your answers (using words).

(a) How many subsets of the set $[10] = \{1, 2, \ldots, 10\}$ contain at least one odd integer?

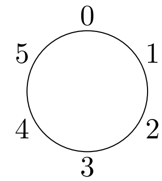
(b) In how many ways can six people be seated in a circle if two arrangements are considered the same whenever each person has the same neighbors (not necessarily on the same side)? For example,



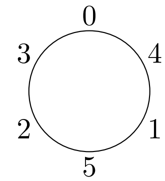
is the same as



and



but not



.

(c) A *permutation* of a finite set S is a bijective map $w : S \to [n]$, where $n = |S|$.
  (i) How many permutations $w : [6] \to [6]$ are there?
  (ii) How many permutations $w : [6] \to [6]$ satisfy $w(1) \neq 2$?

A *cycle* of a permutation is a sequence $(c_1, c_2, \ldots, c_\ell)$ such that

$$w : c_1 \mapsto c_2, \quad w : c_2 \mapsto c_3, \quad \ldots \quad w : c_{\ell-1} \mapsto c_\ell, \quad w : c_\ell \mapsto c_1.$$

For example, the permutation $w : [4] \to [4]$ given by

$$1 \mapsto 4, \quad 2 \mapsto 2, \quad 3 \mapsto 1, \quad 4 \mapsto 3$$

has exactly two cycles: $(1, 4, 3)$ and (2).
  (iii) How many permutations $w : [6] \to [6]$ have exactly one cycle? (Hint: this is *almost* the same question as (b).)
  (iv) How many permutations $w : [6] \to [6]$ have exactly two cycles of length 3?

(d) There are four people who want to sit down, and six distinct chairs in which to do so. In how many ways can this be done?

Exercise 3. (a) Explain why

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

directly using product and division rules.

(b) Give a combinatorial proof of the identity

$$\sum_{i=0}^{n} \binom{a}{i}\binom{b}{n-i} = \binom{a+b}{n},$$

where $a, b, n \in \mathbb{N}$ and $a, b \geq n$. [Hint: Consider two disjoint sets A and B, with $|A| = a$ and $|B| = b$. How many subsets does $A \sqcup B$ have?]

1.3. **Good combinatorial answers.** Recall the Taylor series expansions

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots = \sum_{n=0}^{\infty} x^n, \qquad \text{and} \qquad e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Warmup. Use substitution (function composition), differentiation, and integration to give the series expansions for the following functions, for $c \in \mathbb{R}$, $c \neq 0$.

$$\frac{1}{1-cx}, \qquad \frac{1}{1-x^c}, \qquad e^{cx}, \qquad \frac{1}{(1-x)^2}, \qquad \frac{1}{(1-x)^3}, \qquad \ln|1-x|.$$

[For example, $\frac{d}{dx}\frac{1}{1-x} = -(1-x)^{-2}(-1) = \frac{1}{(1-x)^2}$, so the series expansion of $\frac{1}{(1-x)^2}$ is the derivative of the series expansion of $\frac{1}{1-x}$]

"The basic problem of enumerative combinatorics is that of counting the number of elements of a finite set. Usually we are given an infinite collection of finite sets S_n where n ranges over some index set I (such as the nonnegative integers $\mathbb{N}$), and we wish to count the number $f(n)$ of elements in each S_n *simultaneously*."

Example 1. Some examples that we've seen already include the following.

(1) If $S_n = [n]$, then $f(n) = n$.
(2) If $S_n = 2^{[n]} = \{$ subsets of $[n]$ $\}$, then $f(n) = 2^n$.
(3) If $S_n = \{$ permutations of $[n]$ $\}$, then $f(n) = n!$.

For all of these, i ranges over the indexing set $I = \mathbb{N}$.

But what passes for a satisfactory answer? Sometimes we get nice closed expressions for $f(n)$. But sometimes finding such a closed expression is much more difficult, and we settle for more open-ended "determinations" of $f(n)$. The most common answers we settle for fall into one of the following categories.

(1) Closed formulas.
(2) Recurrence relations.
(3) Algorithms for computing $f(i)$.
(4) An *asymptotic formula* for $f(i)$ (an estimate for large i).
(5) A *generating function* for $f(i)$ (a series whose coefficients store $f(i)$).

Examples of the main classes of combinatorial answers include the following.
Closed formulas. See example 1.

Recurrence relations. Let $f(n)$ be the number of subsets of $[n]$ that do not contain two consecutive integers, For example, for $n = 4$, we have the subsets

$$\emptyset, \quad \{1\}, \quad \{2\}, \quad \{3\}, \quad \{4\}, \quad \{1,3\}, \quad \{1,4\}, \quad \text{and} \quad \{2,4\}.$$

So $f(4) = 8$. Note that if I wanted to list the corresponding subsets of [5], they would each either have a 5 in it or not:

$$\text{without: } \emptyset, \quad \{1\}, \quad \{2\}, \quad \{3\}, \quad \{4\}, \quad \{1,3\}, \quad \{1,4\}, \quad \text{and} \quad \{2,4\};$$
$$\text{with: } \{5\}, \quad \{1,5\}, \quad \{2,5\}, \quad \{3,5\}, \quad \text{and} \quad \{1,3,5\}.$$

If the subset has a 5, then removing it gives you a subset of [3] with no consecutive integers (and you get every such subset by taking a 5 away from some "good" subset of [5]). If the subset doesn't have a 5, then it's actually a "good" subset of [4] (and every such subset appears). So $f(5) = f(4) + f(3)$. In general, by partitioning the set of "good" subsets of $[n]$ into those with n or without n, we can see that $f(n) = f(n-1) + f(n-2)$. So by calculating $f(0) = 1$ and $f(1) = 2$, we can *recursively* calculate

$$\begin{aligned} f(2) &= 1 + 2 = 3, \\ f(3) &= 2 + 3 = 5, \ \checkmark \\ f(4) &= 3 + 5 = 8, \ \checkmark \\ f(5) &= 5 + 8 = 13, \end{aligned}$$

and so on.

[Note that we do have a closed formula for $f(n)$, given by

$$f(n) = \frac{1}{\sqrt{5}}(\tau^{n+2} - \bar{\tau}^{n+2}), \qquad \text{where } \tau = \frac{1}{2}(1+\sqrt{5}), \bar{\tau} = \frac{1}{2}(1+\sqrt{5}).$$

However, it is not a very satisfying combinatorial answer since it is built from irrational numbers.]

Algorithms. In order to give the recurrence relation above, I had to give an algorithm first. As a note on style, a satisfying algorithmic should take less time to run than calculating $f(n)$ by brute force.

Asymptotic formulas. When you can't find an explicit formula for $f(n)$, you might be able to find another function $g(n)$ that satisfies

$$\lim_{n\to\infty} f(n)/g(n) = 1, \qquad \text{written } f(n) \sim g(n),$$

i.e. $f(n)$ and $g(n)$ get closer and closer as n grows larger.

For example, the number of $n \times n$ matrices with 1's and 0's such that every row and column has exactly three 1's is given by

$$f(n) = 6^{-n}(n!)^2 \sum_{\substack{\alpha,\beta,\gamma\in\mathbb{N} \\ \alpha+\beta+\gamma=n}} \frac{(-1)^{\beta})(\beta+3\gamma)!2^{\alpha}3^{\beta}}{\alpha!\beta!(\gamma!)^2 6^{\gamma}}$$

This is a closed formula, but it's terrible. Try checking that $f(3) = 1$, let alone calculating $f(30)$.

On the other hand, one can also show that

$$f(n) \sim e^{-2}36^{-n}(3n)!.$$

Generally, you don't want to use these sorts of estimates for small values of n; but they're great for estimating $f(n)$ for large values of n.

Generating functions. A *generating function* is a formal power series (formal meaning that we're not going to think too hard about convergence) whose coefficients are either $f(n)$; an *exponential generating function* has coefficients $f(n)/n!$:

$$\text{Generating function} \qquad F(x) = \sum_{n\in I} f(n)x^n;$$

$$\text{Exponential generating function} \qquad G(x) = \sum_{n\in I} f(n)\frac{x^n}{n!}.$$

For example, the generating function for the sequence $f(n) = 1$, $n \in \mathbb{N}$, is

$$(1)x^0 + (1)x^1 + (1)x^2 + \cdots = \sum_{n\in\mathbb{N}} x^n = \frac{1}{1-x};$$

the exponential generating function for the same sequence is

$$(1)\frac{x^0}{0!} + (1)\frac{x^1}{1!} + (1)\frac{x^2}{2!} + \cdots = \sum_{n\in\mathbb{N}} \frac{x^n}{n!} = e^x.$$

Note that in this case, we were able to give a nice closed formula for the formal series. This is the best case situation!

As another example, if S_n is the set of permutations of $[n]$, then $f(n) = n!$. So the generating function for $f(n)$ is $F(x) = \sum_{n\in\mathbb{N}} n!x^n$. Since this only converges at $x = 0$ (since factorials grow faster than exponentials for large n), there is no closed formula for $F(x)$. However, the corresponding exponential generating function is

$$G(x) = \sum_{n\in\mathbb{N}} n!\frac{x^n}{n!} = \sum_{n\in\mathbb{N}} x^n = \frac{1}{1-x}.$$

2. Generating functions

Again, for a sequence $f(n)$, $n \in I$, the corresponding

$$\text{the corresponding } \textit{generating function} \text{ is given by} \quad F(x) = \sum_{n\in I} f(n)x^n, \text{ and}$$

$$\text{the corresponding } \textit{exponential generating function} \text{ is given by} \quad G(x) = \sum_{n\in I} f(n)\frac{x^n}{n!}.$$

These power series are called *formal* because we are mostly not concerned with questions of convergence and divergence; the term x^n or $x^n/n!$ just marks the place where $f(n)$ is written.

If $F(x) = \sum_{n\in I} a_n x^n$, we write

$$a_n = [x^n]F(x);$$

similarly, if $G(x) = \sum_{n\in I} a_n x^n/n!$, write

$$a_n = n![x^n]G(x).$$

One of the strengths of storing our counting sequences in generating functions is that we can use basic series operations to combine generating functions and prove other interesting identities (as we'll see in a bit). For example, we can add series, as

$$\sum_{n\in I} a_n x^n + \sum_{n\in I} b_n x^n = \sum_{n\in I}(a_n + b_n)x^n,$$

and

$$\sum_{n\in I} a_n x^n/n! + \sum_{n\in I} b_n x^n/n! = \sum_{n\in I}(a_n + b_n)x^n/n!.$$

Similarly, we can multiply series as

$$\left(\sum_{n\in I} a_n x^n\right)\left(\sum_{n\in I} b_n x^n\right) = \sum_{n\in I} c_n x^n, \qquad \text{where } c_n = \sum_{\substack{i\in I\\ n-i\in I}} a_i b_{n-i}, \tag{2.1}$$

and

$$\left(\sum_{n\in I} a_n x^n/n!\right)\left(\sum_{n\in I} b_n x^n/n!\right) = \sum_{n\in I} c_n x^n/n!, \qquad \text{where } c_n = \sum_{\substack{i\in I\\ n-i\in I}}^{n} \binom{n}{i} a_i b_{n-i}. \tag{2.2}$$

[*Aside:* For those of you who have studied ring theory, these operations make the set of formal power series with coefficients in K (where K is a field, like $\mathbb{C}$), denoted $K[[x]]$, into a very special kind of ring. It's a principal ideal domain, where every ideal is of the form x^n for some n. So in particular, it has the property of unique factorization. Moreover, the map

$$[x^n] : K[[x]] \to K, \quad F(x) \mapsto [x^n]F(x)$$

is linear. The power series in several variables, $K[[x_1, x_2, \ldots, x_m]]$, is also a unique factorization domain, but not generally a principal ideal domain.]

Let

$$\mathbb{C}[[x]] = \left\{ \sum_{n\in\mathbb{N}} a_n x^n \;\middle|\; a_n \in \mathbb{C} \right\}$$

be the set of formal power series with coefficients in $\mathbb{C}$. If $F(x), G(x) \in \mathbb{C}[[x]]$ satisfy

$$F(x)G(x) = 1, \qquad \text{then write } G(x) = F(x)^{-1}$$

(not to be confused with $F^{-1}(x)$, the inverse under function composition). Such a multiplicative inverse exists if and only if $a_0 = [1]F(x) = F(0) \neq 0$. You can see this by solving the equations

$$1 = c_0 = a_0 b_0 \qquad \text{and} \qquad 0 = c_n = \sum_{i=0}^{n} a_i b_{n-i} \text{ for } n > 0,$$

for b_n, from comparing coefficients of (2.1). Here, multiplicative inverses satisfy all the properties we would like. For example,

$$\begin{aligned}
&\text{if } H(x) = F(x)G(x), \text{ then } G(x) = H(x)F(x)^{-1} = H(x)/F(x);\\
&(F(x)G(x))^{-1} = G(x)^{-1}F(x)^{-1};\\
&(F(x)^{-1})^{-1} = F(x); \text{ etc..}
\end{aligned}$$

Exercise 4.

(a) Give both the generating and exponential generating functions for

$$f(n) = 3^n; \qquad g(n) = 3; \qquad h(n) = 3n; \quad \text{and} \quad k(n) = n!3^n; \qquad \text{for } n \in \mathbb{N}.$$

For each, give your answer in series form. Whenever possible, also give your answer in closed form.

(b) Verify the rule for multiplying basic and exponential formal series for $I = \mathbb{N}$, for the first 4 coefficients. In other words, calculate c_n for $n = 0, 1, 2, 3$ by multiplying out the left hand side of

$$(a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots)(b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots) = c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots,$$

and comparing coefficients (and similarly for the exponential case).

(c) Verify that

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

by solving for b_n in the equation

$$\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} c_n x^n,$$

where

$$a_0 = -a_1 = c_0 = 1, \text{ and } a_n, c_n = 0 \text{ otherwise},$$

i.e., $\sum_{n=0}^{\infty} a_n x^n = 1 + x$ and $\sum_{n=0}^{\infty} c_n x^n = 1$. [See EC1, Example 1.1.5; but be more explicit.]

Using generating functions to obtain combinatorial results. So far, we have a combinatorial and a basic counting approaches for studying the binomial coefficients. Now let's use generating functions. Let $x_1, x_2, \ldots, x_n$ be independent variables (they satisfy no relations, aside from commutation).

Lemma 2.1 (EC1, (1.17))**.** *For $S = \{x_1, x_2, \ldots, x_n\}$, we have*

$$(1 + x_1)(1 + x_2) \cdots (1 + x_n) = \sum_{T \subseteq S} \prod_{x_i \in T} x_i. \tag{2.3}$$

Proof. We show (2.3) by induction on n. For $n = 0$, we have $S = \emptyset$, so that

$$\prod_{x_i \in \emptyset} (1 + x_i) = 1 \sum_{T \subseteq \emptyset} \prod_{x_i \in T} x_i.$$

Now assuming (2.3) for a fixed n, we show the same identity for $S' = S \sqcup \{x_{n+1}\}$. To this end, we have

$$\begin{aligned}\prod_{x_i \in S'} (1 + x_i) &= \left(\prod_{x_i \in S} (1 + x_i) \right) (1 + x_{n+1}) \overset{\text{IHOP}}{=} \left(\sum_{T \subseteq S} \prod_{x_i \in T} x_i \right) (1 + x_{n+1}) \\ &= \sum_{T \subseteq S} \prod_{x_i \in T} x_i + x_{n+1} \sum_{T \subseteq S} \prod_{x_i \in T} x_i = \sum_{\substack{T \subseteq S' \\ x_{n+1} \notin T}} \prod_{x_i \in T} x_i + \sum_{\substack{T \subseteq S' \\ x_{n+1} \in T}} \prod_{x_i \in T} x_i = \sum_{T \subseteq S'} \prod_{x_i \in T} x_i.\end{aligned}$$

□

Now, evaluating $x_i = x$ for $i = 1, \ldots, n$ in (2.3) gives

$$(1 + x)^n = \sum_{T \subseteq S} \prod_{x_i \in T} x = \sum_{T \subseteq S} x^{|T|} = \sum_{k=0}^{n} \binom{n}{k} x^k, \tag{2.4}$$

since $|S| = n$, and the term x^k appears exactly $\left|\binom{S}{n}\right|$ times in the sum $\sum_{T \subseteq S} x^{|T|}$ by definition. Equation (2.4) is known as the *Binomial Theorem*. An alternative but equivalent version is

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}, \tag{2.5}$$

which we can surmise by setting $x = a/b$ as follows. If $a = b = 0$, the identity is trivial. Otherwise, assume without loss of generality that $b \neq 0$ (if $b = 0$, set $x = b/a$). Then

$$(a + b)^n = b^n((a/b) + 1)^n = b^n \sum_{k=0}^{n} \binom{n}{k} (a/b)^k = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

The third equality in (2.4) is an instance of the argument that if $\mathcal{S}$ is a collection of finite sets such that $\mathcal{S}$ contains exactly $f(n)$ sets of size n, then

$$\sum_{S \in f S} x^{|S|} = \sum_{n \in \mathbb{N}} f(n) x^n.$$

More generally, if $g : \mathbb{N} \to \mathbb{C}$, then

$$\sum_{S \in \mathcal{S}} g(|S|) x^{|S|} = \sum_{n \in \mathbb{N}} g(n) f(n) x^n.$$

Note that you can generate lots of identities with the binomial theorem by evaluating x at specific values. For example, for $x = -1/2$, we have

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

For $n = 3$, this is

$$\begin{aligned}\frac{1}{8} = \left(\frac{1}{2}\right)^3 &= \binom{0}{3}\left(-\frac{1}{2}\right)^0 + \binom{1}{3}\left(-\frac{1}{2}\right)^1 + \binom{2}{3}\left(-\frac{1}{2}\right)^2 + \binom{3}{3}\left(-\frac{1}{2}\right)^3 \\ &= 1 - 3/2 + 3/4 - 1/8 \quad \checkmark.\end{aligned}$$

Exercise 5. For each of the following identities,

(i) check by hand for $n = 3$;
(ii) verify using the binomial theorem, evaluating for specific values of x;
(iii) give a combinatorial proof of the identity.

(a) $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$;

(b) $\displaystyle\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$ for $n > 0$
[Hint: for (ii), differentiate first].

(c) $\displaystyle\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$ for $n > 0$
[Hint: For (iii), rewrite the identity by moving the negative terms to the right (generally good practice for combinatorial proofs). Construct a bijection between the set E_n of all subsets of $[n]$ that have an even number of elements and O_n, the odd counterpart.].

Exercise 6. Use the multiplication rules for exponential series to show that

$$1 = e^x e^{-x} \qquad \text{implies} \qquad \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0 \quad \text{for } n > 0$$

(our third proof, making EC1, Example 1.1.6 more explicit).

2.1. **A bit of technical stuff about the difference between Taylor series and formal power series.** So far, we've made assertions like

$$\sum_{n\geq 0} c^n x^n = \frac{1}{1-cx} \qquad \text{by substitution.}$$

But be careful with substitutions. For example, we can say that

$$\sum_{n\in\mathbb{N}} \frac{(x+1)^n}{n!} = e^{x+1} = ee^x = e\sum_{n\in\mathbb{N}} \frac{x^n}{n!} \tag{2.6}$$

by substitution. But this is actually a statement made about converging functions, not about formal power series. The series $\sum_{n\in\mathbb{N}} \frac{(x+1)^n}{n!}$ is not a formal power series in x until we simplify is and collect like powers of x, for which we have no formal process (which I'll explain below). Remember a *formal* power series is any element of $K[[x]]$; to make substitutions, we need to introduce the notion of convergence.

In calculus, you probably saw a lot about *pointwise* convergence, i.e. we say

$$\sum_{n\geq 0} \frac{1}{1-x} = \sum_{n\geq 0} x^n, \qquad \text{for } |x| < 0,$$

since $\sum_{n\geq 0} x^n$ converges to $1/(1-x)$ for any *point* x strictly between -1 and 1. However, we really aren't interested in evaluating infinite series like functions. We have to think about the left hand side of (2.6) as a sequence of series (polynomials; *partial sums*) given by

$$F_i = \sum_{n=0}^{i} \frac{(x+1)^n}{n!}, \text{ for } i = 0, 1, \ldots,$$

and ask if that sequence converges to some formal power series.

To this end, we introduce a notion of convergence for series that is not pointwise. Let $F_i(x) \in \mathbb{C}[[x]]$, for $i \in I$ (for simplicity, let I be an infinite subset of $\mathbb{N}$) be a sequence of formal power series. We say the sequence $(F_i(x))_{i\in I}$ *converges* to the formal power series $F(x) = \sum_{n\in\mathbb{N}} a_n x^n \in \mathbb{C}[[x]]$ as $i \to \infty$, written

$$\lim_{i\to\infty} F_i(x) = F(x),$$

if for all $n \in \mathbb{N}$, there is a $\delta(n) \in \mathbb{N}$ for which

$$[x^n]F_i(x) = a_n \qquad \text{for all } i \geq \delta(n).$$

Namely, for each n, the coefficient of x^n eventually stabilizes to the constant a_n.

Example: Consider the sequence

$$F_i(x) = \sum_{n\in\mathbb{N}} \min(i,n) x^n \text{ for } i = 0, 1, 2, \ldots.$$

For example,

$$\begin{aligned} F_0(x) &= 0 \\ F_1(x) &= 0 + x + x^2 + x^3 + \cdots, \\ F_2(x) &= 0 + x + 2x^2 + 2x^3 + \cdots, \\ F_3(x) &= 0 + x + 2x^2 + 3x^3 + \cdots, \end{aligned}$$

and so on. The sequence $(F_i(x))_{i\in I}$ converges to

$$\lim_{i\to\infty} F_i(x) = \sum_{n\in\mathbb{N}} n x^n, \qquad \text{where } \delta(n) = n.$$

Non-example: Consider the sequence

$$F_i(x) = \sum_{n\in\mathbb{N}} \left(\frac{1}{2}\right)^i x^n.$$

You might think that since each coefficient $\left(\frac{1}{2}\right)^i \to 0$ as $i \to \infty$, the sequence $(F_i(x))_{i\in\mathbb{N}}$ might converge to the series 0. However, since the sequence $\left(\frac{1}{2}\right)^i$ doesn't *stabilize* (which is different from converging), the series $(F_i(x))_{i\in\mathbb{N}}$ does not formally converge.

The main difference between the function-theoretic definition of convergence and the formal power series definition of convergence is that the formal definition means that in the limit, one can compute the coefficient of any x^n for a fixed n in *finite* time. Combinatorics is discrete math; we like finite conditions.

For an equivalent definition, define the *degree* of a series $F(x)$, denoted $\deg(F(x))$, as the least integer n for which $[x^n]F(x) \neq 0$ (the opposite notion from degrees of polynomials). For example, the degree of

$$x^5e^x = x^5 + x^6 + \frac{x^7}{2!} + \frac{x^8}{3!} + \cdots \qquad \text{is } 5.$$

Although this is the opposite from the notion of the degree of a polynomial, it satisfies the same sort of identities, like

$$\deg(F(x)G(x)) = \deg(F(x)) + \deg(G(x)).$$

Then the sequence $(F_i(x))_{i\in I}$ converges to $F(x)$ if and only if

$$\lim_{i\to\infty} \deg(F(x) - F_i(x)) = \infty.$$

Namely, for any $n \in \mathbb{N}$, there is some $\delta(n) \in \mathbb{N}$ for which $\deg(F(x) - F_i(x)) > n$ for all $i \geq \delta$, since the first N terms of $F_i(x)$ and $F(x)$ line up. In fact, we can say that $(F_i(x))_{i\in I}$ converges without having to know what it converges to; we have $(F_i(x))_{i\in I}$ converges if and only if

$$\lim_{i\to\infty} \deg(F_{i+1}(x) - F_i(x)) = \infty.$$

Now, taking it one step further, let $(G_i(x))_{i\in I}$ be a sequence of series, and consider

$$F_j(x) = \sum_{i\in I_{\leq j}} G_i(x),$$

which is also a sequence of functions. Then we say

$$\sum_{i\in I} G_i(x) = F(x) \qquad \text{if and only if} \qquad \lim_{j\to\infty} F_j(x) = F(x).$$

Similarly,

$$\prod_{i\in I} G_i(x) = F(x) \qquad \text{if and only if} \qquad \lim_{j\to\infty}\left(\prod_{i\in I_{\leq j}} G_i\right) = F(x),$$

providing $G_i(0) = 1$ for all $i \in I$.

Let's look back at (2.6). Note that in this example,

$$G_i(x) = \frac{(x+1)^i}{i!} = \frac{1}{i!}\sum_{\ell=0}^{i}\binom{i}{\ell}x^\ell = \frac{1}{i!}(1 + ix + \binom{i}{2}x^2 + \cdots + x^i), \text{ for } i \in \mathbb{N}.$$

But then, for example, the constant term of $F_j = \sum_{i=1}^{j} G_i$ is

$$[0]F_j(x) = \sum_{i=0}^{j}\frac{1}{i!} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \cdots + \frac{1}{j!}.$$

Though this coefficient converges to e as $j \to \infty$, it never stabilized. So function-theoretically, (2.6) holds; however, formally, (2.6) does not. Again, one cannot compute the coefficient of x^n for any n in *finitely* many steps. In a formally convergent example, we would have

$$[x^n]\sum_{i\geq 0} G_i(x) = [x^n]\sum_{i=0}^{\delta(n)} G_i(x),$$

which is a finite sum.

To this end, we have the following propositions.

Proposition 2.2 (EC1, Prop. 1.1.8 and 1.1.9). *Let $(G_i(x))_i$ be a sequence of formal power series in $\mathbb{C}[[x]]$.*

(1) The series

$$\sum_i G_i(x) \quad \text{converges if and only if} \quad \lim_{i\to\infty} \deg(G_i(x)) = \infty.$$

(2) If $G_i(0) = 1$ for all i, then the series

$$\prod_i G_i(x) \quad \text{converges if and only if} \quad \lim_{i\to\infty} \deg(G_i(x) - 1) = \infty.$$

The whole point of all of this is to determine when I'm allowed to compose series. For example, if $F(x) = \sum_{n\in\mathbb{N}} a_n x^n$, then $F(3x)$ seems to be ok formally, but $F(x+1)$ does not. Formally, define the composition $F(G(x))$ for $G(x) \in \mathbb{C}[[x]]$ by

$$\sum_{n\in\mathbb{N}} a_n (G(x))^n.$$

Then since

$$\deg((G(x))^n) = n \deg(G(x)) = \begin{cases} 0 & \text{if } \deg(G(x)) = 0, \\ d \geq n & \text{if } \deg(G(x)) \neq 0, \text{i.e. } G(0) = 0, \end{cases}$$

Proposition 2.2 tells us that if $F(x)$ has infinitely many non-zero terms, then

$$F(G(x)) \text{ is well-defined if and only if } G(0) = 0.$$

Exercise 7. Use $\frac{1}{1-x} = \sum_{n\in\mathbb{N}} x^n$ and $e^x = \sum_{n\in\mathbb{N}} \frac{x^n}{n!}$ as *definitions* of $\frac{1}{1-x}$ and e^x, i.e.

$$\frac{1}{1-e^x} \text{ is short-hand for } F(G(x)), \text{ where } \begin{array}{l} F(x) = \sum_{n\in\mathbb{N}} x^n, \text{ and} \\ G(x) = \sum_{n\in\mathbb{N}} \frac{x^n}{n!}. \end{array}$$

Which of the following expressions are well-defined formal power series? Why? For those expressions that are well-defined, give their first few terms.

$$\text{(i) } e^{x+1} \qquad \text{(ii) } e^{x+3x^2} \qquad \text{(iii) } e^{e^x} \qquad \text{(iv) } e^{e^x-1} \qquad \text{(v) } \frac{1}{1-xe^x} \qquad \text{(vi) } \frac{1}{xe^x}$$

Example 2. Let $F(n) \in \mathbb{C}[[x]]$ satisfy $F(0) = 0$ (i.e. $[1]F(n) = 0.$). For any $\lambda \in \mathbb{C}$, define

$$\binom{\lambda}{k} = (\lambda)_k/k! = \lambda(\lambda-1)(\lambda-2)\cdots(\lambda-(k-1))/k!$$

and

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

For example, you can use calculus to calculate the Taylor series for $\sqrt{1+x}$ to get

$$\begin{aligned}
\sqrt{1+x} &= 1 + \frac{1}{2}x - \frac{1}{4}\frac{x^2}{2!} + \frac{3}{8}\frac{x^3}{3!} - \frac{5*3*1}{2^4}\frac{x^4}{4!} + \cdots \\
&= 1 + \frac{1}{2}x + \frac{1}{2}\left(\frac{1}{2}-1\right)\frac{x^2}{2!} + \frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\frac{x^3}{3!} + \frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\left(\frac{1}{2}-3\right)\frac{x^4}{4!} + \cdots \\
&= \sum_{k=0}^{\infty} \binom{1/2}{k} x^k.
\end{aligned}$$

Note that this is a direct consequence of how derivatives work!

We can actually think of λ as an indeterminate, and define $(1+F(x))^\lambda$ as an element of $\mathbb{C}[[x,\lambda]]$. It satisfies the expected product and sum rules with $(1+F(x))^\mu$ in $\mathbb{C}[[x,\lambda,\mu]]$.

As a final formal operation, we can also define derivatives formally on $\mathbb{C}[[x]]$ by

$$F'(x) = \sum_{n=0}^{\infty} a_n n x^n \qquad \text{for } F(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Fortunately, this formal operation satisfies all the differentiation rules we want it to:

$$(F+G)' = F' + G', \quad (FG)' = F'G + FG', \quad (F(G(x)))' = G'(x)F'(G(x)).$$

Note that we don't have to worry about composition in chain rule, since it's still $G(x)$ on the inside, and the well-defined question about composition only depends on the inside function.

If we want, we can also solve differential equations as before, like

$$g' = f'/f, g(0) = g, f(0) = 1 \qquad \text{has solution} \qquad f(x) = e^{g(z)}.$$

The same thing goes formally for formal power series $F(x)$, $G(x)$, $F(0) = 1$, $G(0) = 0$, $G = F'/F$. (See EC1, Example 1.1.11).

Exercise 8. (EC1, exercise 1.8)

(a) Use the generalized binomial theorem to expand $\frac{1}{\sqrt{1-4x}}$ in series form.

(b) Calculate $\frac{(2n)!}{n!}$ for $n = 1, 2, 3$. What is $\frac{(2n)!}{n!}$ in general?

(c) Calculate $\binom{-1/2}{k}$ for $k = 1, 2, 3$ What is $\binom{-1/2}{k}$ in general? [Note that you can factor $\frac{1}{2}$ from every term in the numerator. Then use part (b).]

(d) Conclude

$$\frac{1}{\sqrt{1-4x}} = \sum_{n=0}^{\infty} \binom{2n}{n} x^n.$$

(e) Give a combinatorial proof of the identity $2\binom{2n-1}{n} = \binom{2n}{n}$.

(f) Find $\sum_{n=0}^{\infty} \binom{2n-1}{n} x^n$.

3. Compositions, multisets, and balls into boxes

Consider the problem of counting the number of ways to pick bagels at a bagel shop. Suppose you want to buy in a dozen bagels, and when you get to the shop, there are three kinds of bagels to choose from: plain, raisin, and sesame. How many different choices can you make? How about if you wanted to make sure you get at least one of every kind?

3.1. Compositions.

A *composition* of $n \in \mathbb{N}$ is a sequence $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ such that

$$\alpha_i \in \mathbb{Z}_{>0} \qquad \text{and} \qquad \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = n.$$

For example, the compositions of 3 are

$$(3), \quad (2,1), \quad (1,2), \quad \text{and} \quad (1,1,1).$$

We call the terms of the sequence *parts*, and the number of parts the *length*. If α has ℓ parts, we can call it an ℓ-combination.

If α is an ℓ-composition of n, define

$$S_\alpha = \{\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3, \ldots, \alpha_1 + \cdots + \alpha_{\ell-1}\} \subseteq [n-1].$$

The map $\theta : \alpha \to S_\alpha$ is a bijection between ℓ-compositions of n and $(\ell-1)$-subsets of $[n-1]$, we hace that there are

$$\binom{n-1}{\ell-1} \ \ell\text{-compositions of } n, \qquad \text{and} \qquad 2^{n-1} \text{ compositions of } n > 0.$$

A *weak composition* of n is a sequence $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_\ell)$ such that

$$\alpha_i \in \mathbb{Z}_{\geq 0} \qquad \text{and} \qquad \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = n.$$

For example, the weak 2-compositions of 3 are

$$(3,0), \quad (0,3), \quad (2,1), \quad \text{and} \quad (1,2).$$

3.2. **Stars and bars.** To each ℓ-composition α of n, associate a picture of n stars and $\ell-1$ bars which separate the n stars into groups with α_1 stars, then α_2, stars, and so on. For example, the composition $(1,3,1,1,2)$ goes with the picture

$$*|***|*|*|**.$$

One way to count the number of such arrangements is to think about the spaces between the stars an places where a bar can be placed or not. There are $n-1$ such spaces. If we want an ℓ-composition, we want to choose $\ell-1$ spaces to place bars. So there are $\binom{n-1}{\ell-1}$ such arrangements.

Another way to think about these arrangements is as length $n+\ell-1$ sequences of stars and bars, with exactly n stars. But since there is at least one star to either side of the bars, these arrangements are actually in bijection with stars and bars arrangements with $n-\ell$ stars (remove one from each compartment) and $\ell-1$ bars, now with no restrictions. There are $\binom{n-\ell+\ell-1}{\ell-1} = \binom{n-1}{\ell-1}$ of these (think of it as $n-\ell+\ell-1$ coin tosses where heads comes up exactly $\ell-1$ times).

Exercise 9.

(a) (i) For each composition α of 3, give S_α.
   (ii) Prove that the map $\theta : \alpha \to S_\alpha$ is a bijection between ℓ-compositions of n and $(\ell-1)$-subsets of $[n-1]$. [Hint: define θ^{-1} and show that it is well-defined.]

(b) Describe a bijection between weak ℓ-compositions of n and arrangements of stars and bars, and conclude how many weak ℓ-compositions there are of n.

(c) Give a bijection between E_n, the set of compositions of n with an even number of even parts, and O_n, the set of compositions of n with an odd number of even parts.
[For example, $E_3 = \{(1,1,1),(3)\}$ and $O_3 = \{(2,1),(1,2)\}$.]

   Use your bijection to conclude how many compositions there are with an even number of even parts.

(d) Show that the total number of all parts in all compositions of n is $(n+1)2^{n-2}$.
[For example, the compositions of 3 are $(3),(2,1),(1,2)$, and $(1,1,1)$, which all together have 8 parts; and $8 = (3+1)2^{3-2}$.]
[Hint: Use a stars and bars argument: if you line up all the compositions, how many bars appear in total? Explain why the total number of parts is equal to the total number of bars, plus the total number of compositions.]

3.3. **Integer solutions.** Consider the linear equation

$$x_1 + x_2 + \cdots + x_\ell = n. \tag{3.1}$$

Positive integer solutions to (3.1) are in bijection with ℓ-compositions of n; Non-negative integer solutions are in bijection with weak ℓ-compositions of n.

Exercise 10.

(a) Explain why there are $\binom{n-1}{\ell}$ positive integer solutions to

$$x_1 + x_2 + \cdots + x_\ell < n$$

and $\binom{n+\ell}{\ell}$ non-negative integer solutions to

$$x_1 + x_2 + \cdots + x_\ell \leq n$$

by setting up a linear equations that have the appropriate number of solutions.

(b) (i) How many solutions are there to the equation

$$x_1 + x_2 + x_3 = 10,$$

where x_1, x_2, and x_3 are nonnegative integers? How many solutions are there if x_1, x_2, and x_3 are positive integers?

(ii) How many solutions are there to the equation

$$x_1 + x_2 + x_3 \leq 10,$$

where x_1, x_2, and x_3 are nonnegative integers?

3.4. **Balls into boxes.** Now let's get back to the bagel problem. Drawing the 12 bagels as a line of 12 stars. Use two bars to separate the stars into three categories. Assign the first category to plain, the second to raisin, and the third to sesame. For example, 2 plain, 1 raisin, and 9 sesame is drawn like

$$**|*|*********;$$

similarly, all plain looks like

$$************||.$$

So there are $\binom{12+3-1}{3-1}$ possible choices. If you want at least one of any kind, there are $\binom{12-1}{3-1}$ possible choices.

This problem is generally known as the "indistinguishable balls into distinguishable boxes" problem. Here, the balls are the choices; the boxes are the categories of bagels.

Exercise 11.

(a) Suppose you've got eight varieties of doughnuts to choose from at a doughnuts shop.
   (i) How many ways can you pick 6 doughnuts?
   (ii) How many ways can you pick a dozen doughnuts?
   (iii) How many ways can you pick a dozen doughnuts with at least one of each kind?

(b) How many different combinations of pennies, nickels, dimes, quarters, and half dollars can a jar contain if it has 20 coins in it?

3.5. **Multisets.** A *multiset* is like a set, only with possibly repeated elements, like

$$\{1, 1, 2, 5, 6, 6, 6\}.$$

More precisely a finite multiset m on a set S is a pair (S, ν) where $\nu : S \to \mathbb{N}$ is a function satisfying $\sum_{x \in S} \nu(x) < \infty$. For example, the above multiset is the pair

$$S = \{1, 2, 5, 6\} \qquad \text{and} \qquad \nu(1) = 2, \nu(2) = 1, \nu(5) = 1, \nu(6) = 3.$$

In other words, ν tells you the multiplicity of each element of S in M. Note that $|M| = \sum_{x \in S} \nu(x) \infty$. For shorthand, we can write

$$M = \{x^{\nu(x)} \mid x \in S\}.$$

For example, the multiset above is
$$\{1^2, 2, 5, 6^3\}.$$
Denote the size-k multisets on S by
$$\left(\!\!\binom{S}{k}\!\!\right) = \{(S,\nu) \mid \nu : S \to \mathbb{N}, \sum_{x\in S} \nu(x) = k\}; \quad \text{and let } \left|\left(\!\!\binom{S}{k}\!\!\right)\right| = \left(\!\!\binom{|S|}{k}\!\!\right).$$
We say $\left(\!\binom{n}{k}\!\right)$ as "*n multichoose k*". Another name for subsets of S are *combinations of S without repetitions*; similarly, multisets on S are *combinations of S (with repetitions).*

If $M' = (S, \nu')$ is another multiset on S, then we say that M is a *submultiset* of M if $\nu'(x) \le \nu(x)$ for all $x \in S$. The number of submultisets of M is $\prod_{x\in S}(\nu(x) + 1)$, since for each $x \in S$ there are $\nu(x) + 1$ possible values of $\nu'(x)$.

So what is $\left(\!\binom{n}{k}\!\right)$? Well, let
$$S = \{y_1, \dots, y_n\}, \qquad \text{and} \qquad x_i = \nu(y_i).$$
Then the k-multisets on S are in bijection with nonnegative integer solutions to $x_1 + \cdots + x_n = k$. So
$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{n-1}.$$

We can also give bijective proofs directly, using the combinatorial definitions of $\left(\!\binom{n}{k}\!\right)$ and $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$, i.e.
$$\left(\!\!\binom{n}{k}\!\!\right) = \left|\left(\!\!\binom{S}{k}\!\!\right)\right| \text{ where } |S| = n \qquad \text{and} \qquad \binom{n+k-1}{k} = \left|\left(\!\!\binom{S}{k}\!\!\right)\right| \text{ where } |S| = n+k-1.$$
So construct a bijection between the k-subsets of $[n+k-1]$ and the k-multisets on $[n]$ as follows. First let $A = \{a_1, a_2, \dots, a_k\} \subseteq [n+k-1]$ so that $a_1 < a_2 < \cdots < a_k$. Let $b_i = a_i - i + 1$. Then $\{b_1, \dots, b_k\}$ is a multiset on $[n]$ (why??). The inverse function is $a_i = b_i + i - 1$; we just need to argue that $\{a_1, \dots, a_k\}$ is a proper subset of $[n+k-1]$. This proof illustrates the technique of *compression*, where we convert a strictly increasing sequence to a weakly increasing sequence.

Exercise 12. (a) List the 3-multisets on [2].
(b) How many 5-multisets are there on [8]?
(c) How many multisets are there on [8]? (Of any size)
(d) Describe a bijection between k-multisets on $[n]$ and stars and bars arrangements with k stars and $n-1$ bars.

Generating function. Recall when we were looking at the binomial coefficients, we calculated in equation (2.3) that if $S = \{x_i, \dots, x_n\}$, then
$$(1 + x_1)(1 + x_2)\cdots(1 + x_n) = \sum_{T\subseteq S} \prod_{x_i\in T} x_i,$$
so that, as in (2.4),
$$(1 + x)^n = \sum_{T\subseteq S} \prod_{x_i\in T} x = \sum_{T\subseteq S} x^{|T|} = \sum_{k=0}^{n} \binom{n}{k} x^k.$$
Similarly, we can show that
$$(1 + x_1 + x_1^2 + \cdots)(1 + x_2 + x_2^2 + \cdots)\cdots(1 + x_n + x_n^2 + \cdots) = \sum_{M=(S,\nu)} \prod_{x_i\in S} x_i^{\nu(x_i)},$$

where the sum is over all finite multisets M on S. Note that the closed form of the lefthand side is

$$\prod_{i=1}^{n}(1-x_i)^{-1}.$$

Now set $x_i = x$ for all i. Then

$$\begin{aligned}(1+x+x^2+\cdots)^n &= \sum_{M=(S,\nu)} x^{\nu(x_i)+\cdots+\nu(x_n)} \\ &= \sum_{M=(S,\nu)} x^{|M|} \\ &= \sum_{k\in\mathbb{N}} \left(\!\!\binom{n}{k}\!\!\right) x^k.\end{aligned}$$

On the other hand, since

$$1+x+x^2+\cdots = (1-x)^{-1},$$

we have

$$(1+x+x^2+\cdots)^n = (1-x)^{-n} = \sum_{k\in\mathbb{N}} \binom{-n}{k}(-1)^k x^k.$$

You will show on the homework that $\binom{-n}{k}(-1)^k = \binom{n+k-1}{k}$, so that comparing coefficients of x^k gives

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

Exercise 13.

(a) Prove

$$(1+x_1+x_1^2+\cdots)(1+x_2+x_2^2+\cdots)\cdots(1+x_n+x_n^2+\cdots) = \sum_{M=(S,\nu)} \prod_{x_i\in S} x_i^{\nu(x_i)},$$

by induction on n.

(b) Show algebraically that $\binom{-n}{k}(-1)^k = \binom{n+k-1}{k}$.

(c) (a) Write the generating function (both series and closed form) for the number of weak compositions of n with k parts.
[Hint: This should look something like the generating function for multisets.]

(b) Write the generating function (both series and closed form) for the number of (not weak) compositions of n with k parts.

(c) Write the generating function for the number of weak compositions of n with k parts, all less than j.

(d) Give a generating function proof that the number of weak compositions of n into k parts, with each part less than j, is

$$\sum_{\substack{r,s\in\mathbb{N}\\ r+sj=n}} (-1)^s \binom{k+r-1}{r}\binom{k}{s}.$$

3.6. **Multinomial coefficient.** Think about n choose k as dividing n distinguishable items into two distinguishable categories, with k items in the first, and $n-k$ in the second. As a generalization, let $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ be a composition of n; then let $\binom{n}{\alpha_1,\ldots,\alpha_\ell}$ be the number of ways of taking n distinguishable items, and placing α_1 into one category, α_2 into a second category, ..., and α_ℓ into an ℓth category. For example, say you have a party committee with 10 people, and you want to pick 3 people to work on publicity, 3 people to organize food, and 4 people to work on decorations; there are $\binom{10}{3,3,4}$ ways to do this (whatever that is). In this way, $\binom{n}{k} = \binom{n}{k,n-k}$. The number $\binom{n}{\alpha_1,\ldots,\alpha_\ell}$ is called a *multinomial coefficient*.

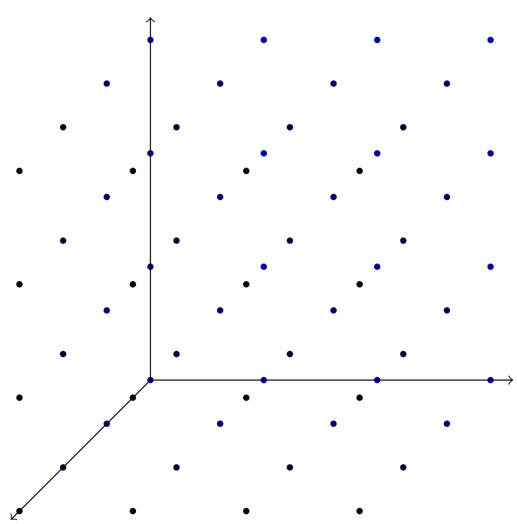
We can reinterpret $\binom{n}{\alpha_1,\ldots,\alpha_\ell}$ as permutations of a size-n multiset of the form $\{k_1^{\alpha_1}, \ldots, k_\ell^{\alpha_\ell}\}$. Define a *permutation* of a finite multiset M is a bijection $w : [|M|] \to M$; equivalently, it's a word $w(1)w(2)\cdots w(|M|)$ made by rearranging the elements of M. For example, the multiset $\{a, a, b\}$ has $3 = \binom{3}{2,1}$ permutations/words:

$$\begin{array}{ccc} w : 1 \mapsto a, 2 \mapsto a, 3 \mapsto b, & w : 1 \mapsto a, 2 \mapsto a, 3 \mapsto a, & w : 1 \mapsto a, 2 \mapsto a, 3 \mapsto a, \\ \text{word: } aab, & \text{word: } aba, & \text{word: } baa. \end{array}$$

Let cfS_M be the set of permutations of $M = (S, \nu)$ where $S = \{x_1, \ldots, x_n\}$; then

$$|\mathcal{S}_M| = \binom{|M|}{\nu(x_1), \ldots, \nu(x_n)}.$$

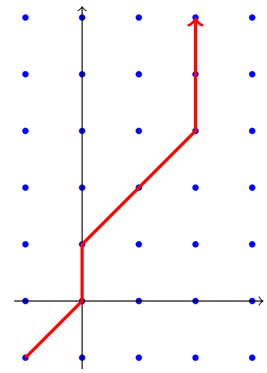
Lattice paths. Visualize $\mathbb{Z}^d$ as a lattice of points



A *lattice path* of length k with steps in a set $S \subset \mathbb{Z}^d$ is a sequence of points $v_0, v_1, v_2, \ldots, v_k$ such that $v_{i+1} - v_i \in S$. For example, if $d = 2$ and $S = \{(1, 1), (0, 1)\}$, one lattice path starting at

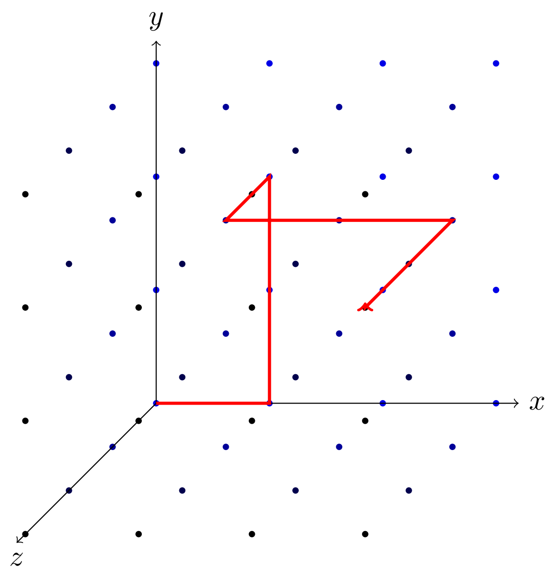
$(-1,-1)$ is

$$((-1,-1),(0,0),(0,1),(1,2),(2,3),(2,4),(2,5)), \qquad \text{drawn as}$$



As another example, if $d = 3$ and $S = \{(1,0,0),(0,1,0),(0,0,1)\}$, one lattice path starting at $(0,0,0)$ is

$$\begin{array}{l}((0,0,0),(1,0,0),(1,1,0),(1,2,0),(1,2,1), \\ \quad (2,2,1),(3,2,1),(3,2,2),(3,2,3),(3,2,3))\end{array} \qquad \text{drawn as}$$



Exercise 14.

(a) **Deriving multinomial coefficients algebraically.** Let $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ be a composition of n. Use the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ to compute $\binom{n}{\alpha_1,\ldots,\alpha_\ell}$, noting that you can first choose the α_1 items from n, then α_2 from $n - \alpha_1$, then α_3 from $n - (\alpha_1 + \alpha_2)$, and so on.

(b) **Multinomial theorem.** Following our proof of the binomial theorem, show that

$$(x_1 + x_2 + \cdots + x_\ell)^n = \sum_{\substack{(\alpha_1,\ldots,\alpha_\ell)\in\mathbb{N}^\ell \\ \alpha_1+\cdots+\alpha_\ell=n}} \binom{n}{\alpha_1,\ldots,\alpha_\ell} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_\ell^{\alpha_\ell}.$$

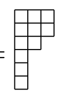
[Hint: Recall that the key computation for the binomial theorem was that, for $S = \{x^{(1)}, \ldots, x^{(n)}\}$,

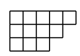
$$\prod_{x^{(i)}\in S} (1 + x^{(i)}) = \sum_{T\subseteq S} \prod_{x^{(i)}\in T} x^{(i)}, \qquad \text{so that} \qquad (1+x)^n = \sum_{T\subseteq S} \prod_{x^{(i)}\in T} x = \sum_{T\subseteq S} x^{|T|}.$$

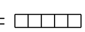
The former we had to prove by induction on n. Now fix ℓ, and let $S = \{x_i^{(j)} \mid 1 \le i \le \ell, 1 \le j \le n\}$ (so that there are n distinct variables associated to each x_i), and walk through a similar proof.]

(c) **Lattice paths.** Proposition 1.2.1 in EC1 says the following. Let $v = (a_1, \dots, a_d) \in \mathbb{N}^d$, and let e_i denote the ith unit coordinate vector in $\mathbb{Z}^d$. The number of lattice paths in $\mathbb{Z}^d$ from the origin $(0, 0, \dots, 0)$ to v with steps in $\{e_1, \dots, e_d\}$ is given by the multinomial coefficient $\binom{a_1+\cdots+a_d}{a_1,\dots,a_d}$.
  (i) Check this proposition for $d = 2$ with the point $v = (2, 3)$.
  (ii) Check this proposition for $d = 3$ with the point $v = (1, 1, 2)$.
  (iii) Prove this theorem (spell out the book's proof with more details).

(d) **Integer partitions.** An (integer) partition $\lambda = (\lambda_1, \dots, \lambda_\ell)$ of n is a composition of n satisfying $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_\ell > 0$. We draw partitions as n boxes piled up and to the left into a corner, with λ_1 boxes in the first row, λ_2 boxes in the second row, and so on. For example,

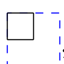
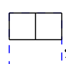
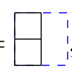
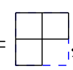
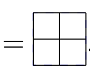
$\lambda = (3, 3, 2, 1, 1, 1) =$  is a partition of 11,

$\lambda = (5, 4, 3) =$  is a partition of 12,

$\lambda = (5) =$  is a partition of 5, and

$\lambda = \emptyset$ is a partition of 0.

The six partitions to fit in a 2×2 square are

$\emptyset =$ , $(1) =$ , $(2) =$ , $(1,1) =$ , $(2,1) =$ , and $(2,2) =$ .

Use lattice paths to count the number of integer partitions fitting into a $m \times n$ rectangle.

4. Aside: Fibonacci numbers

A *recurrence relation* for a sequence $(a_n)_{n\in\mathbb{N}}$ is an identity determining the nth term from the previous terms. For example, the *Fibonacci recurrence* is

$$a_n = a_{n-1} + a_{n-2}. \qquad (*)$$

Any recurrence relation needs *initial conditions* to determine the corresponding sequence. For example, with the Fibonacci recurrence,

the initial conditions $a_0 = 0, a_1 = 3$ determines the sequence $0, 3, 3, 6, 9, 15, \dots$;
the initial conditions $a_0 = 1, a_1 = -2$ determines the sequence $1, -2, -1, -3, -4, -7, \dots$.

The *Fibonacci numbers* are defined recursively by $f_n = f_{n-1} + f_{n-2}$ with initial conditions $f_0 = 0$ and $f_1 = 1$. So $f_2 = 0 + 1 = 1$, $f_3 = 1 + 1 = 2$, $f_4 = 2 + 1 = 3$, and so on. Note that any sequence that satisfies $(*)$, and that has initial conditions consisting of two consecutive Fibonacci numbers, can be expressed in terms of f_i's (see the example below).

Exercise 15.

(a) For each of the following, give examples for small values of n. Then express the following numbers in terms of the Fibonacci numbers.

(i) **Example:** The number of subsets S of the set $[n] = \{1, 2, \ldots, n\}$ such that S contains no two consecutive integers.

Answer: Let a_n be the number of good subsets of $[n]$. Note that $a_1 = |\{\emptyset, \{1\}\}| = 2$ and $a_3 = \{\emptyset, \{1\}, \{2\}\} = 2$.
Now divide S into 2 cases: either it contains n or it doesn't. Since every good subset without n is also a good subset of $[n-1]$, and vice versa, the number of good subsets without n is a_{n-1}. Similarly $S \mapsto S - \{n\}$ is a bijection between good subsets of $[n]$ containing n and good subsets of $[n-2]$, the number of good subsets of $[n]$ containing n is a_{n-2}. So $a_n = a_{n-1} + a_{n-2}$, with $a_1 = 2 = f_3$, $a_2 = 3 = f_4$. This is the same recurrence that determines f_n, but shifted so that $a_n = f_{n+2}$.
So there are $\boxed{f_{n+2}}$ good subsets of $[n]$. $\square$

NOTE: For many of these, this is the strategy you want. Make a recurrence relation that looks like the Fibonacci recurrence, and shift appropriately. For at least one, you'll want to use a previous part.

(ii) The number of compositions of n into parts greater than 1.
(iii) The number of compositions of n into parts equal to 1 or 2.
(iv) The number of compositions of n into odd parts.
(v) The number of sequences $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ of 0s and 1s such that $\varepsilon_1 \leq \varepsilon_2 \geq \varepsilon_3 \leq \varepsilon_4 \geq \cdots$.
(vi) The number of sequences $(T_1, T_2, \ldots, T_k)$ of subsets T_i of $[n]$ such that $T_1 \subseteq T_2 \supseteq T_3 \subseteq T_4 \supseteq \cdots$.
(vii) The sum $\sum \alpha_1 \alpha_2 \cdots \alpha_\ell$ over all 2^{n-1} compositions $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ of n.
[Hint: this sum counts the number of ways of inserting at most one vertical bar in each of the $n-1$ spaces between stars in a line of n stars, and then circling one star in each compartment. Now try replacing bars, un-circled stars, and circled stars by 1's, 2's, and 1's, respectively. Use a previous part.]

(b) Consider the identity

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

(i) Check this identity for $n = 2$ and 3.
(ii) Prove this identity recursively by showing that it satisfies the Fibonacci recurrence, and that it holds for the first 2 values.
(iii) Prove this identity combinatorially. Namely, first show combinatorially that the number of k-subsets of $[n-1]$ containing no two consecutive integers is $\binom{n-k}{k}$, and then use (i).

(c) Note that EC1, Example 1.1.12 computes the generating function for the sequence $(a_i)_{i \in \mathbb{N}}$, where $a_i = f_{i+1}$ (so $a_0 = 1, a_1 = 1, a_2 = 2, a_3 = 3$, and so on). Repeat this computation for $(f_i)_{i \in \mathbb{Z}_{>0}}$, making the appropriate changes to accommodate the shift.

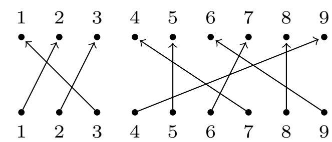
5. Permutations

"Permutations of sets and multisets are among the richest objects in enumerative combinatorics. A basic reason for this fact is the wide variety of ways to represent a permutation combinatorially. We have already seen that we can represent a set permutation either as a word or a function." For a finite set S, a *permutation* is a bijection $w : S \to S$, or equivalently, a bijection $w : [n] \to S$. For the latter, writing $w(i) = w_i$, the corresponding word is $w_1 w_2 \cdots w_n$. We will see several more representations as we go along.

We will also see several *statistics* of permutations. Here statistic usually means things like number of permutations satisfying some condition, or number of fixed points in a given permutation, or number of permutations that have no fixed points, and so on.

5.1. **Cycles and Inversions.** We follow Section 1.3 of EC1 closely here, with one addition. There is an additional representation of a permutation that I would like us to consider. For a permutation $w : [n] \to [n]$, the *diagrammatic* representation is given by drawing two rows pf n vertices, one above the other, labeled by $1, \ldots, n$. Then draw arrows$a \to b$ from the bottom row to the top row according whenever $w : a \mapsto b$. For example, the permutation

$w = (312)(5)(8)(9674)$ has diagrammatic representation



Exercise 16. For each of the following permutations of [9], give whichever of the following is not already given.

(i) The function representation.
(ii) The word representation.
(iii) The standard cycle representation.
(iv) The digraph representation.
(v) The word given by the fundamental bijection (the ˆ word).
(vi) The diagrammatic representation.

(a) $w : [9] \to [9]$, the permutation given by

$$1 \mapsto 8, \quad 2 \mapsto 9, \quad 3 \mapsto 7, \quad 4 \mapsto 4, \quad 5 \mapsto 6, \quad 6 \mapsto 5, \quad 7 \mapsto 1, \quad 8 \mapsto 3, \quad 9 \mapsto 2.$$

(b) $v = (6325)(1)(9478)$.
(c) The permutation u determined by $\hat{u} = 123456798$.

Exercise 17. For each of the permutations in 16, give the cycle type, and the number of permutations of [9] that have the same cycle type. Additionally, give the following.

(a) For w in 16(a), verify the equation $n = \sum_i ic_i$.
(b) For v in 16(b), verify that v has the same number of cycles as $\hat{v}$ has left-to-right maxima (be sure to identify the left-to-right maxima).
(c) For u in 16(c), what is $t^{\text{type}(u)}$?

Exercise 18. Show that the number of permutations $w \in \mathcal{S}_n$ fixed by the fundamental bijection $\mathcal{S}_n \xrightarrow{\hat{}} \mathcal{S}_n$ (i.e. $|\{w \in \mathcal{S}_n \mid \hat{w} = w\}|$) is the Fibonacci number F_{n+1}.

Exercise 19.

(a) For $Z_n = \frac{1}{n!}\sum_{w\in\mathcal{S}_n} t^{\text{type}(w)}$, calculate Z_1, Z_2, Z_3, and Z_4 explicitly (verifying the formulas between (1.25) and (1.26) in EC1).
(b) For $E_k(n) = \frac{1}{n!}\sum_{w\in\mathcal{S}_n} c_k(w)$, verify that

$$E_k(n) = \frac{\partial}{\partial t_k} Z_n(t_1, t_2, \ldots, t_n)\Big|_{\substack{t_i=1 \\ i=1,\ldots,n}}.$$

(c) Give a combinatorial proof of $E_k(n) = 1/k$ by (i) explaining why there are $\binom{n}{k}(k-1)!$ k-cycles, (ii) explaining why each k-cycle appears in $(n-k)!$ permutations, and (iii) computing $E_k(n)$ using these two values.

Exercise 20. (a) Compute the signless Stirling numbers of the first kind $c(n, k)$ for $n = 1, 2, 3, 4$ and $k = 1, \ldots, n$ (i) directly, and (ii) using the recursion. Then give the Stirling numbers of the first kind $s(n, k)$ for $n = 1, 2, 3, 4$ and $k = 1, \ldots, n$.

(b) Verify $\sum_{k=0}^{n} c(n, k)t^k = t(t + 1)(t + 2)\cdots(t + n - 1)$ for $n = 0, 1, 2, 3, 4$.

Exercise 21. (Proving Proposition 1.3.7)

(a) Verify that $\sum_{k=0}^{n} c(n, k)t^k = n!Z_n(t, t, \ldots, t)$ for $n = 1, 2, 3$, and then explain why this identity holds in general.

(b) Cary out another example for the third proof of Proposition 1.3.7, again for $n = 9$ and $k = 4$.

(c) Walk through and complete the third proof of Proposition 1.3.7.

(d) Read the fourth proof and example 1.3.9. Cary out another example for $n = 9$ and $k = 4$ for a sequence $(a_1, \ldots, a_n)$ of your choice.

Exercise 22. Using only the combinatorial definitions of the signless Stirling numbers $c(n, k)$, give formulas for $c(n, 1)$, $c(n, n)$, $c(n, n - 1)$, and $c(n, n - 2)$.

5.2. **Descents.** We follow Section 1.4 of EC1 closely here.

Exercise 23. Inversions and descents.

(a) For each of $w \in \mathcal{S}_3$, write w in word form and give (i) w^{-1}, (ii) $I(w)$, (iii) $\mathrm{inv}(w)$, (iv) $\mathrm{code}(w)$, (v) $D(w)$, (vi) $\mathrm{des}(w)$, and (vii) $\mathrm{maj}(w)$. (Make a table.)

(b) Use your calculations in (a) to verify
  (i) $I(w^{-1})$ and $\mathrm{code}_i(w) = \#\{j > i \mid w(j) < w(i)\}$ are equivalent definitions of $\mathrm{code}(w)$,
  (ii) Corollary 1.3.13,
  (iii) Proposition 1.3.14,
  (iv) the proof of Prop 1.3.14 (show the bijection between inversions (i, j) in w and inversions (w_i, w_j) in w^{-1}); and
  (v) equation (1.41),

  for $n = 3$.

Let $\mathrm{des}(w) = |D(w)| = |\{i \mid w_i > w_{i+1}\}|$. Let

$$A_d(x) = \sum_{w \in \mathcal{S}_d} x^{1+\mathrm{des}(w)} = \sum_{k=1}^{d} A(d, k)x^k,$$

where

$$A(d, k) = |\{w \in S_k \mid \mathrm{des}(w) = k - 1\}|.$$

We call A_d the *Eulerian polynomial* and $A(d, k)$ the *Eulerian numbers*, where $A(0, k) = \delta_{0,k}$. For example,

$$\begin{aligned} A_0 &= 1; \\ A(1, 0) = 0, A(1, 1) &= |\{1\}| = 1, \text{ so } A_1 = x; \\ A(2, 0) = 0, A(2, 1) &= |\{12\}| = 1, A(2, 2) = |\{21\}| = 1, \text{ so } A_2 = x + x^2; \end{aligned}$$

and so on.

Write $w = (a_1a_2\cdots a_{i_1})(a_{i_1+1}\cdots a_{i_2})\cdots(a_{i_{k-1}+1}\cdots a_d)$ in standard form, so that $a_1, a_{i_1+1}, \cdots, a_{i_{k-1}+1})$ are the largest in their cycles, and $a_1 < a_{i_1+1} < \cdots < a_{i_{k-1}+1}$. If $w(a_i) \neq a_{i+1}$ (jumping to a new cycle), then $a_i < a_{i+1}$. Thus for $i \neq d$, $a_i < a_{i+1}$ if and only if $w(a_i) \geq a_i$. Thus

$$d - \mathrm{des}(\hat{w}) = |\{i \in [d] \mid w(i) \geq i\}|.$$

We call i such that $w(i) \geq i$ a *weak excedance*, and i such that $w(i) > i$ is an *excedance.*

Proposition 5.1 (Prop. 1.4.4)**.** *For $d \geq 0$,*

$$\sum_{m \geq 0} m^d x^m = \frac{A_d(x)}{(1-x)^{d+1}}.$$

Proof. We prove this by induction on d. For $d = 0$, $A_0 = 1$, and both sides are the geometric series $\sum_{m \geq 0} x^m$.

Now assuming this identity holds for a fixed d, differentiate both sides and multiply by x to get

$$\sum_{m \geq 0} m^{d+1} x^m = \sum_{m > 0} m^{d+1} x^m = x \frac{A_d'(x)(1-x)^{d+1} + A_d(x)(d+1)(1-x)^d}{(1-x)^{2d+2}} = x \frac{A_d'(x)(1-x) + A_d(x)(d+1)}{(1-x)^{d+2}}$$

Thus if we wish to show our identity for $d+1$, we want to show that

$$\frac{A_{d+1}(x)}{(1-x)^{d+2}} = \sum_{m \geq 0} m^{d+1} x^m = x \frac{A_d'(x)(1-x) + A_d(x)(d+1)}{(1-x)^{d+2}},$$

i.e. that

$$\begin{aligned}
\sum_{k=1}^{d+1} A(d+1,k) x^k &= A_{d+1}(x) \\
&= x((1-x)A_d'(x) + (d+1)A_d(x)) \\
&= x(1-x) \sum_{k=1}^{d} A(d,k) k x^{k-1} + (d+1) x \sum_{k=1}^{d} A(d,k) x^k \\
&= \sum_{k=1}^{d} A(d,k) k x^k - \sum_{k=2}^{d+1} A(d,k-1)(k-1) x^k + (d+1) \sum_{k=2}^{d+1} A(d,k-1) x^k \\
&= A(d,1)x + A(d,d)x^{d+1} + \sum_{k=2}^{d+1} (A(d,k)k + (d+1)A(d,k-1) - A(d,k-1)(k-1)) x^k \\
&= A(d,1)x + A(d,d)x^{d+1} + \sum_{k=2}^{d+1} (kA(d,k) + (d-k+2)A(d,k-1)) x^k
\end{aligned}$$

Comparing coefficients on both sides, we wish to show

$$A(d+1,k) = kA(d,k) + (d-k+2)A(d,k-1).$$

The left hand side is permutations of $d+1$ with $k-1$ descents. We can get such a permutation from a permutation of d in one of two ways. Either take a permutation w of d with $k-1$ descents (of which there are $A(d,k)$) and insert $d+1$ after any w_i for $i \in D(w)$, or at the end (k ways); or take a permutation w of d with $k-2$ descents (of which there are $A(d,k-1)$) and insert $d+1$ after any w_i for $i \notin D(w)$ ($d-(k-2)$ ways). Thus we have proved the recursion. (We prove the extremal values in the homework.) $\square$

A bit of algebraic manipulation, using Prop. 1.4.4, also gives us the following (see proof in book).

Proposition 5.2 (Prop. 1.4.5)**.** *For $d \geq 0$,*

$$\sum_{d \geq 0} A_d(x) \frac{t^d}{d!} = \frac{1-x}{1-xe^{(1-x)t}}.$$

Let $\mathrm{maj}(w) = \sum_{i \in D(w)} i$, called the *major index* of w. We will give a bijective proof that the number of permutations of n that have k inversions is the same as as the number of permutations of n that have major index k. The generating function version of this statement is

$$\sum_{w \in \mathcal{S}_n} q^{\mathrm{inv}(w)} = (\mathbf{n})! = \sum_{w \in \mathcal{S}_n} q^{\mathrm{maj}(w)}.$$

To this end, we have the following proposition.

Proposition 5.3 (Prop. 1.4.6)**.** *We have*

$$\sum_{w \in \mathcal{S}_n} q^{\mathrm{maj}(w)} = (\mathbf{n})!.$$

Proof. Define a bijection $\varphi : \mathcal{S}_n \to \mathcal{S}_n$ recursively as follows. Let $w = w_1 \cdots w_n \in \mathcal{S}_n$. Define a sequence of permutations $\gamma_1 \cdots \gamma_n$, where γ_i is a permutation of $\{w_1, w_2, \dots, w_i\}$.

First, let $\gamma_1 = w_1$. Then define $\gamma_2, \gamma_3, \dots$, in order as follows, viewing them as words. If the last letter of γ_i is greater than w_{i+1}, then split γ_i at each letter greater than w_{i+1}; if the last letter is smaller, split γ_i at each letter that's smaller that w_{i+1}. Cyclicly shift each grouping to the right, placing the last group first, and then put w_{i+1} at the end. (So the last letter of γ_i is always w_i.) Set $\varphi(w) = \gamma_n$. (See Example 1.4.7, with $w = 683941725$.)

Now the claim is that

$$\mathrm{maj}(w) = \mathrm{inv}(\varphi(w)),$$

where again

$$\mathrm{inv}(w) = |\{(i,j) \mid i < j, w_i > w_j\}$$
$$\mathrm{maj}(w) = \sum w_i > w_{i+1} i.$$

Let $\eta_k = w_1 \cdots w_k$, and show by induction that $\mathrm{inv}(\gamma_k) = \mathrm{maj}(\eta_k)$.

First, $\mathrm{inv}(\gamma_1) = \mathrm{maj}(\eta_1) = 0$. So assume that $\mathrm{inv}(\gamma_k) = \mathrm{maj}(\eta_k)$ for a fixed $k < n$.

Case 1: The last letter w_k of γ_k is greater than w_{k+1}: Then $k \in D(w)$, so since $\mathrm{maj}(\eta_{k+1}) = \mathrm{maj}(\eta_k w_{k+1}) \mathrm{maj}(\eta_k) + k$, we need to show that $\mathrm{inv}\gamma_{k+1} = k + \mathrm{inv}\gamma_k$. When we break γ_k into its compartments, the last letter in each compartment C is the largest. So when we cyclicly rotate the compartments, we introduce $|C| - 1$ new inversions from that compartment. Then appending w_{k+1}, we add a new inversion per each of the $m = |\{C\}|$ compartments. So there are

$$m + \sum_C (|C| - 1) = k$$

new inversions total.

Case 2: The last letter w_k of γ_k is smaller than w_{k+1}: Homework.

It remains to show that φ is a bijection. Basically, the task is to unwind the recursive definition of γ_i. See the last paragraph of the poof of 1.4.6.

□

Exercise 24.

(a) Eulerian numbers.
   (i) Verify that $A_3 = x + 4x^2 + x^3$ and $A_4 = x + 11x^2 + 11x^3 + x^4$.
   (ii) In general, for $d > 0$, what are $A(d, 1)$ and $A(d, d)$?
   (iii) What is $A_d(1)$?

(b) Excedances.

(i) Show that $w = w_1 w_2 \cdots w_d$ has k weak excedances if and only if $u = u_1 u_2 \cdots u_d$, defined by $u_i = d + 1 - w_{d-i+1}$, has $d - k$ excedances.
(ii) Show w has $d - 1 - j$ descents if and only if $w_d w_{d-1} \cdots w_1$ has j descents.
(iii) Show that

$$A(d, k+1) = |\{w \in \mathcal{S}_d \mid w \text{ has } k \text{ excedances}\}|$$

and

$$A(d, k+1) = |\{w \in \mathcal{S}_d \mid w \text{ has } k+1 \text{ weak excedances}\}|.$$

(c) Complete the proof of Prop. 1.4.6 by proving that $\mathrm{inv}(\gamma_k) = \mathrm{maj}(\eta_k)$ implies $\mathrm{inv}(\gamma_{k+1}) = \mathrm{maj}(\eta_{k+1})$ in the case where the last letter w_k of γ_k is smaller than w_{k+1}.

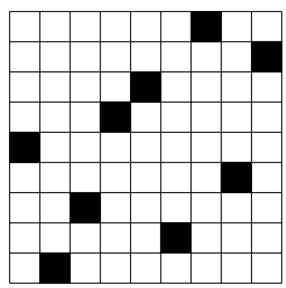
5.3. **Geometric representations of permutations.** We follow Section 1.5 of EC1 closely here.
The *permutation matrix* associated to w is

$$(P_w)_{i,j} = \begin{cases} 1 & \text{if } w(i) = j, \\ 0 & \text{otherwise.} \end{cases}.$$

This is a square $n \times n$ matrix with exactly one 1 in every row and column. Namely, there is a 1 in row i and column w_i. This establishes a bijection between permutations of $[n]$ and $n \times n$ binary matrices with exactly one 1 in every row and column. Sometimes, for visual effect, we replace 1's and 0's with other more distinctive symbols; for example, a $n \times n$ grid with boxes filled in corresponding to the 1's is easier to parse visually.

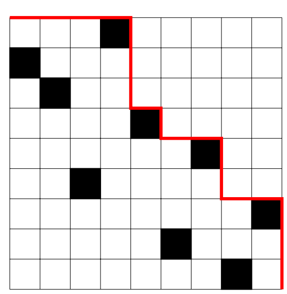
For example, the permutation $w = 795418362$, the corresponding permutation matrix is

$$\begin{bmatrix} 0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0 \\ 0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0 \\ 0&1&0&0&0&0&0&0&0 \end{bmatrix}$$

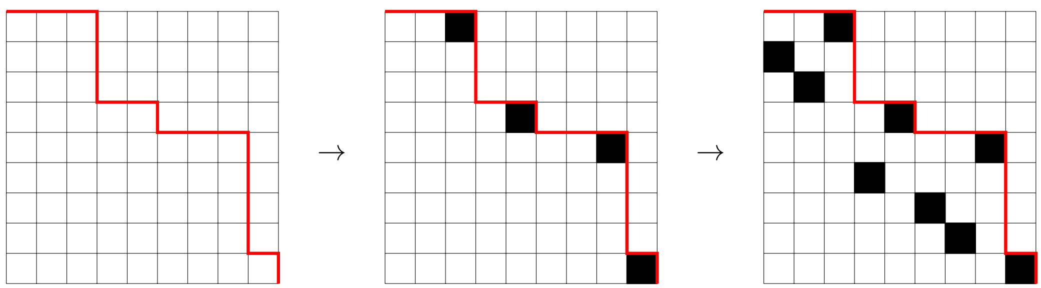


Decreasing subsequences: $w_{i_1} > \cdots > w_{i_k}$ inside a word $w = w_1 w_2 \cdots w_n$. We might like to count the number $f(n)$ of permutations of n with no decreasing sequences of length, say, 3 (we call these "321 avoiding", since 321 tells you that the pattern you're looking for is high-medium-low). Associate to each permutation a lattice path which starts in the upper-left corner, ends on the lower-right corner, and walks as south-west as it can while keeping 1's below it:

$w = 412573968$:



Notice that each such lattice path will always stay above the diagonal. Now, to each such lattice path, we can assign the permutation that has 1's at the "inside corners", and then fills in the rest of the "missing 1's" in increasing order:



$= 312584679.$

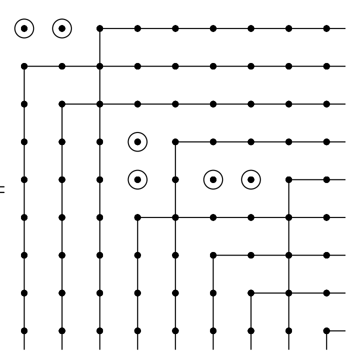
This is exactly the same as saying that there is a unique permutation with (1) a given set S of left-to-right maxima, (2) a specific set of locations for the left-to-right maxima, and (3) with $[n] - S$ appearing in increasing order. In the example, the left-to-right maxima are $S = \{3, 5, 8, 9\}$, and their locations are $\{1, 4, 5, 9\}$. This is the unique 321-avoiding permutation with left-to-right maxima determined by that path. So 321-avoiding permutations are in bijection with elementary lattice paths from $(0, 0)$ to (n, n) staying above (in this orientation) the diagonal. These are called *Dyck paths*.

Preview: There are $C_n = \frac{1}{n+1}\binom{2n}{n}$ of these, where C_n is the nth Catalan number. Most famous book exercise in enumerative combinatorics: Problem 6.19 from EC2.

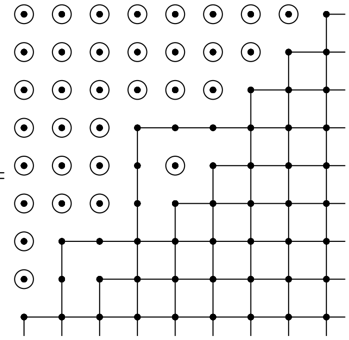
Pattern avoidance: Let u be a permutation of $[m]$ where $m \leq n$. We say that a permutation w of $[n]$ is *u-avoiding* if there is no subsequence of w (in word form) that follows the same pattern as u. So, for example, a 321-avoiding permutation is one which has no decreasing subsequence of length 3. On the other hand, a permutation $w = w_1 w_2 \cdots w_n$ is 312-avoiding if it does not contain any subsequence satisfying $i_1 < i_2 < i_3$ and $w_{i_1} > w_{i_3} > w_{i_2}$.

(Dot) diagram of w: Dram an $n \times n$ array of dots. If $w(i) = j$, draw a horizontal line to the right and a vertical line down from (i, j) (matrix coordinates: ith row, jth column). The diagram of w is the set of dots not crossed out:

$D_{312584679} =$

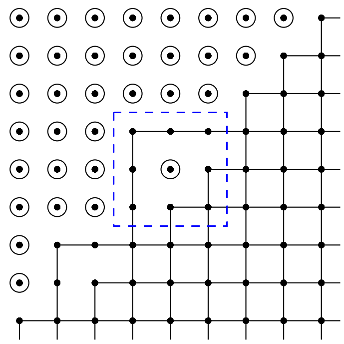
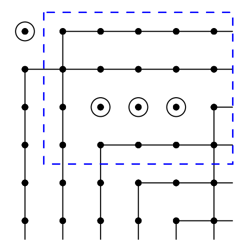


$= \{(1, 1), (1, 2), (4, 4), (5, 4), (5, 6), (5, 7)\},$

$D_{987465231} =$  $=$ $\{(1,1),\ldots,(1,8),$
$(2,1),\ldots,(2,7),$
$(3,1),\ldots,(3,6),$
$(4,1),(4,2),(4,3),$
$(5,1),(5,2),(5,3),(5,5),$
$(6,1),(6,2),(6,3),$
$(7,1),(8,1)\}$

Note that the inversion table $I = (a_1, \ldots, a_n)$ are given by a_i is the number of elements in column i of D_w. Similarly, code is read off of the rows. Finally, $D_w^t = D_{w^{-1}}$.

Consider the pattern $u = 132$. You can see 132 patterns in diagrams where there's windows like:

$D_{987465231}$:  or D_{216345} : 

where there's a circled dot SE of some crossed-out dots. So the 132-avoiding permutations are exactly the ones whose dot diagrams consist of unbroken collections of circled dots. In other words, there is some non-negative integer sequence $\lambda_1 \geq \lambda_2 \geq \cdots$ such that for all $i \geq 0$, the ith row of D_w consists of the first λ_i dots in that row:

$$D_w = \{(i,j) \mid 1 \leq j \leq \lambda_i\}.$$

Note that since the number of circled dots in D_w is $|D_w| = \mathrm{inv}(w)$, we have that $\lambda = (\lambda_1, \lambda_2, \ldots)$ is a *partition* of $\mathrm{inv}(w)$ (recall that a partition is a composition whose parts weakly decrease). We call the dot arrangement with λ_i dots in row i, left-aligned, the *Ferrers diagram* for the partition λ.

Now notice that the Ferrers diagrams appearing are not for just any partitions of $0, 1, \ldots, \binom{n}{2}$; in row i, there are at least i dots that are crossed out (there are only $n - i$ numbers bigger than i that could have appeared in rows on and above i). So λ satisfies $\lambda_i \leq n - i$. The result is that the 132-avoiding permutations are in bijection with partitions λ of $0, 1, \ldots, \binom{n}{2}$ satisfying $\lambda_i \leq n - i$. However, in the same fashion as Exercise 14, these are in bijection with elementary lattice paths from $(0,0)$ to (n,n) staying above the line $y = x$. (Another example of a set counted by Catalan numbers!!)

Proposition 5.4 (EC1 Prop. 1.5.1)**.** *Let $\mathcal{S}_{132}(n)$ denote the 132-avoiding permutations of $[n]$. Then*

$$\sum_{w \in \mathcal{S}_{132}(n)} q^{\mathrm{inv}(w)} = \sum_{\lambda} q^{|\lambda|},$$

where λ ranges over partitions $\lambda : \lambda_1 \geq \lambda_2 \cdots \geq 0$ satisfying $\lambda_i \leq n - i$, and where $|\lambda| = \sum_i \lambda_i$.

Exercise 25. (a) Write the permutation matrices for each $w \in \mathcal{S}_3$.

(b) Verify that there are 14 of permutations of 4 that have no decreasing sequences of length 3 by counting the corresponding lattice paths, and listing the corresponding permutations.
(c) For each of the following pairs (u, w), decide whether or not w is u-avoiding.
   (i) $u = 132$, $w = 7421365$;
   (ii) $u = 132$, $w = 5671234$;
   (iii) $u = 1234$, $w = 1765423$;
   (iv) $u = 1234$, $w = 73164258$;
   (v) $u = 1234$, $w = 123$.
(d) Draw the dot diagrams for each $w \in \mathcal{S}_3$.
(e) Inverses of dot diagrams.
   (i) Verify that $D_w^t = D_{w^{-1}}$ for $w \in \mathcal{S}_3$.
   (ii) Explain why $D_w^t = D_{w^{-1}}$.
(f) Inversion tables from dot diagrams.
   (i) Verify that the inversion tables for the permutations $w \in \mathcal{S}_3$ can be read off of the dot diagrams D_w.
   (ii) Explain why the inversion table $I = (a_1, \ldots, a_n)$ corresponding to a permutation w is given by
$$a_i = \#\{ \text{ circled dots in column } i \text{ of } D_w \ \}.$$
(g) Verify that the 132-avoiding permutations of [3] are Ferrers diagrams, and in bijection with the set of partitions of $0, 1, \ldots, \binom{3}{2}$ satisfying $\lambda_i \leq n - i$.
(h) Verify Proposition 1.5.1, i.e. that
$$\sum_{w \in \mathcal{S}_{132}(n)} q^{\mathrm{inv}(w)} = \sum_{\lambda} q^{|\lambda|},$$
for $n = 3$

Exercise 26. Go to

http://www-math.mit.edu/∼rstan/ec/catalan.pdf

and

http://www-math.mit.edu/∼rstan/ec/catadd.pdf

to see the famous Catalan numbers problem (6.19 in EC2). Skim the whole thing. Pick three unrelated parts and verify that the corresponding combinatorial set is of size 14 for $n = 4$. Unrelated means that if you do one part concerning lattice paths, then the other parts you pick should not concern lattice paths. Parts (h) and (i) are equivalent to the lattice paths you did in problem 25(b), so don't repeat these. You may need to look up or ask for definitions, but examples for $n = 3$ are almost always given.

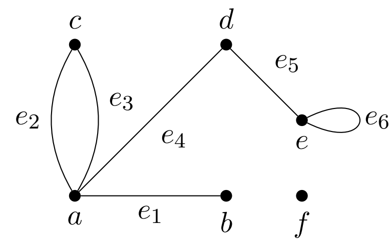
6. Graphs

See the appendix on graph theory terminology. A *graph* is a triple (V, E, ϕ), where V is a se of vertices, E is a set of edges, and ϕ is a function that assigns to each edge $e \in E$ a 2-element multiset of vertices,
$$\phi : E \to \left(\!\!\binom{V}{2}\!\!\right).$$
We draw graphs as vertices represented by points, and edges e represented as simple curves connecting the points in $\phi(e)$. For example, for
$$V = \{a, b, c, d, e, f\}, \qquad E = \{e_1, e_2, e_3, e_4, e_5, e_6\},$$

and

$$\phi : e_1 \mapsto \{a, b\}, \quad e_2 \mapsto \{a, c\}, \quad e_3 \mapsto \{a, c\}, \quad e_4 \mapsto \{a, d\}, \quad e_5 \mapsto \{d, e\}, \quad e_6 \mapsto \{e, e\},$$

we draw



The edge e_6 is called a *loop*. We say a vertex a is *adjacent* to a vertex b if there is an edge mapping to $\{a, b\}$. We say that and edge is *incident* to a vertex if the edge connects to the vertex. If vertices u and v are adjacent, we we say that they are *neighbors*, and that u is in the *neighborhood* $N(v)$ of v (and vice-versa). If $A \subseteq V$, then

$$N(A) = \bigcup_{v \in A} N(v).$$

The *degree* $\deg(v)$ of a vertex v is the number of edge ends attached to v.

Classes of graphs:

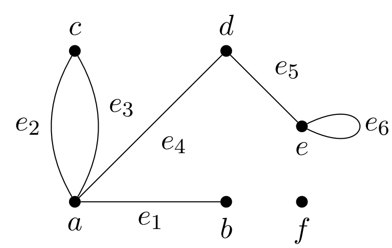
A graph is *simple* if there are no loops and every pair of vertices has at most one edge between them (i.e. ϕ is injective).

A graph is a *multigraph* if there are no loops, but there could be multiple edges between two vertices.

So

$$\{ \text{ graphs } \} \supset \{ \text{ multigraphs } \} \supset \{ \text{ simple graphs } \}.$$

Example 3. Back to the graph



This graph is not simple nor is it a multigraph. The neighborhood $N(a)$ is $\{b, c, d\}$; the neighborhood $N(e)$ is $\{d, e\}$; and the neighborhood $N(\{a, e\}) = \{b, c, d, e\}$. The degrees of the vertices are given by

$$\deg(a) = 4, \quad \deg(b) = 1, \quad \deg(c) = 2, \quad \deg(d) = 2, \quad \deg(e) = 3, \quad \deg(f) = 0.$$

Theorem 6.1 (The handshake theorem and related theorems)**.**

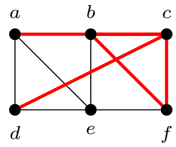
(1) Handshake theorem: $2|E| = \sum_{v \in V} \deg v$

(2) Corollary: There are an even number of odd vertices.

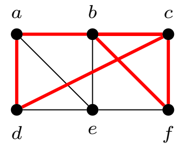
A *walk* of length n from vertex u to vertex v is a sequence $v_0 e_1 v_1 e_2 v_2 \cdots e_n v_n$ such that $v_i \in V$, $e_i \in E$, $v_0 = u$, $v_n = v$, and any two consecutive terms are incident. If G is simple then the sequence $v_0 v_1 \ldots v_n$ of vertices suffices to determine the walk. A walk is *closed* if $v_0 = v_n$, a *trail* if the e_is are distinct, and a *path* if the v_is (and hence the e_is) are distinct. If $n \geq 1$ and all the v_is are distinct except for $v_0 = v_n$, then the walk is called a *cycle*.

We call a walk (resp. path, trail) from u to v a $u - v$ *walk* for short.

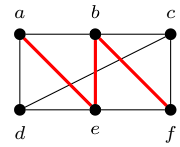
Walk:
a, b, c, f, b, c, d



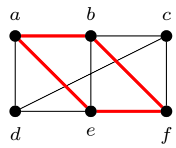
Closed walk:
a, b, c, f, b, c, d, a



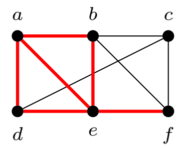
Path:
a, e, b, f



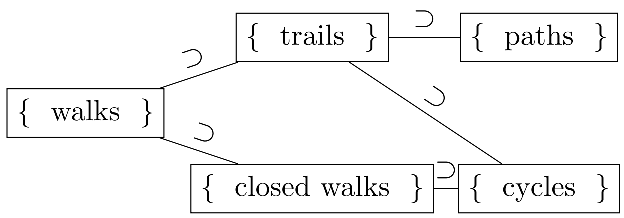
Cycle:
a, b, f, e, a



Trail:
a, e, b, a, d, e, f



Note:

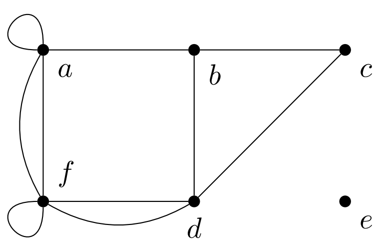


A *maximal path* is a path that cannot be extended on either end to be a longer path.

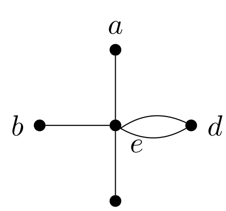
A graph is *connected* if it is nonempty and any two distinct vertices are joined by a walk. A graph without cycles is called a *forest*; a connected graph without cycles is called a *tree*.

Exercise 27. Let

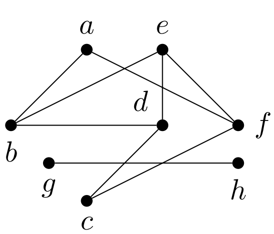
$G_1 =$



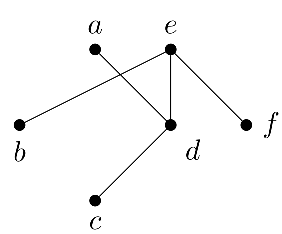
$G_2 =$



$G_3 =$



$G_4 =$

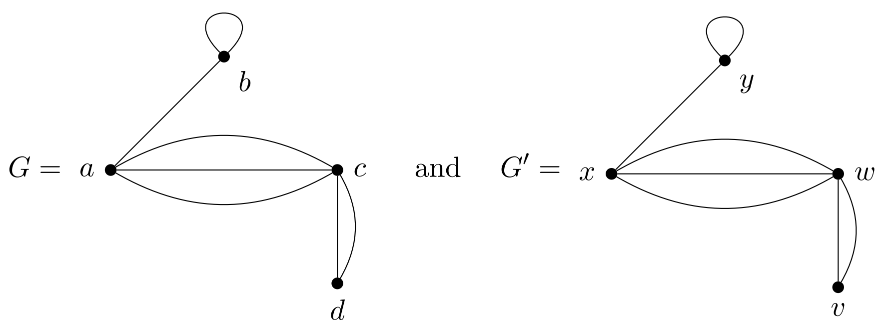


(a) For each of the graphs G_1, G_2, G_3, and G_4, answer the following questions.
   (i) Is G_i simple? Is G_i a multigraph? Is G_i connected? Is G_i a forrest? Is G_i a tree?
   (ii) What is the neighborhood of a? of e? of $\{a, e\}$?
   (iii) What is the degree of a? of e?
   (iv) Verify the handshake theorem.
   (v) Verify that there are an even number of vertices of odd degree.
(b) List three cycles in G_3.
(c) Give an example of a closed walk in G_1 that is not a cycle.
(d) Give an example of a path of length 3 in G_3.
(e) Give an example of a maximal path in G_4.
(f) Give a walk of length 5 in G_2 that is not closed, not a path, and not a trail.
(g) Give a trail of length 8 in G_1.
(h) What's the longest length of the trails in G_2?

6.0.1. *Graph isomorphisms.* We say two graphs G and G' are isomorphic if there is a relabeling of the vertices of G that transforms it into G'. In other words, there is a bijection

$$f : V \to V'$$

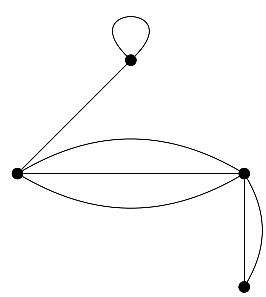
such that the induced map on E is a bijection $f : E \to E'$. For example,



are isomorphic via the map

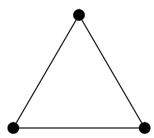
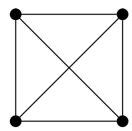
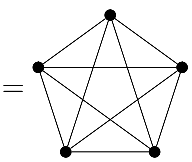
$$a \mapsto x, \quad b \mapsto y, \quad c \mapsto x, \quad d \mapsto v.$$

Recall that an equivalence relation is a pairing $\sim$ that is reflexive, symmetric, and transitive. Then on the set of all graphs, isomorphism is an equivalence relation. An *equivalence class* is the set of all things that are pairwise equivalent. For an equivalence class of graphs, we draw the associated *unlabeled graph.* For example, the equivalence class of graphs corresponding to G above is



6.0.2. *Special graphs.* Recall that a simple graph is a graph with no loops or multiple edges.

Cycles. A *cycle* C_n is the equivalence class of simple graphs on n vertices $\{v_1, v_2, \ldots, v_n\}$ so that v_i is adjacent to $v_{i\pm1}$ (v_1 is adjacent to v_n).
Wheels. A *wheel* W_n is the the cycle C_n together with an additional vertex that is adjacent to every other vertex.
Complete graphs. The *complete graph* on n vertices, denoted K_n, is the simple graph on v vertices so that $N(v) = V - \{v\}$ for all all $v \in V$. For example,

$K_3 =$  $K_4 =$  $K_5 =$ 

Bipartite graphs. A graph is *bipartite* if V can be partitioned into two nonempty subsets V_1 and V_2 so that no vertex in V_i is adjacent to any other vertex in V_i for $i = 1$ or 2.

For example, the *complete bipartite graph* $K_{n,m}$ is the class of simple graphs corresponding to the graph where

$$V_1 = \{v_1, \ldots, v_n\} \quad V_2 = \{u_1, \cdots u_m\}$$
$$N(v_i) = V_2 \qquad \text{and} \qquad N(u_i) = V_1$$

for all i. We call an example of such V_1 and V_2 *vertex classes* of G.

One way to show that a graph is bipartite is to "color" the vertices two different colors, so that no two vertices of the same color are adjacent. A more thorough classification of bipartite graphs is as follows.

Theorem 6.2. *A graph is bipartite if and only if it does not contain an odd cycle as a subgraph.*

Proof. Suppose G is bipartite with vertex classes V_1 and V_2. Let $v_0e_1v_1 \ldots e_nv_n$ be a cycle in G. Assume, without loss of generality, that $v_0 \in V_1$. Then $v_1 \in V_2$, $v_2 \in V_1$, and so on. Since $v_n = v_0$ is in V_1, n must be even.

Now suppose G does not contain an odd cycle. A graph is bipartite if and only if every connected component is bipartite, so we may assume that G is connected; for non-connected graphs, execute the following argument on each component individually.

Pick a vertex $v \in V$, and let $V_1 = \{u \in V \mid d(v,u) \text{ is odd }\}$ and $V_2 = V - V_1$. Now let $v, v' \in V_i$, for $i = 1$ or 2. A shortest path from v to v' is of even length; if there were in addition an edge between v and v', then closing the path from v to v' with that edge would yield an odd cycle. Thus the vertices in V_i are pairwise non-adjacent.

□

Note that the maximum number of edges a graph with vertex classes V_1 and V_2 could have is $|V_1||V_2|$. So a bipartite graph with $|V| = n$ can have no more than

$$\max_{k=1,2,\ldots,n-1} k(n-k) = \lfloor n^2/4 \rfloor$$

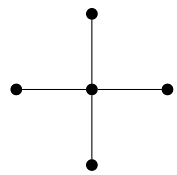
edges. This maximum is attained in $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. By Theorem 6.2, the same can be said for graphs with no odd cycles.
Hypercubes. The n-dimensional hypercube Q_n is the simple graph with vertices

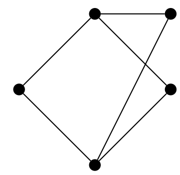
$$V = \{ \text{ bit strings of length } n \ \},$$

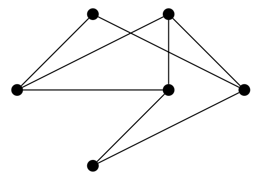
and two vertices are adjacent exactly when the bitstrings differ in exactly one digit.

Exercise 28. (a) Draw C_6, W_6 K_6, and $K_{5,3}$.

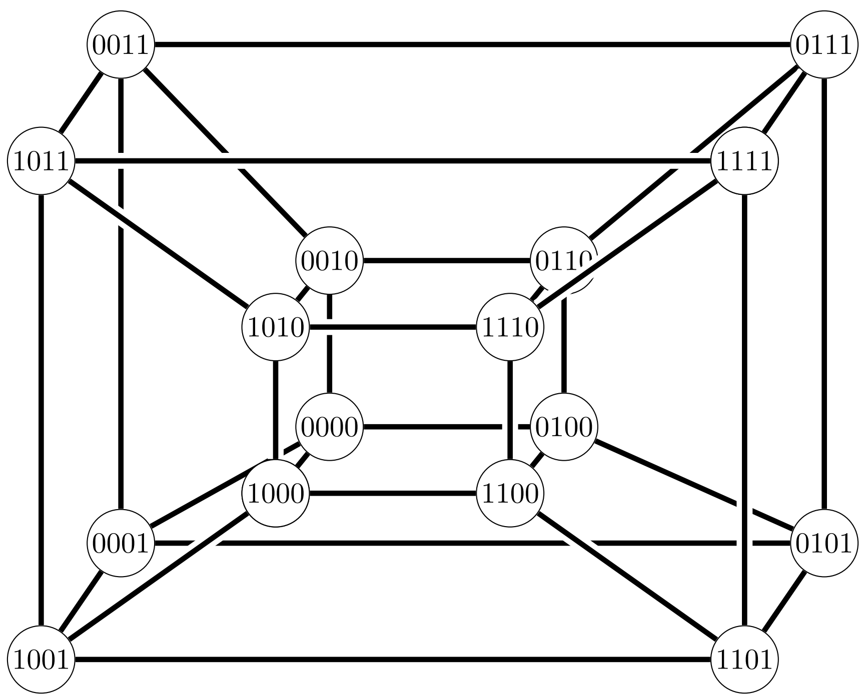
(b) Which of the following are bipartite? Justify your answer.







(c) Hypercubes are bipartite.
   (i) The following is the 4-cube:



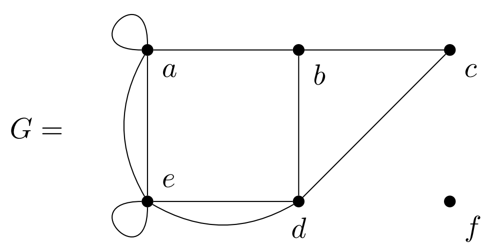
Shade in the vertices that have an even number of 0's. Explain why the 4-cube is bipartite.
   (ii) Explain why Q_n is bipartite in general.
   [Hint: consider the parity of the number of 0's in the label of a vertex.]

6.1. **Graph invariants.** Sometimes it's hard to see whether two graphs are isomorphic or not from their pictures. To show that two graphs are isomorphic, you need to find an isomorphism. To show that they're not isomorphic, you have to show that *no isomorphism exists*, which can be harder. So we look for properties of the graphs that are preserved by isomorphisms. These are called *(graph) invariants*.

Examples of graph invariants:

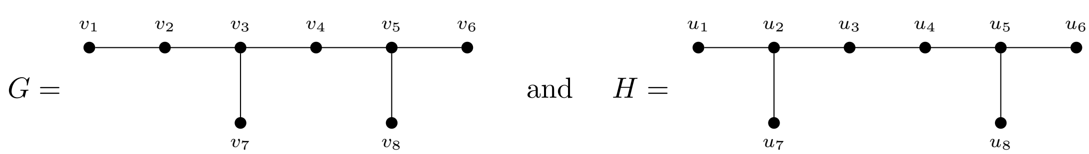
(1) Number of vertices.
(2) Number of edges.

(3) Degree sequences.
The *degree sequence* of a graph is the list of degrees of vertices in the graph, given in decreasing order. For example, the degree sequence of



is 6, 5, 4, 3, 2, 0.

You can get finer invariants by looking closer at the degree sequences. For example, suppose two graphs have the degree sequence 4, 4, 3, 3, 2, and in one of them, the two vertices of degree 4 are adjacent. If the two vertices of degree 4 are not adjacent in the other graph, then the two graphs are not isomorphic. For example:



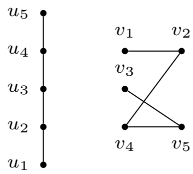
Both of these graphs have the degree sequence $3, 3, 2, 2, 1, 1, 1, 1$. But in G, there's a vertex of degree 1 adjacent to a vertex of degree 2, where as no vertex of degree 1 is adjacent to a vertex of degree 2 in H. So $G \ncong H$.

We call a graph *regular* if its degree sequence is constant (all the vertices have the same degree).

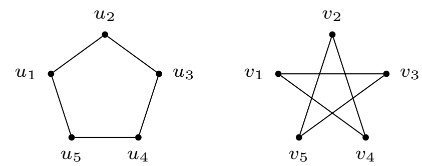
Exercise 29.

(a) For each of the following pairs of graphs, first list their degree sequences. Then decide whether they are isomorphic or not. If not, say why. If they are, give a bijection on the vertices that preserves the edges, and draw the unlabeled graph that represents the corresponding isomorphism class of graphs.

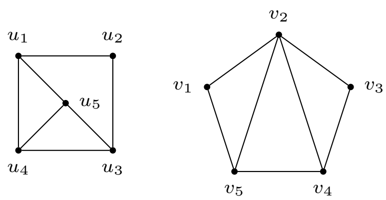
(i)



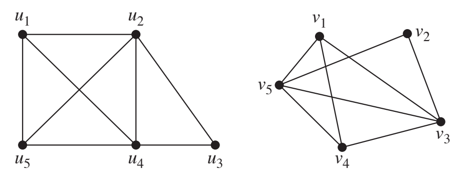
(ii)



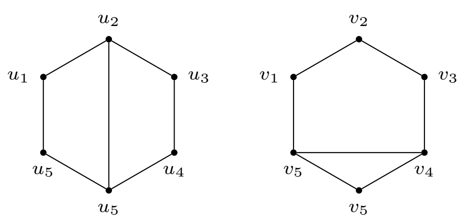
(iii)



(iv)



(v)



(b) How many isomorphism classes are there for graphs with 4 vertices? Draw them.
(c) How many edges does a graph have if its degree sequence is $4, 3, 3, 2, 2$? Draw a graph with this degree sequence. Can you draw a simple graph with this sequence?
(d) For which values of n, m are these graphs regular? What is the degree?

(i) K_n (ii) C_n (iii) W_n (iv) Q_n (v) $K_{m,n}$

(e) How many vertices does a regular graph of degree four with 10 edges have?
(f) Show that every non-increasing finite sequence of nonnegative integers whose terms sum to an even number is the degree sequence of a graph (where loops are allowed). Illustrate your proof on the degree sequence 7,7,6,4,3,2,2,1,0,0. [Hint: Add loops first.]
(g) Show that isomorphism of simple graphs is an equivalence relation.

6.1.1. *New graphs from old.* Let $G = (V, E)$ be a simple graph.
Subgraph. A subgraph of a graph $G = (V, E)$ is a graph $H = (W, F)$ such that $W \subseteq V$ and $F \subseteq E$.
Induced subgraph. Let $W \subseteq V$. The *subgraph induced* by W is $G[W]$, the subgraph of G with vertex set W, and every edge of G connecting two vertices in W. For $v \in V$, the graph $G - v$ is the induced subgraph corresponding to $V - \{v\}$.

By systematically choosing subsets of vertices, we can more easily count the total number of subgraphs by considering the induced subgraphs one at a time. For example...
Edge operations.

(1) Subtraction: Let $e \in E$. Then $G - e$ is the subgraph of G with vertex set V and edge set $E - \{e\}$. If $F \subseteq E$, the graph G is the subgraph with vertex set V and edge set $E - F$.
(2) Addition: For an edge e on the vertex set V but not in E, $G + e$ is the graph containing G satisfying $(G + e) - e = G$.
(3) Contraction: Let $e \in E$. The graph G/e is the graph obtained by *contracting* the edge e, which means merging the vertices that are incident to e and deleting any resulting loops or multiple edges. Note that G/e is not in general a subgraph of G!

Unions. The union of two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ is

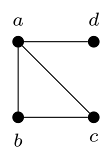
$$G_1 \cup G_2 = (V_1 \cup V_2, E_1 \cup E_2).$$

Complements. Consider G as a subgraph of K_n where $|V| = n$. The complement of the graph G is

$$\bar{G} = (V, E(K_n) - E).$$

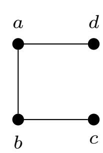
In other words, G and $\bar{G}$ have the same vertex set, but u and v are adjacent in $\bar{G}$ if and only if u and v are not adjacent in G.

Exercise 30. (a) Consider the graph

$G =$ 

(i) Give an example of a subgraph of G that is not induced.
(ii) How many induced subgraphs does G have? List them.
(iii) How many subgraphs does G have?
(iv) Let e be the edge connecting a and d. Draw $G - e$ and G/e.
(v) Let e be the edge connecting a and c. Draw $G - e$ and G/e.
(vi) Let e be an edge connecting d and c. Draw $G + e$.
(vii) Draw $\bar{G}$.

(b) Show that

$G =$ 

is isomorphic to its complement.

(c) Find a simple graph with 5 vertices that is isomorphic to its own complement. (Start with: how many edges must it have?)

Back to connectivity. Recall that a graph $G = (V, E, \phi)$ is connected if and only if every pair of vertices $u, v \in V$ have a walk starting at u and ending at v. In fact, there are several equivalent definitions. To this end, let $d(u, v)$ be the *distance* between u and v in G, given by the length of the shortest walk (which must be a path) between then.

Theorem 6.3. *Let $v \in W \subset V$ be a vertex in a graph G, where W is the vertex set of a single connected component of G. Then the following hold.*

(1) $W = \{y \in V \mid G$ contains a walk from u to v $\}$.
(2) $W = \{y \in V \mid G$ contains a trail from u to v $\}$.

(3) $W = \{y \in V \mid G$ *contains a path from* u *to* v $\}$.
(4) $W = \{y \in V \mid d(u, v) < \infty\}$.

Now suppose G is connected – can we measure *how* connected it is? For example, if G models a network of computer servers, we would like it to be stable with regard to power outages and downed lines. Namely, if one server goes down, we don't want one part of the network to be separated from another.

If the subgraph induced by $V - \{v\}$ is not connected, we call v a *cut vertex*. Similarly, if $G - e$ is not connected, we call e a *cut edge*. Note that if there is no cut-vertex, then there is no cut edge.

Theorem 6.4. *For a simple graph with at least 3 vertices, the following are equivalent.*

(i) G *is connected and contains no cut vertex.*
(ii) Every two vertices in V *are contained in some cycle.*
(iii) Every two edges in E *are contained in some cycle.*
(iv) For any three vertices $u, v, w \in V$*, there is a path from* u *to* v *containing* w.

If $W \subset V$ has the property that $G[V - W]$ is not connected, we call W a *vertex cut.* If $F \subset E$ has the property that $G - E$ is not connected, we say F is an *edge cut.*

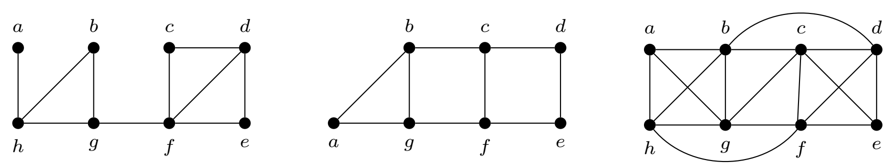
Let $\kappa(G)$ be the fewest number of vertices needed to disconnect a graph. We call $\kappa(G)$ the (vertex) connectivity of G. Note that the complete graph has no vertex cut. We define $\kappa(K_n) = n - 1$. So

$$0 \leq \kappa(G) \leq |V| - 1.$$

The larger the κ, the more connected the graph. We say G is k-connected if $\kappa(G) \geq k$.

Let $\lambda(G)$ be the fewest number of vertices needed to disconnect a graph. We call $\lambda(G)$ the edge connectivity of G. Note that we can disconnect a graph by removing all the edges around a single vertex. So $\lambda(G) \leq \min_{v \in V} \deg(v)$. Moreover, if we remove all the vertices adjacent to v, then v is isolated. In a non-complete graph, we can always find a vertex that is not adjacent to all other vertices. So

$$\kappa(G) \leq \lambda(G) \leq \min_{v \in V} \deg(v).$$



Exercise 31.

(a) Draw the isomorphism classes of connected graphs on 4 vertices, and give the vertex and edge connectivity number for each.
(b) Show that if v is a vertex of odd degree, then there is a path from v to another vertex of odd degree.
(c) Prove that for every simple graph, either G is connected, or $\bar{G}$ is connected.
(d) Recall that $\kappa(G)$ is the vertex connectivity of G and $\lambda(G)$ is the edge connectivity of G. Give examples of graphs for which each of the following are satisfied.
  (i) $\kappa(G) = \lambda(G) < \min_{v \in V} \deg(v)$
  (ii) $\kappa(G) < \lambda(G) = \min_{v \in V} \deg(v)$
  (iii) $\kappa(G) < \lambda(G) < \min_{v \in V} \deg(v)$
  (iv) $\kappa(G) = \lambda(G) = \min_{v \in V} \deg(v)$
(e) For Theorem 6.4, pick any of parts (ii)–(iv) and show that it's equivalent to part (i):

Theorem: For a simple graph with at least 3 vertices, the following are equivalent.

(i) G is connected and contains no cut vertex.
(ii) Every two vertices in V are contained in some cycle.
(iii) Every two edges in E are contained in some cycle, and G contains no isolated vertices.
(iv) For any three vertices $u, v, w \in V$, there is a path from u to v containing w.

6.2. **Trees and forests.** Recall that a graph is a forrest if it contains no cycles, i.e. is *acyclic*. A graph is a tree if it is a connected forrest.

Theorem 6.5. *A simple graph G is a forest if and only if for every pair $u, v \in V$ of distinct vertices, G contains at most one $u - v$ path.*

Proof. If $v_0, v_1, \ldots, v_\ell = v_0$ is a cycle in G, then v_0, v_1 and $v_0, v_{\ell-1}, \ldots, v_1$ are two $v_0 - v_1$ paths.

Conversely, let $P_1 = (u, v_1, \ldots, v_{\ell-1}, v)$ and $P_2 = (u, u_1, \ldots, u_{k-1}, v)$ be two $u - v$ paths. Let i be minimal such that $u_{i+1} \neq v_{i+1}$ and let j be minimal such that $j \geq i$ and u_{j+1} is in P_1, i.e. $u_{j+1} = v_r$. Then $v_i, v_{i+1}, \ldots, v_r, u_j, u_{j-1}, \ldots, u_{i+1}, u_i = v_i$ is a cycle in G. $\square$

Theorem 6.6. *The following are equivalent for a simple graph G.*

(1) G is a tree.
(2) G is a minimal connected graph*, i.e. every edge in E is a cut edge.*
(3) G is a maximally acyclic graph, i.e. G is acyclic, and adding any edge between two non-adjacent vertices creates a cycle.

Proof. Suppose G is a tree, and $e \in E$ has endpoints u and v. Then u, v is one $u - v$ path in G. However, if $G - e$ was connected, then there would be a second $u - v$ path in G, contradicting the previous theorem. So e is cut.

Complete on the homework.

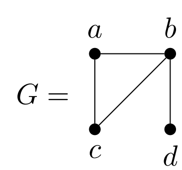
$\square$

Exercise 32. Complete the proof of the following theorem:
The following are equivalent for a simple graph G.

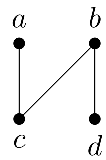
(1) G is a tree.
(2) G is a *minimal connected graph*, i.e. every edge in E is a cut edge.
(3) G is a maximally acyclic graph, i.e. G is acyclic, and adding any edge between two non-adjacent vertices creates a cycle.

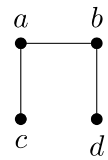
6.2.1. *Spanning trees.* A *spanning subgraph* of $G = (V, E, \phi)$ is a subgraph H with vertex set V. A *spanning tree* of a graph G is a spanning subgraph that is a tree. Every connected graph G has at least one spanning tree, because we can take a minimal connected spanning subgraph by deleting edges until we can't anymore.

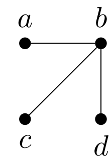
For example, the graph



has exactly three spanning trees:

$T_1 =$ 

, $T_2 =$ 

, and $T_3 =$ 

.

Constructing and counting spanning trees. Let G be a simple graph.

<u>Method 1:</u> Pick a vertex $v \in V$ and set $V_i = \{u \in V \mid d(u,v) = i\}$, for $i = 0, 1, \dots$. Note that if $u_i \in V_i$, $i > 0$, and $v, z_1, z_2, \dots, z_{i-1}, u_i$ is a $v - u$ path, then $d(v, z_j) = j$ for each j. In particular, $V_j \neq \emptyset$. Moreover, for every $u_i \in V_i$, u_i is adjacent to some u_{i-1} in V_{i-1}. For each $u \in V_i$, pick a unique $u' \in V_{i-1}$ that is adjacent to u. Let T be the subgraph of G with vertex set V and edges given by including exactly the edges between corresponding u, u' pairs. Then T is spanning, and for each $u \in V$, u is in some V_i, as so has a neighbor which is closer to v that u, which has a neighbor that is closer, and so on. So every vertex is connected to v, and therefore by symmetry and transitivity, T is connected. Moreover, for any $W \subseteq V$, if w is a maximally distant vertex from v, then w has exactly one neighbor in v. SO any path from v through w can only move farther away from v in G.

Note that if $k = \max_{u \in V} d(u,v)$, then $V_i \neq \emptyset$ for $0 \leq i \leq k$, and $V = \bigsqcup_{i=0,\dots,k} V_i$. As an aside, the *diameter* of a graph is

$$\operatorname{diam}(G) = \max_{u,v \in V} d(u,v),$$

and the radius of a graph is

$$\operatorname{rad} = \min_{v \in V} \max_{u \in V} \operatorname{diam}(u,v).$$

For any $v \in V$ with $\max_{u \in V} d(u,v) = \operatorname{rad}(G)$, then the resulting spanning tree using this method has radius $\operatorname{rad}(G)$.

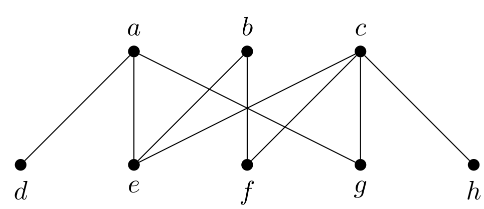
<u>Method 2:</u> Pick a vertex $v \in V$. Let T_1 be the subgraph of G that is just v as an isolated vertex. Construct $T_2, T_3, \dots$, inductively as follows. If $k < n = |V|$, then by the connectedness of G, there must be a vertex $u \in V - V(T_k)$ that is adjacent in G to some vertex $w \in V(T_k)$. Let T_{k+1} be obtained by adding u to T_k along with the edge between u and w. Then similarly as before, T_{k+1} is acyclic and connected, and T_n is a spanning tree.

The spanning trees constructed by either of these two methods have n vertices and $n-1$ edges. In the first construction, there is a bijection between $V - \langle v \}$ and $E(T)$, given by u maps to the edge uu'. In the second method, $|E(T_k)| = k - 1$. But each tree has exactly one spanning tree, so we have the following corollary.

Corollary 6.7. *A connected simple graph with n vertices is a tree if and only if it has $n-1$ edges.*

Exercise 33. (a) How many spanning trees does C_5 have?

(b) Let

$G =$ 

(i) Calculate $\operatorname{diam}(G)$ and $\operatorname{rad}(G)$.

(ii) Fix the vertex a, and give $V_i = \{u \in V \mid d(u, a) = i\}$.

(iii) Build a spanning tree using Method 1 from the notes using $v = a$ (show your steps!). What is the radius of the resulting tree?

(iv) Find a central vertex v, i.e. one for which $\max_{u \in V} d(u, v) = \mathrm{rad}(G)$, and build a spanning tree using Method 1 from the notes using that vertex (show your steps!). What is the radius of the resulting tree?

(v) Build a spanning tree using Method 2 from the notes starting with T_1 being the isolated vertex a (show your steps!).

For a connected graph G, let $t(G)$ be the number of spanning trees in G. Notice that for any fixed edge, you can split the spanning trees into two categories: those that contain e and those that don't. Every spanning tree of G that doesn't contain e is also a spanning tree of $G - e$, so

$$|\{ \text{ spanning trees of } G \text{ not containing edge } e \ \}| = t(G - e).$$

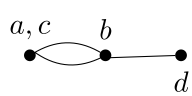
Now let's redefine G/e for the purpose of counting trees. Let G/e be the graph gotten by glueing the endpoints of e and deleting e **but not deleting any remaining loops or multiple edges!** Then

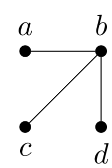
$$|\{ \text{ spanning trees of } G \text{ containing edge } e \ \}| = t(G/e).$$

So

$$t(G) = t(G - e) + t(G/e).$$

For example, if e is the edge joining a and c above,

$G/e =$ 

and $G - e =$ 

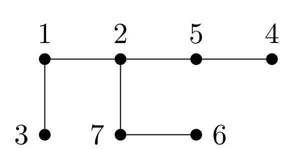
.

The two spanning trees of G/e correspond to T_1 and T_2 above, and the one spanning tree of $G - e$ corresponds to T_3 above.

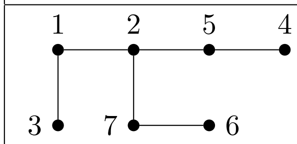
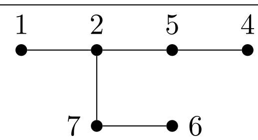
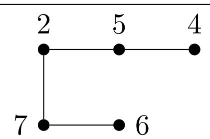
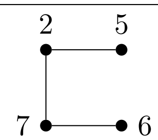
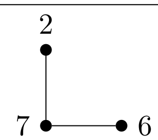
6.3. **Prüfer code.** To every labeled tree for n vertices, there is an associated sequence of numbers between 1 and n of length $n - 2$, called a *Prüfer code*. Also, for every Prüfer code (sequence of length $n - 2$ of numbers between 1 and n) there is a tree. You get them as follows:

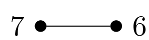
Prüfer code from tree: Remove the lowest leaf possible and record the value of its neighbor. Iterate until there are two leaves left. Your code should have $|V| - 2$ numbers.

For example, let

$T =$ 

Removing the lowest leaf, one-by-one, and recording the neighbors, looks like

remaining tree:	lowest leaf:	its neighbor:	new code:
	3	1	1
	1	2	$1, 2$
	4	5	$1, 2, 5$
	5	2	$1, 2, 5, 2$
	2	7	$1, 2, 5, 2, 7$

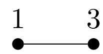
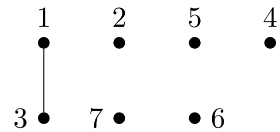
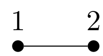
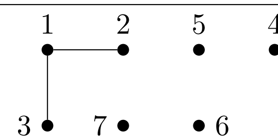
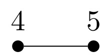
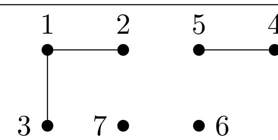
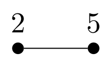
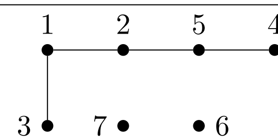
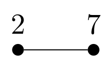
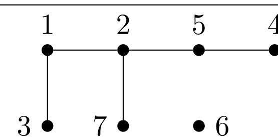
Now removing vertex 2 leaves you with  , so you're done. Thus the Prüfer code for T is $1, 2, 5, 2, 7$.

You can also reverse this process as follows:

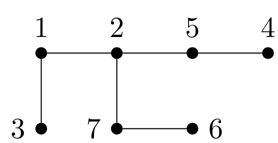
Tree from Prüfer code: At the very beginning, draw a bar (|) at the end of your code of length $n - 2$. Also draw n vertices, labeled from 1 to n.

Now, take the the first number, say i, from your code and identify the least number from 1 to n missing from your code, say j. Draw an edge from i to j, drop i from the front of your code, and add j to the end (after the bar). Iterate until you've cycled the bar to the front. Then draw and edge between the two numbers that are missing from your code.

For example, starting with the Prüfer code $1, 2, 5, 2, 7$, this process goes as follows:

new code:	lowest number missing:	new edge:	tree so far:
$1, 2, 5, 2, 7 \mid$	3		
$2, 5, 2, 7 \mid 3$	1		
$5, 2, 7 \mid 3, 1$	4		
$2, 7 \mid 3, 1, 4$	5		
$7 \mid 3, 1, 4, 5$	2		

After this last step, the code is $\mid 3, 1, 4, 5, 2$, which is totally cycled. So we look for the last two missing numbers, which are 6 and 7, so we finish the graph off with an edge between 6 and 7:



,

and we're done. Notice that we got back to the tree we started with.

It turns out that these two processes are inverses of each other. What's more, is the following:

Theorem 6.8. *For each n, there is a bijection between Prüfer codes, i.e. sequences of numbers between 1 and n of length $n-2$, and labeled trees on n vertices.*

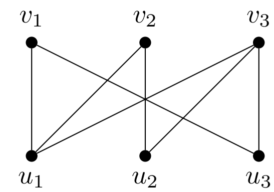
Since there are exactly n^{n-2} Prüfer codes of length $n-2$, there must be exactly n^{n-2} labeled trees on n vertices. This is known as Cayley's formula:

Theorem 6.9 (Cayley's formula)**.** *There are n^{n-2} labeled trees on n vertices.*

Since every labeled tree with n vertices is a spanning tree of (the labeled) K_n, and vice versa, we get the following corollary.

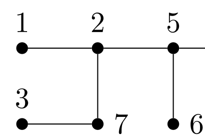
Corollary 6.10. *There are n^{n-2} spanning trees in K_n.*

Exercise 34. (a) Use the recurrence relation $t(G) = t(G - e) + t(G/e)$ to count the number of spanning trees of



Remember to keep multiple edges!!

(b) What is the Prüfer code for the following labeled tree?



Check your answer by reversing the process and building the tree from the code.

(c) Draw the tree whose Prüfer code is $2, 2, 5, 3, 6$. Check your answer by calculating the Prüfer code that goes with your tree.

(d) Draw a labeled K_3 (labeled with $1, 2, 3$), and list all the spanning trees, and the corresponding Prüfer code. Verify that there is a bijection between the labeled trees on 3 vertices and the length-1 Prüfer codes.

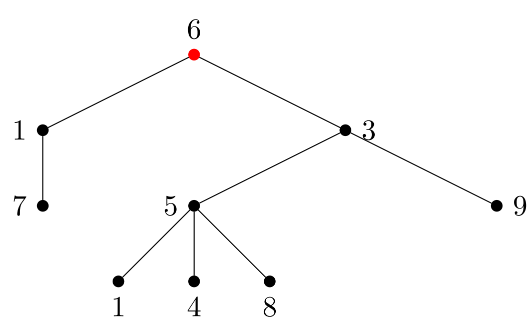
(e) How many spanning trees does K_7 have?

(f) How many labeled trees are there on 14 vertices?

6.4. **Rooted trees.** A *rooted tree* T is a tree with a designated choice of vertex, called the *root*. Note that in EC1, a "tree" is actually a rooted tree, and our notion of tree is called a "free tree". In the appendix on graph theory, a (rooted) tree is equivalently defined as follows: a (rooted) tree T is a finite set of vertices such that:

(1) One specially designated vertex is called the root of T, and
(2) The remaining vertices (excluding the root) are partitioned into $m \geq 0$ disjoint nonempty sets $T_1, \ldots, T_m$, each of which is a tree. The trees $T_1, \ldots, T_m$ are called subtrees of the root.

For example, we would usually say for the rooted tree



where 6 is the root, that $V(T) = [9]$. But in but in this alternate definition, we say $T = [9]$, with root 6 and subtrees T_1, T_2; the subtree T_1 has vertices $\{2, 7\}$ and root 2; the subtree T_2 has vertices

$\{1, 3, 4, 5, 8, 9\}$, root 3, and subtrees T_3, T_4; the subtree T_3 has vertices $\{1, 4, 5, 8\}$, root 5, and subtrees T_5, T_6, T_7 consisting of $\{1\}$, $\{4\}$, $\{8\}$, respectively, while T_4 consists of the single vertex 9.

By convention, we draw the root at the top. Then draw successive neighbors below, so that you end up with a *leveled graph*, with

$$V_i = \{u \in V(T) \mid d(u, \text{root}) = i\}$$

appearing on level i. Call the neighbors of $u \in V$ below u the *children* or *successors* of u and the (unique) neighbor above u the *parent* or predecessor of u. In our example, 5 has successors $1, 4, 8$ and parent 3. The leaves of T are the vertices with no successors.

Recall that a partially ordered set, or *poset*, P is a set together with a binary relation $\leq$ reflexive, antisymmetric, and transitive:

(1) $a \leq a$ (reflexivity);
(2) if $a \leq b$ and $b \leq a$, then $a = b$ (antisymmetry);
(3) if $a \leq b$ and $b \leq c$, then $a \leq c$ (transitivity).

As opposed to a *totally ordered set*, not all elements in a partially ordered set are pairwise comparable. One example of a poset is the poset of subsets of $[n]$ with order given by containment:

$$\begin{array}{ccc} & \{1,2\} & \\ \diagup & & \diagdown \\ \{1\} & & \{2\} \\ \diagdown & & \diagup \\ & \emptyset & \end{array}$$

A tree then defines a poset: we say $a \geq b$ if there is a downward moving $a - b$ path. As a note, the picture of the tree is the *Hasse* diagram of the corresponding poset: the elements of P drawn as vertices with an edge between a and b, a above b, whenever $a > b$ and no c such that $a > c > b$.

A *plane tree* or *ordered tree* has an order on the immediate subtrees of T, usually drawn left to right. An m-ary tree is one for which the number of subtrees is m, though those subtrees can be empty. They are drawn with equal angles between the children, so that the empty trees are visible as missing vertices. See Figures 4.34 and 4.35 in EC1. A 2-ary tree is called a binary tree. An m-ary tree is *complete* if no subtree is empty. In Figure 4.35 of EC1, only the first tree is complete.

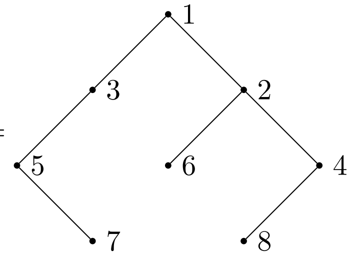
The length $\ell(T)$ of a tree T is equal to the length of the longest path away from the root. This is incidentally also its length as a poset. The complete m-ary tree of length ℓ is the unique (up to isomorphism) complete m-ary tree with every maximal chain of length ℓ; it has a total of $1 + m + m^2 + \cdots + m^\ell$ vertices.

A rooted tree (with or without additional structure, such as being a plane tree or binary tree) on a linearly ordered vertex set (such as $[n]$) is *increasing* if every path from the root is increasing.

6.4.1. *Permutations as rooted trees.* This is the end of section 1.5 in EC1.

Let $w = w_1 w_2 \cdots w_n$ be any word on the alphabet ordered P with no repeated letters (a permutation on $\{w_1, w_2, \ldots, w_n\} = \{p_1, p_2, \ldots, p_n\}_\leq$). Define a plane binary tree $T(w)$ as follows. If $w = \emptyset$, then $T(w) = \emptyset$. If $w \neq \emptyset$, then let i be the least element (letter) of w. Thus w can be factored uniquely in the form $w = uiv$. Now let i be the root of $T(w)$, and let $T(u)$ and $T(v)$ be the left and right subtrees of i. This procedure yields an inductive definition of $T(w)$. The left successor of a vertex j is the least element k to the left of j in w such that all elements of w between k and j (inclusive) are $\geq j$, and similarly for the right successor.

For example,

$T(57316284) =$ 

The correspondence w?? $\mapsto T(w)$ is a bijection between S_n and increasing binary trees on n vertices; that is, binary trees with n vertices labelled $1, 2, \ldots, n$ such that the labels along any path from the root are increasing. To obtain w from $T(w)$, read the labels of w in *symmetric order*: first the labels of the left subtree (in symmetric order, recursively), then the label of the root, and then the labels of the right subtree.

Let $w = w_1 w_2 \cdots w_n \in \mathcal{S}_n$. Define w_i to be

a double rise or double ascent	if $w_{i-1} < w_i < w_{i+1}$
a double fall or double descent	if $w_{i-1} > w_i > w_{i+1}$
a peak	if $w_{i-1} < w_i > w_{i+1}$
a valley	if $w_{i-1} > w_i < w_{i+1}$

where we set $w_0 = w_{n+1} = 0$. It is easily seen that the property listed below of an element i of w corresponds to the given property of the vertex i of $T(w)$:

Element i of w	Vertex i of $T(w)$ has precisely the successors below
double rise	right
double fall	left
valley	left and right
peak	none

Proposition 6.11. *(a) The number of increasing binary trees with n vertices is $n!$.*
(b) The number of such trees for which exactly k vertices have left successors is the Eulerian number $A(n, k+1)$.
(c) The number of complete (i.e. every vertex is either an endpoint or has two successors) increasing binary trees with $2n+1$ vertices is equal to the number E_{2n+1} of alternating permutations ($w_1 > w_2 < w_3 > w_4 < \cdots$) in S_{2n+1}.

Given $w = w_1 w_2 \cdots w_n \in S_n$, construct an (unordered) tree T?(w) with vertices $0, 1, \ldots, n$ by defining vertex i to be the successor of the rightmost element j of w which precedes i and which is less than i. If there is no such element j, then let i be the successor of the root 0. See Figure 1.9 of EC1.

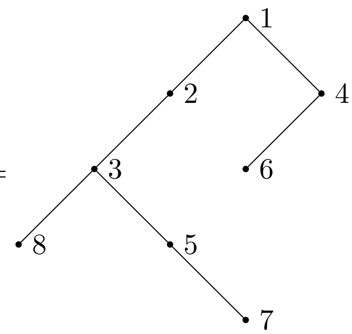
The correspondence w?? $\mapsto T$?(w) is a bijection between $\mathcal{S}_n$ and increasing trees on $n+1$ vertices. The successors of 0 are just the left-to-right minima (or retreating elements) of w (i.e., elements w_i such that $w_i < w_j$ for all $j < i$). Moreover, the leaves of T?(w) are just the elements w_i for which $i \in D(w)$ or $i = n$. Thus in analogy to Proposition 1.5.3 (using Proposition 1.3.1 and the symmetry between left-to-right maxima and left-to-right minima) we obtain the following result.

Proposition 6.12. *(a) The number of unordered increasing trees on $n+1$ vertices is $n!$.*
(b) The number of such trees for which the root has k successors is the signless Stirling number $c(n, k)$.
(c) The number of such trees with k endpoints is the Eulerian number $A(n, k)$.

Exercise 35. (a) For the following permutations w, draw $T(w)$ and $T'(w)$. Verify that $T(w)$ and $T'(w)$ are increasing.

(i) $w = 68412537$
(ii) $w = 12345$
(iii) $w = 54321$

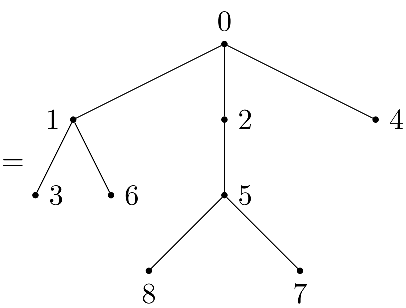
(b) Let

$T =$ 

.

If $T = T(w)$, what is w? Verify that the double rises, double descents, valleys, and peaks of w correspond to the correct behavior of successors.

(c) Let

$T =$ 

.

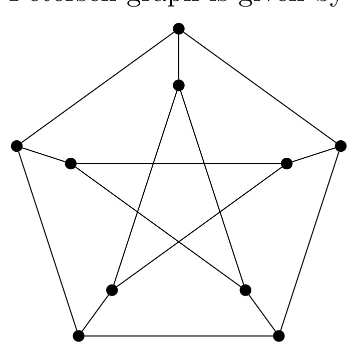
If $T = T'(w)$, what is w? Verify the following for $T'(w)$.

(i) The successors of 0 are just the left-to-right minima of w.
(ii) The leaves are $D(w) \cup \{n\}$.

(d) Briefly justify each of (a)–(c) in Proposition 1.5.3 of EC1.
(e) Briefly justify each of (a)–(c) in Proposition 1.5.5 of EC1.

6.5. **Graph colorings.** This is section 10.8 in Rosen.

Before we start into graph coloring, let's introduce another special graph (to add to paths, cycles, wheels, complete graphs, etc). The Petersen graph is given by



A *coloring* of a simple graph is the assignment of a color to each vertex of the graph so that no two adjacent vertices are assigned the same color. The *chromatic number* of a graph G, denoted $\chi(G)$, is the least number of colors needed for a coloring of this graph.

Example 4. The chromatic number of the Petersen graph is 3.

Example 5. $\chi(K_n) = n$.

Greedy algorithm. We can always find some coloring of a graph by ordering the vertices, ordering the colors, and then one-by-one coloring each vertex with the lowest coloring available.

Notice a graph coloring restricts to a proper coloring of any subgraph. So if G contains K_n as a subgraph, $\chi(G) \geq n$! Let $\omega(G)$ be the largest n for which K_n is a subgraph of G. We call this the *clique number* of G. So

$$\chi(G) \geq \omega(G).$$

Any subgraph of G isomorphic to K_n is called a *clique of size n*.

Now can we think of graphs where $\chi(G) = \omega(G)$ and $\chi(G) > \omega(G)$? i.e. are these really different graph invariants? Absolutely! Take a complete graph K_n. First, this has $\omega(K_n) = \chi(K_n)$. Maybe this is cheating – maybe the complete graph is a special case? Well, no: extend K_n by adding leaves or paths extending from any of the vertices. This preserves both the chromatic number and the clique number.

Now, think about an odd cycle: it is not bipartite, so $\chi(C_{2n+1}) > 2$. But for $n > 1$, there are no cliques of size more than 2! So $\chi(C_{2n+1}) > \omega(G) = 2$.

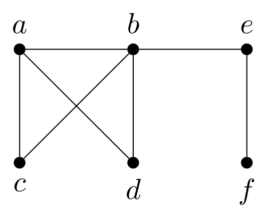
Similarly, we can define the *independence number* $\alpha(G)$ of a graph by the size of the largest collection of vertices that are pairwise non-adjacent (the largest *independent set*). Notice $\alpha(G) = \omega(\bar{G})$. Also, notice that in a coloring of a graph, the vertices of any given color are all independent. So

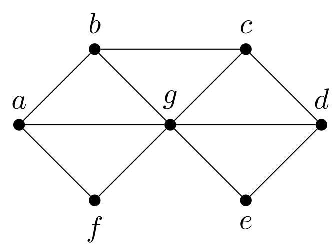
$$\chi(G) \geq |V|/\alpha(G).$$

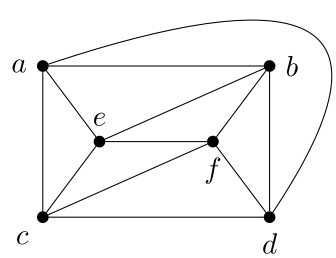
Putting this together with the inequality from before,

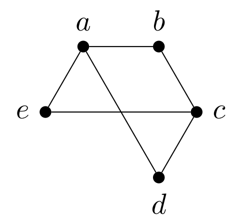
$$\chi(G) \geq \max(\omega(G), |V|/\alpha(G)).$$

Exercise 36. Let A, B, C, and D be the graphs

$A =$ 

$B =$ 

$C =$ 

$D =$ 

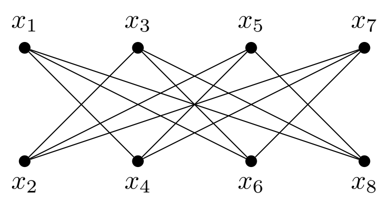
(a) Calculate the chromatic numbers for A, B, C, and D. For each, give an example of a vertex coloring of the corresponding graph using exactly χ colors.
(b) What are the clique and independence numbers of A, B, C, and D? How do ω and $|V|/\alpha$ compare to χ for each graph?
(c) What are the chromatic numbers of

(i) $K_{m,n}$, (ii) C_n, (iii) W_n, and (iv) Q_n?

(d) What are the clique and independence numbers of

(i) $K_{m,n}$, (ii) C_n, (iii) W_n, (iv) Q_n?

How do ω and $|V|/\alpha$ compare to χ for each graph? (You may need to break into cases.)
(e) What are necessary and sufficient conditions for a graph to have chromatic number (i) 1, and (ii) 2.
(f) For A, B, D, and H, pick an ordering of the vertices and color the graph using the greedy algorithm. Compare the colors you used to the chromatic number of the graph. Can you find an ordering that yields a coloring with too many colors?
(g) Explain why the clique number of the complement of a bipartite is no smaller than the number of vertices in each part. (Recall that the *parts* of a bipartite graph are the two collections of pairwise non-adjacent vertices.)
(h) Notice that the graph

$G =$ 

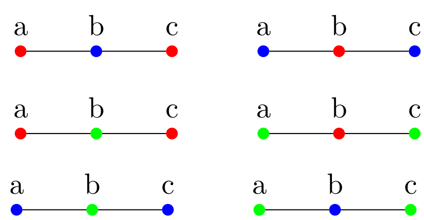
is bipartite, so should have chromatic number 2. Now color this graph using the greedy algorithm. How many colors did you need?

6.5.1. *The chromatic polynomial.* The *chromatic* polynomial, $\chi(G,t)$ counts the number of proper vertex colorings using t colors or fewer colors.

For example, the path graph P_3 on 3 vertices cannot be colored at all with 0 or 1 colors. With 2 colors, it can be colored in 2 ways:

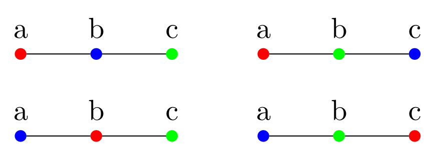
.

With 3 colors, it can be colored in 12 ways (there are $2 * \binom{3}{2} = 6$ with exactly 3 colors:



,

and $3! = 6$ with exactly 3 colors:



).

So $\chi(P_3, t)$ is a degree 3 polynomial, i.e.

$$\chi(P_3, t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

satisfying $\chi(P_3, 0) = \chi(P_3, 1) = 0$, $\chi(P_3, 2) = 2$, and $\chi(P_3, 3) = 12$. So we need to solve

$$\begin{aligned}
0 =& \chi(P_3, 0) = a_0\\
0 =& \chi(P_3, 1) = a_0 + a_1 + a_2 + a_3\\
2 =& \chi(P_3, 2) = a_0 + 2a_1 + 4a_2 + 8a_3\\
12 =& \chi(P_3, 3) = a_0 + 3a_1 + 9a_2 + 27a_3
\end{aligned}$$

Solving this system gives

$$a_0 = 0, \qquad a_1 = 1, \qquad a_2 = -2, \qquad a_3 = 1.$$

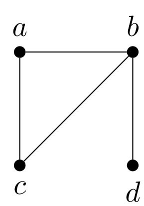
So

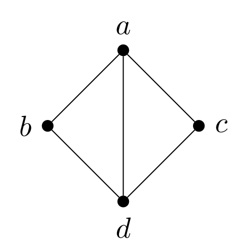
$$\chi(P_3, t) = t - 2t^2 + t^3 = t(t-1)^2.$$

For more notes, see `http://en.wikipedia.org/wiki/Chromatic_polynomial`.
Some properties of the chromatic polynomial that can help you error check:
If G has n vertices, m edges, and connected components $C_1, \ldots, C_\ell$, then...

(1) For each $0 \leq i < \chi(G)$, $t - i$ is a factor of $\chi(G, t)$.
(2) The coefficient of t^n in $\chi(G, t)$ is 1.
(3) The coefficient of t^{n-1} in $\chi(G, t)$ is $-m$.
(4) The coefficients alternate in signs.
(5) $\chi(G, t) = \chi(C_1, t)\chi(C_2, t) \cdots \chi(C_\ell, t)$

Exercise 37. (a) Compute the number of ways to color the graph with each of 1, 2, ..., $|V|$ colors for

$G =$ 

and
$H =$ 

(b) Calculate the chromatic polynomial for G above.
(c) The chromatic polynomial for the cycle C_n is $\chi(C_n, k) = (k-1)^n + (-1)^n(k-1)$.
  (i) Draw all the ways of coloring the 3-cycle with 3 colors. Then compute $\chi(C_3, 3)$ and compare your answers.
  (ii) How many ways are there to color the 5-cycle with 3 colors?
  (iii) How many ways are there to color the 6-cycle with 2 colors?
  (iv) Use $\chi(C_n, k)$ to verify that even cycles are bipartite and odd cycles are not.

7. Inclusion/Exclusion

We will attempt to tackle the following sorts of questions:

(1) In box of cookies, there are 10 with chocolate chips, 15 with nuts, 20 with oatmeal, and 5 with none of the above. There are 6 cookies with both chocolate chips and nuts, 10 with nuts and oatmeal, 2 with oatmeal and chocolate chips, and 1 with all three. How many cookies are there in the box?
(2) A permutation $w \in \mathcal{S}_n$ is a *derangement* if $w(i) \neq i$ for all $i \in [n]$. How many derangements are there in $\mathcal{S}_n$?
(3) Let $S \subset [n-1]$. How many permutations are there with descent set $D(w) = S$? How many permutations are there with descent set $D(w) \subseteq S$?

Venn diagrams. Basic inclusion/exclusion.

Theorem 7.1. *Let*

$$A_1, A_2, \ldots, A_n \subseteq S \quad \text{so that } \bigcup_{i=1}^{n} A_i = S.$$

Then

$$\begin{aligned} |S| &= \sum_{1 \leq i \leq n} |A_i| - \sum_{1 \leq < i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq < i < j < k \leq n} |A_i \cap A_j \cap A_k| - \cdots \\ &= \sum_{\substack{T \subseteq [n] \\ T \neq \emptyset}} (-1)^{|T|+1} \left| \bigcap_{i \in T} A_i \right|. \end{aligned}$$

For example, the answer to the first question above is

$$5 + ((10 + 15 + 20) - (6 + 10 + 2) + (1)) = 33.$$

Now let's count derangements. Instead of counting the number of permutations with some number of fixed points. For each $i \in [n]$, let

$$A_i = \{w \in \mathcal{S}_n \mid w(i) = i\},$$

so that, for example, if $n = 4$

$$A_1 \cap A_3 = \{w \in \mathcal{S}_4 \mid w(1) = 1, w(3) = 3\} = \{1234, 1423\}.$$

Then the derangements $D(n)$ of $\mathcal{S}_n$ are

$$D(n) = \mathcal{S}_n - \cup_{i=1}^{n} A_i,$$

since if w is not a derangement, then w is in at least one of the A_i. Thus, by the inclusion/exclusion principle, the number of permutations that are not derangements is

$$\sum_{\substack{T \subseteq [n] \\ T \neq \emptyset}} (-1)^{|T|+1} |A_T|, \quad \text{where} \quad A_T = \bigcap_{i \in T} A_i = \{w \in \mathcal{S}_n \mid w(i) = i \forall i \in T\}.$$

But the permutations that fix a subset $T \subseteq [n]$ are in bijection with the permutations of $[n] - T$, so $|A_T| = (n - |T|)!$. Thus

$$\begin{aligned}
D(n) &= n! - \sum_{\substack{T \subseteq [n] \\ T \neq \emptyset}} (-1)^{|T|+1} |A_T| = n! - (-1) \sum_{\substack{T \subseteq [n] \\ T \neq \emptyset}} (-1)^{|T|} (n - |T|)! \\
&= n! + \sum_{k=1}^{n} (-1)^k \binom{n}{k} (n-k)! = n! + \sum_{\ell=0}^{n-1} (-1)^{n-\ell} \binom{n}{n-\ell} \ell! \\
&= \sum_{\ell=0}^{n} (-1)^{n-\ell} \binom{n}{\ell} \ell!.
\end{aligned}$$

Recall that

$$e^x = \sum_{i \in \mathbb{N}} x^i / i!, \quad \text{so that} \quad 1/e = e^{-1} = \sum_{i \in \mathbb{N}} (-1)^i / i!,$$

and

$$\binom{n}{\ell} \ell! = \frac{n!}{\ell!(n-\ell)!} \ell! = \frac{n!}{(n-\ell)!},$$

so that

$$D(n) = n!(1 - 1/1! + 1/2! - 1/3! + \cdots + (-1)^n 1/n!) \qquad \text{is approximately } n!/e \text{ for large } n.$$

We can also conclude from the formula $D(n) = n! \sum_{i=0}^{n} (-1)^i / i!$ the recursion formulas

$$\begin{aligned}
D(n) &= nD(n-1) + (-1)^n \\
D(n) &= (n-1)(D(n-1) + D(n-2)).
\end{aligned}$$

7.1. **The general principle of inclusion/exclusion.** Let S be a set of size n, and let K be a field (think $\mathbb{C}$ or $\mathbb{R}$ or $\mathbb{Q}$). Consider

$$\mathrm{Hom}(2^S, K) = \{f : 2^S \to K\}$$

the set of functions f from the power set of S to the field K. This set is itself a vector space: for $f, g \in \mathrm{Hom}(2^S, K)$, $T \in 2^S$, and $c \in K$, the maps

$$f + g : T \mapsto f(T) + g(T) \qquad \text{and } cf : T \mapsto cf(T)$$

are also in $\mathrm{Hom}(2^S, K)$. Since $|2^S| = 2^n$, $\mathrm{Hom}(2^S, K)$ has dimension 2^n.

Theorem 7.2 (Thm. 2.1.1, EC1)**.** *As above, let S be a finite set, let K be a field, and let $V = \mathrm{Hom}(2^S, K)$ be the vector space of functions from the power set of S to the field K. Now define*

$$\phi : V \to V \qquad \textit{by} \qquad \phi(f) : T \mapsto \sum_{U \supseteq T} f(U) \quad \textit{for all } T \in 2^S.$$

Then ϕ^{-1} exists and is given by

$$\phi^{-1}(f) : T \mapsto \sum_{U \supseteq T} (-1)^{|U-T|} f(U) \quad \textit{for all } T \in 2^S.$$

Proof. Define $\psi : V \to V$

$$\phi^{-1}(f) : T \mapsto \sum_{U \supseteq T} (-1)^{|U-T|} f(U).$$

Then

$$\begin{aligned}\psi \circ \phi \circ f(T) &= \sum_{U \supseteq T} (-1)^{|U-T|} \phi(f)(U) \\ &= \sum_{U \supseteq T} (-1)^{|U-T|} \sum_{V \supseteq U} f(V) \\ &= \sum_{V \supseteq T} \left(\sum_{V \supseteq U \supseteq T} (-1)^{|U-T|} \right) f(V).\end{aligned}$$

Setting $m = |V - T|$, we have, for fixed V and T,

$$\left(\sum_{V \supseteq U \supseteq T} (-1)^{|U-T|} \right) = \sum_{i=0}^{m} (-1)^i \binom{m}{i} = \delta_{0,m}$$

(the latter equality gotten by setting $x = -1$ in the binomial theorem). So $\psi \circ \phi(f(T)) = f(T)$. A similar calculation yields $\phi \circ \psi(f(T)) = f(T)$. So $\psi = \phi^{-1}$. □

How do we get back to our previous understanding of inclusion/exclusion? Let S be a set of properties that the elements of some given set A may or may not have. For example, let A $\mathcal{S}_n$, and let S be the set of conditions "$w(i) = i$" (so that $|S| = n$, one condition for each element of $[n]$.) For each $T \subseteq S$, let

$$f_=(T) = \#\{a \in A \mid \ A \text{ satisfies } \textit{exactly} \text{ the properties in } T\},$$

and

$$f_\geq T = \#\{a \in A \mid \ A \text{ satisfies } \textit{at least} \text{ the properties in } T\}.$$

So, in our derangement example, if $T = \{w(1) = 1, w(2) = 2\}$ then

$$f_=(T) = \#\{w \in \mathcal{S}_n \mid w(1) = 1, w(2) = 2, w(i) \neq i \text{ for } i = 3, \dots, n\}$$

and

$$f_\geq(T) = \#\{w \in \mathcal{S}_n \mid w(1) = 1, w(2) = 2\}.$$

Then

$$f_\geq(T) = \sum_{U \supseteq T} f_=(U).$$

But notice that $f_=$ and $f_\geq$ are both maps from the power set 2^S of the set of properties S to any field K containing $\mathbb{N}$ (say $K = \mathbb{C}$ if you like). Then $\phi(f_=) = f_\geq$, so that

$$f_=(T) = \phi^{-1}(\phi(f(T))) = \phi^{-1}(f_\geq(T)) = \sum_{U \supseteq T} (-1)^{|U-T|} f(U).$$

To translate from our previous notation, in our derangement example,

$$|A_i| = f_\geq(\{w(i) = i\}) \quad \text{ and for } \emptyset \subset X \subseteq [n], \quad \left| \bigcap_{i \in X} A_i \right| = f_\geq(\{w(i) = i \mid i \in X\}).$$

Suppose now that $f_=$ (and therefore $f_\geq$) only depends on the size of T (like in our derangement example). In this case, for T of size i,

$$a(n-i) f_=(T) \qquad \text{and} \qquad b(n-i) = f_\geq(T),$$

where $|S| = n$. Note that

$$\begin{aligned} a(0) = b(0) &= \#\{ \text{ elements with all of the properties } \}, \\ a(n) &= \#\{ \text{ elements with none of the properties } \}, \text{ and} \\ b(n) &= |A|. \end{aligned}$$

Then we have that

$$b(m) = \sum_{i=0}^{m} \binom{m}{i} a(i) \qquad \text{and} \qquad a(m) = \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} a(i).$$

For shorthand, for a function $\varphi : \mathbb{N} \to K$, we sometimes write

$$\Delta f(n) = f(n+1) - f(n),$$

so that if $\Delta^k f = \underbrace{\Delta \cdots \Delta}_{k} f$, then you can show

$$\Delta^k f(n) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f(n+i)$$

(see EC1, the discussion above and around equations (1.97)–(1.99)). This is called the *calculus of finite differences*. In particular, in our above example,

$$a(m) = \Delta^m b(0).$$

Exercise 38. (a) *Derangements.* Recall that a derangement of n is a permutation of n with no fixed-points, and that the number of derangements of n is given by

$$D(n) = n! \sum_{i=0}^{n} (-1)^i / i!.$$

(i) Verify this formula for $D(n)$ for $n = 3$.
(ii) Verify the recursive formula $D(n) = nD(n-1) + (-1)^n$ using the above formula for $D(n)$.
(iii) Give a combinatorial proof for the recursive formula $D(n) = (n-1)(D(n-1) + D(n-2))$.

(b) *Fixed-point free functions* Consider the set of functions $\varphi : [n] \to [n]$. Note that this differs from our work on derangements, since φ is not necessarily bijective.
(i) Let S be the set of conditions "$\varphi(i) = i$" (so that $|S| = n$, one condition for each element of $[n]$.) For $T \subseteq S$, describe $f_=(T)$ and $f_\geq(T)$ using set notation.
(ii) How many functions $\varphi : [n] \to [n]$ have no fixed points?
(iii) Let $E(n)$ be the number of fixed-point free functions $\varphi : [n] \to [n]$. Show that

$$\lim_{n \to \infty} E(n)/n^n = 1/e.$$

(c) How many permutations of $[n]$ have no cycle of length k? If $f_k(n)$ denotes this number, then compute $\lim_{n\to\infty} f_k(n)/n!$.
[Hint: for a subset $S \subseteq [n]$ of size k, let A_S be the set of permutations in which there is a k-cycle whose entries are the elements of S. It may or may not be useful to note that A_S and A_T are disjoint exactly when S and T are not disjoint.]

7.2. **Example: surjective functions.** Consider the set of functions $\varphi : [n] \to [m]$. If we put no restrictions, there are m^n of these (there are m choices for the image of 1, m choices for the image of 2, and so on). If we place the restriction that φ must be injective, with $m \geq n$, there are $m(m-1)\dots(m-n+1)$ of these (there are m choices for the image of 1, $m-1$ choices for the image of 2, and so on). However, we'll need inclusion/exclusion to compute the number of *surjective* functions from $[n]$ to $[m]$.

Let p_i be the property that $i \in [m]$ is *not* in the image of f. Then

$$f_{=}(\{p_i\}) = \#\{\varphi : [n] \to [m] \mid \varphi([n]) = [m] - \{i\}\}$$

and

$$f_{\geq}(\{p_i\}) = \#\{\varphi : [n] \to [m] \mid \varphi([n]) \subseteq [m] - \{i\}\},$$

where $\varphi([n])$ is the set of elements of $[m]$ that get mapped to by φ. Similarly,

$$f_{\geq}(\{p_i \mid i \in I\}) = \#\{\varphi : [n] \to [m] \mid \varphi([n]) \subseteq [m] - I\}, \text{ for } I \subseteq [m].$$

Note that $f_{=}$ is exactly as hard to calculate as the number of surjective functions, and $f_{\geq}$ is exactly as easy to calculate the total number of functions. Namely,

$$f_{\geq}(\{p_i \mid i \in I\}) = (m - |I|)^n.$$

Now, the number of surjective functions is

$$\begin{aligned} f_{=}(\emptyset) &= \phi^{-}(\phi(f_{=}(\emptyset))) = \phi^{-1}(f_{\geq}(\emptyset)) \\ &= \sum_{I \subseteq [m]} (-1)^{|I|} f_{\geq}(\{p_i \mid i \in I\}) \\ &= \sum_{I \subseteq [m]} (-1)^{|I|} (m - |I|)^n \\ &= \sum_{j=0}^{m} (-1)^j \binom{m}{j} (m-j)^n \end{aligned}$$

(noting that the term when $j = m$ evaluates to 0).

Recall that we have already counted

the number of ways to place...	is given by...
n distinguishable balls into m distinguishable boxes	m^n
n indistinguishable balls into m distinguishable boxes	$\binom{n+m-1}{n}$

The number or ways to place n indistinguishable balls into m indistinguishable boxes is the same as the number $p(n)$ of integer partitions of n into at most m parts, but actually this is actually a bit of a difficult problem. There is no closed formula for this number, though there is a closed form for its generating function, given by

$$\sum_{n \in \mathbb{N}} p(n) x^n = \prod_{k=1} \frac{1}{1 - x^k}.$$

So finally, how many ways are there to place n distinguishable balls into m indistinguishable boxes? Let

$$S(n,m) = \text{ number of ways to put } n \text{ dist. balls into } \textit{exactly } m \text{ indist boxes,}$$

meaning that none of the m boxes are empty. These $S(n,m)$ are called the *Stirling numbers of the second kind*. Then the number of ways to place n distinguishable balls into m indistinguishable boxes is given by $\sum_{i=1}^{m} S(n,i)$.

To calculate $S(n, m)$, first choose an ordering of the m boxes (so that they're now distinguishable). Then an assignment of n distinguishable balls into those m boxes so that none of them are empty determines a surjective function $\phi : [n] \to [m]$. Since there are $f_=(\emptyset)$ such functions, and $m!$ permutations of the m boxes, we have

$$m!S(n, m) = f_=(\emptyset) = \sum_{j=0}^{n} (-1)^j \binom{m}{j} (m - j)^n.$$

Thus

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

So finally, the number of ways to place n distinguishable balls into m indistinguishable boxes (so that they may or may not be empty) is given by

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

Exercise 39. (a) Compute the following.

(i) S(3,0), (ii) S(3,1), (iii) S(3,2), (iv) S(3,3),
(v) S(n,0), (vi) S(n,1), (vii) S(n, n-1), (viii) S(n,n).

(b) Give a combinatorial proof that the Sterling numbers of the second kind satisfy the recurrence relation

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

(c) Verify the identity

$$\sum_{k \geq 0} S(n, k)s(k, m) = \delta_{m,n}$$

for $n = 3$ and $m = 3$ and $m = 1$. (Recall $s(a, b) = (-1)^{a-b}c(a, b)$ is the Stirling number of the first kind.)

(d) Lunch box examples. Ok to give your answers in terms of $\binom{n}{m}$ or $S(n, m)$ where appropriate.
  - (i) How many ways are there to distribute 6 distinct candy bars into 4 identical lunch boxes so that every lunch box gets at least one candy bar?
  - (ii) How many ways are there to distribute 6 distinct pieces of fruit into 4 identical lunch boxes? (you might leave some empty)
  - (iii) How many ways are there to distribute 6 identical juice boxes into 4 identical lunch boxes? (you might leave some empty)
  - (iv) How many ways are there to distribute 6 identical sandwiches into 4 identical lunch boxes so that every box gets at least one sandwich?
  - (v) How many ways are there to distribute 6 identical carrots into 4 distinct lunch boxes so that every box gets at least one carrot?
  - (vi) How many ways are there to distribute 6 identical bottles of water into 4 distinct lunch boxes? (you might leave some empty)
  - (vii) How many ways are there to distribute 6 distinct cookies into 4 distinct lunch boxes? (you might leave some empty)
  - (viii) How many ways are there to distribute 6 distinct pieces of cheese into 4 distinct lunch boxes so that every box gets a piece of cheese?

Exercise 40. The following argument is very similar to the one that we used to compute $S(n, m)$, and *seems* to be an easier approach to computing the number of ways of putting n distinguishable balls into m indistinguishable boxes. However, it results in the claim that there are $m^n/m!$ such ways, which cannot possibly be true, since $m^n/m!$ is not always an integer (e.g. let $n = 2$ and $m = 3$). Find and explain the flaw in this argument.

> *To count the ways of putting n distinguishable balls into m indistinguishable boxes, first choose an ordering of the m boxes (so that they're now distinguishable). Then an assignment of n distinguishable balls into those m boxes determines a function $\phi : [n] \to [m]$. Since there are m^n such functions, and $m!$ permutations of the m boxes, there are $m^n/m!$ ways of putting n distinguishable balls into m indistinguishable boxes.*

Exhibit the flaw for $n = 2$ and $m = 3$. [Hint: consider the empty boxes. If $n = 2$ and $m = 3$ doesn't make it clear to you, try other examples, but note that you need $m \geq 3$ to see what goes wrong. If you try bigger examples, focus on maps where the image has size $|f([n])| \leq m - 2$.]

Exercise 41. Read and summarize Example 2.2.4 (up to equation (2.15)), taking time to read the first page of Section 1.4.