Math 3U03: Combinatorics

Jeremy Lane

July 7, 2021

Abstract

These incomplete lecture notes are from the course *MATH 3U03: Combinatorics* which I taught Fall 2020 at McMaster University. They have not been edited since then and may contain errors. For more information about these notes, please refer to the preface.

Contents

1 Preface

1.1 About these notes

These notes are presented as-is; I have not edited them since the end of the course, with the exception of adding this preface. They may contain errors. You have been warned.

The textbook for the course, which is referred to throughout the notes, is "How to Count: An Introduction to Combinatorics and Its Applications" by Robert A. Beeler.

Each section of these notes covers a single topic. Typically, a section lasted 1-3 lectures, depending on the material. In places where we followed the textbook closely, the notes are quite sparse. In places where we deviated from Beeler, the notes are more detailed.

During the course, sections of the lecture notes were posted the day of the corresponding lecture. The primary goal of these notes was to provide the students with a scaffold that would help them to follow along in the online lectures. As such, the notes are quite terse, often presented in bullet form. Additional details, where needed, were provided in the lectures.

The secondary purpose was to provide them with a reference which organized the material in the order which we covered it in class. This is because we deviated from the order and contents of Beeler in several places. There are also several additional topics that we covered in the course which were not present in Beeler. The lecture notes were the primary references for those topics.

The notes include a number of practice problems. Some of these problems are from Beeler while others were adapted from other sources or invented by me. I have made no attempt to provide proper attribution for the sources of these practice problems.

Finally, these notes do not represent the sum total of material that I created for the students. I also provided them with supplementary notes on proofs, problem solving, and induction. Those notes have not been included here.

1.2 About MATH 3U03

MATH 3U03 is a third year combinatorics course at McMaster which has several features that affected how I designed the course. This section describes those features. Pedagogical choices for the course are discussed in the following section.

The first concerns the students' backgrounds. This course does not have a proofs course as a pre-requisite. Many of the students in this course have not previously encountered proofs in a serious way, nor have they dealt with fundamental concepts such as finite sets and functions, other than those of real variables. On the other hand, a good number of the students taking this course are actually fourth year students who have encountered proofs several times before and are proficient with those fundamental tools such as sets and functions. Thus, the level of mathematical maturity of the students in this course varies quite broadly. It is important to provide a balance which will entertain students at both ends of this broad spectrum.

The second is a common feature of discrete math courses in math departments. Namely, there are several other courses in the math department and the computer science department which duplicate many of the topics typically covered in MATH 3U03. Thus, it is often the case (e.g. when treating binomial coefficients) that students have seen some version of the material before. When teaching a course such as this one must be intentional in the course design so as to provide the students with something they have not seen before (to ensure the course is intellectually stimulating and sufficiently challenging).

1.3 My approach to MATH 3U03

My primary goals in designing this course were two-fold. The first was to present material that would be engaging and intellectually stimulating for everyone in the broad cohort of students. The second was to present combinatorics with a unique flavour that both distinguishes it from other discrete math courses in the math and computer science departments at McMaster and does justice to the elegance of combinatorial thinking. In order to achieve these goals, I made the following choices in the course design and selection of topics.

The first and most significant choice was to focus heavily on combinatorial proofs. The course begins with combinatorial proofs from the very first lecture. Background material about sets and functions (sections 2-6) is introduced alongside and motivated by combinatorial proofs. This theme continues for most of the remainder of the notes. I found the book "Proofs that Really Count: The Art of Combinatorial Proof" by Benjamin and Quinn to be an excellent instructor reference for this material. As someone who is not a practitioner of combinatorics and who has had little prior exposure to combinatorial proofs, I found this book to be delightful and indispensable. Much of these notes are inspired by, or follow directly, sections from this book. However, since this book is not written at an appropriate level for all the students in my course, I did not provide it as a reference to the students. This was one of the motivations for providing lecture notes. Although Beeler contains some combinatorial proofs, the book is not organized around them as a central theme. Moreover, the background material in Beeler is not presented in the same manner or order.

From my perspective, the approach of focusing on combinatorial proofs was quite successful. A majority of the students, both those with and without prior exposure to proofs, found the material on combinatorial proofs challenging and intellectually stimulating. I believe this is for several reasons. First, combinatorial proofs are direct proofs so they do not require a more detailed background on logic. Second, most combinatorial proofs follow the same pattern so students do not need to focus as much on proof writing because there is somewhat of a template. These two points mean that combinatorial proofs are something which students can master without a prior background in proofs. They were introduced early in the course so that the students with minimal proofs experience could have more time to adapt. Third, combinatorial proofs are quite different from proofs in courses such as analysis and can be rather challenging, even to students who have already mastered proofs in other courses. They require students to take a structured problem solving approach and think carefully about how to break problems into sub-cases and to experiment with many examples. They also present an important challenge to students who have an overly formal view of what proofs are. Those students might ask questions such as "how can it be a proof if there are no symbols?" or "How can talking about committees and some guy named Bob constitute a proof?" and it is, I believe, very important and helpful for those students to confront these questions. Finally, combinatorial proofs provide a distinct flavour which I believe sets this course apart from other courses at McMaster whose topics overlap.

A particularly nice feature of combinatorial proofs is that it is not difficult to cook up problems which require a good deal of thought and strong problem solving skills to solve. This was reflected in the homework for the course. Homework problems were designed to be challenging and to require a good deal of thought before a solution would be evident. They typically required students to be very sytematic, to work with some small examples in order to find a pattern, or to employ other common problem solving strategies. Accordingly, discussion of problem solving strategies was a common theme in the course. Unfortunately, much of these discussions were informal and are not reflected in the lecture notes. It is also worth mentioning that half of the homework problems were peer graded using the platform Kritik. The peer grading component of the course had, in my view, mixed results. I will not go into further details about that here.

The second major choice was to postpone material relating to existence problems and more complex combinatorial structures to the end of the course. Proof by contradiction, constructive proofs, and the pigeonhole principle are not treated until sections 24 and 25. My reasoning was that since these problems require a completely different type of logic from combinatorial proofs, presenting them in a different parts of the course (and treating each in more detail) would be helpful to the students. These techniques also fit thematically with the more advanced topics of block designs and finite projective planes which necessarily came at the end of the course. On the other hand, because existence problems appeared so late in the course, students had less time to accustom themselves with these techniques (proof by contradiction, constructive proofs, and the pigeonhole principle) and some communicated to me that they found this particularly difficult. One might argue that it makes sense to cover these topics somewhat earlier, while still leaving the more complex applications (such as block designs) to the end of the course.

1.4 Overview

Finally, I will give a brief overview of what these notes contain and the topics that we covered in MATH 3U03.

Sections 1-6 introduce combinatorial proofs alongside the pre-requisite material about finite sets and functions. Although the material involving sets and functions is rather straightforward, the combinatorial proofs presented in this section will be challenging for most students. It is worth noting that we also covered advanced induction techniques and there was previously a more detailed section on proofs and problem solving. Those sections have been removed because they are quite unpolished.

Sections 7-14, 16, and 17 cover combinatorial numbers (factorials, binomial coefficients, etc.) with a very heavy emphasis on their combinatorial interpretations and combinatorial proofs.

The principle of inclusion and exclusion is covered in Section 15.

Sections 18-23 cover generating functions and recurrence relations. I chose not to cover some of the less exciting material from the corresponding sections of Beeler. On the other hand, I include some interesting material which does not appear in Beeler, such as exponential generating functions and a discussion explaining how the method of characteristic polynomials for linear homogeneous recurrence relations is related to the method of characteristic polynomials for linear homogeneous differential equations.

Sections 24-30 deal with existence problems and more complex combinatorial structures. I chose to cover block designs because they are an interesting and broad topic that has many applications and is an active area of modern research. Another up-shot was that Beeler provides a fairly good coverage of this material. Block designs are also give a very good playground full of examples of non-trivial existence problems. My favourite lecture from this section of the course was the last, where we discussed how the card game Spot It! is modelled on a finite projective plane. Unfortunately, this lecture did not make it into the notes, but one can find writings on this topic in various blogs with a quick internet search.

Although they are often presented in discrete math courses alongside combinatorics, graphs are not covered in this course. The reason is that 3U03 is a sister course to MATH 3V03: Graph Theory (which is not anti-requisite). As such, it did not make sense to duplicate material between the two courses.

2 What is Combinatorics?

- Three types of common combinatorial problems:
  1. Give a formula for the number of elements in a finite set (aka its *cardinality*).
  2. Show two finite sets have the same cardinality.
  3. Prove an equation by interpreting both sides as the cardinality of the same finite set.
- **Example:** Count the number of ways c_n to triangulate a regular $n{+}2$-gon by diagonals.

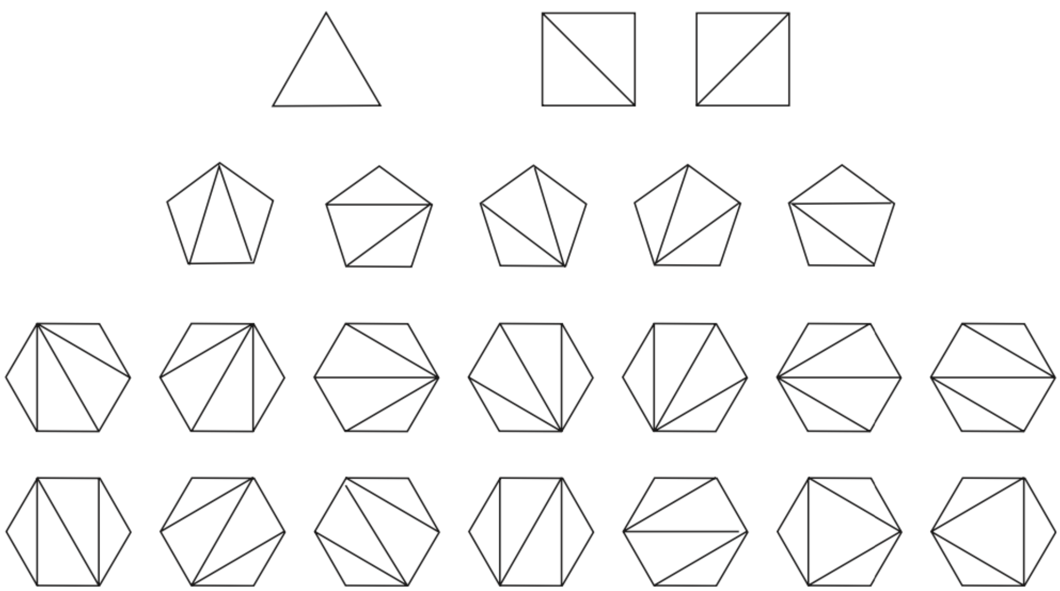

Figure 1.1. Triangulated polygons.

Figure 1: Enumerating triangulations of regular polygons. From "Catalan Numbers" by Richard Stanley.

- The first few numbers in this sequence:

$$1, 2, 5, 14, 42, 132, 429, 1430, \ldots$$

- **Example:** Count the number of ways p_n to "parenthensize" or "bracket" the product x^{n+1}.

$$xx$$

$$(xx)x, \quad x(xx)$$

$$x(x(xx)), \quad x((xx)x), \quad (xx)(xx), \quad (x(xx))x, \quad ((xx)x)x$$

- **Example:** You are at the south-west corner in a city whose streets form a $n \times n$ square grid. A railroad runs diagonally though the city from the south-west corner to the north-east corner. Count the number of ways d_n to walk to the north-east corner never walking south, nor west, nor crossing to the north-west side of the tracks.

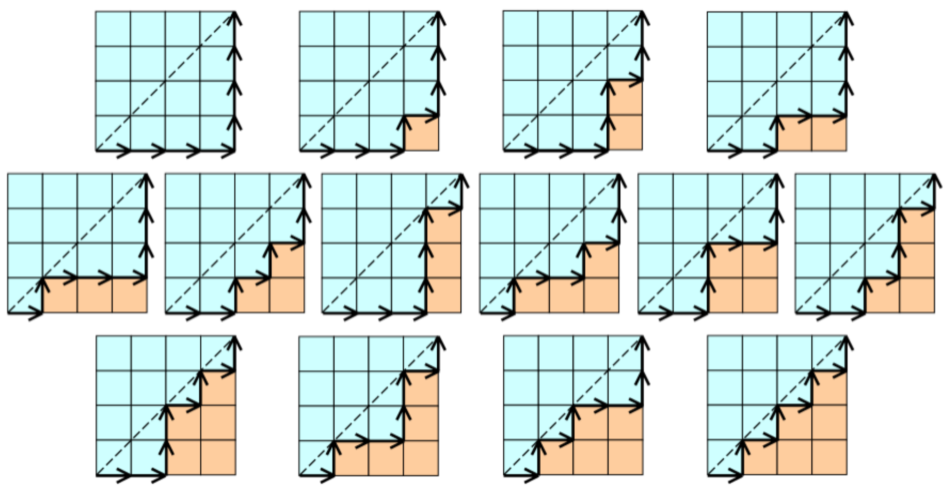

Figure 2: 14 ways to walk through a 4×4 city without crossing the tracks or going backwards. From the Wikipedia page on Catalan numbers.

The following result is an example of a solution to a combinatorial of type 1.

Theorem 2.1. *For any positive integer n,*

$$c_n = \frac{(2n)!}{(n+1)(n!)^2}.$$

The following result is an example of a solution to a combinatorial of type 2.

Theorem 2.2. *For any positive integer n,*

$$c_n = p_n = d_n.$$

We will come back to these results later in the course.

In order to excel at these sorts of problems, we need to have a strong foundation. Here is the plan for the next few weeks:

- Introduce/review fundamental concepts for the course (sets, functions, proofs, induction). These are described in Chapter 1 of How To Count.
- Introduce the three combinatorial problems mentioned above (counting, bijections, combinatorial proofs).

Existence problems are another major type of combinatorial problem not discussed in this introduction. They typically involve very different tools from those used on the types of problems discussed here. We will see existence problems towards the end of the course.

2.1 Practice Problems

1. In Lecture 1 (and in the lecture notes) we see a complete list of all the triangulations of the pentagon (5-gon) and hexagon (6-gon). Give a thorough explanation of why these lists are complete. i.e.

    - Show that every triangulation of the pentagon is one of the 5 triangulations listed in the lecture.
    - Show that every triangulation of the hexagon is one of the 14 triangulations listed in the lecture notes.

    (In order to answer these questions correctly, you need to find a way to be systematic.)

3 Sets and Counting

3.1 Sets

- A *set* is a collection of distinct objects (called *elements* of the set).
- A set is *finite* if it has finitely many elements.
- The *cardinality* of a finite set S is the number of elements it contains. Denoted $|S|$. (Some people denote it $\#S$).
- Examples:
  - $\mathbb{R}$ the set of real numbers. Not finite.
  - $\mathbb{Q}$ the set of rational numbers. Not finite.
  - $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ the set of integers. Not finite.
  - $\mathbb{N} = \{0, 1, 2, \dots\}$ the set of non-negative integers. Not finite.
  - $S = \{1, 4, 9, -3, 8\}$. Finite. $|S| = 5$.
  - Given an integer $n \geq 1$, denote
    $$[n] = \{1, 2, 3, \dots, n\},$$
    the set of integers 1 to n. It is finite. $|[n]| = n$.
  - The empty set is the set containing 0 elements. Denoted $\emptyset$ or, less commonly, $\{\}$.

3.2 Counting

For the rest of the lecture, all sets are finite.

- As we mentioned in the previous lecture, a fundamental problem in combinatorics is counting the number of elements in a finite set, i.e. computing its cardinality.
- Although it may sound easy, it can often be tricky. We need strategies to help us.
- One of the simplest and most useful strategies for counting is to subdivide the set into smaller, disjoint sets, count them individually, and then add the results.

In order to state this mathematically, we need to recall the notions of union and intersection:

- The *union* of two sets A and B is the set $A \cup B$ consisting of all elements of A or B.
- The *intersection* of two sets A and B is the set $A \cap B$ consisting of elements of both A and B.

- Example: $A = \{0, 1, 2\}$, $B = \{1, 3, 4\}$, $C = \{0, 2\}$.
  - $A \cup B = \{0, 1, 2, 3, 4\}$. $A \cap B = \{1\}$.
  - $B \cup C = \{0, 1, 2, 3, 4\}$. $B \cap C = \emptyset$.

Proposition 3.1. *(The addition principle, Prop 1.3.6) If $S = A \cup B$ and $A \cap B = \emptyset$, then*

$$|S| = |A| + |B|.$$

- Thus, the problem of counting S reduces to two hopefully easier problems: counting A and counting B.
- Two sets are *disjoint* if their intersection is empty.
- A decomposition of a set S into two disjoint sets (i.e. $S = A \cup B$ and $A \cap B = \emptyset$) is called a *partition* of S.

3.3 Example: Tilings of a $1 \times n$ chessboard

(see How To Count, Exercise 1.6.6.)

- Given a positive integer n, consider a $1 \times n$ chessboard.
- Let T_n be the set of tilings of a $1 \times n$ chessboard using 1×1 tiles (squares, also known as 1-ominos) and 1×2 tiles (dominos, also known as 2-ominos). Let $t(n) = |T_n|$.
- (By a *tiling* we mean that we cover the chessboard with squares and dominos so that the pieces do not overlap)
- Ex: $t(1) = 1$, $t(2) = 2$, $t(3) = 3$, $t(4) = 5$.

We can derive a recurrence relation for the numbers $t(n)$ as follows:

- Let A_n be the set of tilings in T_n that end with a square. Let B_n be the set of tilings in T_n that end with a domino.
- (By 'end with a square' we mean that the rightmost tile is a square)
- The sets A_n and B_n form a partition of T_n, i.e. $A_n \cup B_n = T_n$ and $A_n \cap B_n = \emptyset$.
- Thus,
$$t(n) = |T_n| = |A_n| + |B_n|.$$
- What is the cardinality of A_n and B_n?
- Every element of A_n corresponds to a tiling of a $1 \times (n - 1)$ chessboard. Thus, $|A_n| = |T_{n-1}| = t(n - 1)$.

- Every element of B_n corresponds to a tiling of a $1 \times (n-2)$ chessboard. Thus, $|B_n| = |T_{n-2}| = t(n-2)$.
- Putting this together, we have a *recurrence relation* for $|T_n|$. For $n \geq 2$,

$$t(n) = t(n-1) + t(n-2).$$

  The values $t(1)$ and $t(2)$ are the *initial values* of this recurrence.

- Our derivation of this identity is a very basic example of a combinatorial proof: we proved it by counting the number of elements in the set T_n a different way.

It is possible to derive a closed formula for $t(n)$ (see the exercises).

3.4 Example: A combinatorial proof of an identity about tilings

We want to prove the following identity:

Proposition 3.2. *For all positive integers $n \geq 2$,*

$$2t(n) = t(n+1) + t(n-2).$$

Proof. Suppose we have two $1 \times n$ chessboards. One coloured red and the other coloured blue. Let D_n be the set of all ways to tile one of the two chessboards with squares and dominos. We show that both the LHS and the RHS count the total number of ways to tile one of the two chessboards.

- First, we claim that the LHS counts the number of ways to tile one of the two chessboards, $|D_n|$:
- We can partition the set of all tilings of the red or blue chessboards as $D_n = R_n \cup B_n$ where R_n is the set of all tilings of the red chessboard, and B_n is the set of all tilings of the blue chessboard.
- Then, the total number of tilings is

$$|D_n| = |R_n| + |B_n| = t(n) + t(n) = 2t(n).$$

We need to show that the RHS counts the same thing. We partition the set a different way:

- Let A_n be the set of all tilings of the blue chessboard that end with a domino and let C_n be the set of all tilings of the red or blue $1 \times n$ chessboard that are not in A_n. Since A_n and C_n form a partition of D_n,

$$|D_n| = |C_n| + |A_n|.$$

- Removing the last domino from any such tiling in A_n produces a tiling of the blue $1 \times (n-2)$ chessboard, so
$$|A_n| = t(n-2).$$
- Finally, we want to show that $|C_n| = t(n+1)$.
- Every tiling in C_n is either a a tiling of the red chessboard, or a tiling of the blue chessboard that ends with a square.
- Adding a square to the end of the red tiling produces a tiling of a $1 \times (n+1)$ chessboard which ends with a square.
- Replacing the last square of the blue tiling with a domino produces a tiling of a $1 \times (n+1)$ chessboard which ends with a domino.
- Thus, the number of elements in C_n equals the number of tilings of a $1 \times (n+1)$ chessboard:
$$|C_n| = t(n+1).$$
- Combining everything above, we have shown that
$$|D_n| = |C_n| + |A_n| = t(n+1) + t(n-2). \qquad \square$$

3.5 Practice Problems

1. In this lecture we defined the sets T_n, $n = 1, 2, 3, \ldots$. We listed all the elements of the set T_1, T_2, T_3, and T_4. List all the elements of T_5, Be sure to explain why your list is complete.
2. Recall that the Fibonacci numbers are a sequence F_n, $n = 1, 2, 3, ...$ that is defined using a recurrence relation:
$$F_0 = 0, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2}.$$
   (a) Compute the Fibonacci numbers up to $n = 6$ (using the definition given here).
   (b) What is the relationship between the numbers $t(n)$ and the Fibonacci numbers? (give a proof of the relationship you found)

   A closed formula for F_n is derived in How To Count, Proposition 1.2.2.
3. (How To Count, Exercise 1.6.4) In chess, a king is *attacking* if it is adjacent to another piece. Find a recurrence relation and initial values for $k(n)$, the number of ways of placing non-attacking kings on a $1 \times n$ chessboard.
4. (How To Count, Exercise 1.6.7) Find a recurrence relation and initial values for $W(n)$, the number of words of length n from the alphabet $\{a, b, c\}$ with no adjacent a's.
5. Give a combinatorial proof of the following identity:
   - For all positive integers $n \geq 2$, $3t(n) = t(n+2) + t(n-2)$.

4 Counting with bijections

4.1 Functions

- Suppose X and Y are sets (finite or infinite).
- A *function with domain X and codomain Y* is a rule that assigns to every element $x \in X$, a unique element $y \in Y$.
- $f\colon X \to Y$ denotes a function named 'f' with domain X and codomain Y. It sends $x \in X$ to $f(x) \in Y$.
- Be careful: the domain and codomain are part of the definition of a function.
- Example: All of the following functions are different.
  - $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$.
  - $f\colon [0,\infty) \to \mathbb{R}$, $f(x) = x^2$.
  - $f\colon \mathbb{R} \to [0,\infty)$, $f(x) = x^2$.
  - $f\colon [0,\infty) \to [0,\infty)$, $f(x) = x^2$.

4.2 Injective, Surjective, Bijective

- A function $f\colon X \to Y$ is *injective* (one-to-one) if $f(x_1) = f(x_2)$ implies that $x_1 = x_2$. In other words, the number of elements of X that are sent to an element y in Y is at most 1.
- The *image* of a function $f\colon X \to Y$ is the set of values,

  $$f(X) = \{f(x) \in Y \mid x \in X\}.$$

  Another common notation for the image of a function is $\mathrm{Im}(f)$.
- A function $f\colon X \to Y$ is *surjective* (onto) if its image equals its codomain; $f(X) = Y$. In other words, every element y in Y is the image of at least one element x in X under f.
- A function $f\colon X \to Y$ is *bijective* if it is both injective and surjective. In other words, every element in Y is the image of exactly one element in X.
- Example:
  - $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Not injective. Not surjective
  - $f\colon [0,\infty) \to \mathbb{R}$, $f(x) = x^2$. Injective but not surjective.
  - $f\colon \mathbb{R} \to [0,\infty)$, $f(x) = x^2$. Surjective but not injective.
  - $f\colon [0,\infty) \to [0,\infty)$, $f(x) = x^2$. Bijective.

- Recall that we can compose functions. f bijective is the same as saying that f has an inverse.
- Example: Let $\mathcal{A} = \{a, \ldots, z, A, \ldots, Z\}$. Define the 'capitalization function,'
$$Cap\colon \mathcal{A} \to \mathcal{A},$$
which sends a letter in $\mathcal{A}$ to its capitalization, e.g. $Cap(a) = A$, $Cap(A) = A$.
  - Define a new function by modifying the codomain of Cap so that it is surjective.
  - We can also define a new function by restricting the domain. Describe all the largest possible domains of $\mathcal{A}$ such that Cap is injective. What are the inverse functions?

4.3 Counting sets using functions

For the rest of the lecture, all sets are finite.

Theorem 4.1. *(Theorem 1.4.4) Let $f\colon X \to Y$ be a function. Then:*

- *If $f\colon X \to Y$ is injective, then $|X| \leq |Y|$.*
- *If $f\colon X \to Y$ is surjective, then $|X| \geq |Y|$.*
- *If $f\colon X \to Y$ is bijective, then $|X| = |Y|$.*

The most important consequence of this theorem is that we can use bijections to count:

- Suppose we want to count the number of elements in X (i.e. we want to compute the cardinality, $|X|$).
- If we can find a bijection $f\colon X \to Y$ and we can count the number of elements in Y, then we can count the number of elements in X since $|X| = |Y|$.
- This counting strategy is helpful if it is somehow easier or we already know how to count Y.

4.4 Example: words in the alphabet $\{a, b\}$

- For any positive integer n, let W_n denote the set of words of length n, using only the letters a and b, which do not contain aa. Let $w(n) = |W_n|$.
- In math, a *word* is just a string of characters. It does not need to be a word in an English dictionary (or any dictionary for that matter). The set of letters/characters used to form words is called an *alphabet.* In math the alphabet can be whatever we want.

- Examples:
  - $W_1 = \{a, b\}$.
  - $W_2 = \{ab, ba, bb\}$.
  - $W_3 = \{bbb, bba, bab, abb, aba\}$.
- For any positive integer n, define a function $f \colon W_n \to T_{n+1}$ as follows:
  - Suppose we have a word w in W_n.
  - Write the letters of w in order, in the first n squares of the $1 \times (n+1)$ chessboard. Write 'b' in the last square.
  - On every pair of sequential squares with 'ab', place a domino. On the remaining squares (which will all have 'b's in them), place a square.
  - The result is an element of T_{n+1}; a tiling of the $1 \times n$ chessboard by squares and dominos. This tiling is $f(w)$.
- The function $f \colon W_n \to T_{n+1}$ is a bijection.
  - Exercise.
- Thus, $w(n) = t(n + 1)$.
- This example is also discussed, in a different way, in How to Count, Section 1.6.

4.5 Practice problems

1. Show that the function $f \colon W_n \to T_{n+1}$ defined in the lecture is a bijection.
2. A binary string is composed entirely of zeros and ones. Define $b(n)$ to be the number of binary strings of length n that have no adjacent ones. Use a bijective proof to derive a recurrence relation and initial values for $b(n)$. (compare with How To Count, Exercise 1.6.3)
3. Let $\tau(n)$ be the number of tilings of a $1 \times n$ chessboard using 1×1 squares and 1×3 'triominoes.' Use a bijective proof to derive a recurrence relation and initial values for $\tau(n)$. (Compare with How To Count, Exercise 1.6.8)
4. Let $s(n)$ be the number of ways of writing n as an ordered sum where the summands are either 1 or 2. Use a bijective proof to derive a recurrence relation and initial values for $s(n)$. (Compare with How To Count, Exercise 1.6.5)
5. Enumerate (list) all functions $f \colon [n] \to \{0, 1\}$ for $n = 1, 2, 3, 4$. How many are there?
6. Enumerate all bijections $f \colon [3] \to [3]$. How many are there?
7. Recall the following definitions:

- Given functions $f\colon X \to Y$ and $g\colon Y \to Z$, we can form their *composition*:
$$g \circ f\colon X \to Z$$
by the rule $g \circ f(x) = g(f(x))$.
- A function $g\colon Y \to X$ is an *inverse* of a function $f\colon X \to Y$ if:
  - $f \circ g(y) = y$ for all $y \in Y$, and
  - $g \circ f(x) = x$ for all $x \in X$.

8. Show that a function $f\colon X \to Y$ is a bijection if and only if it has an inverse $g\colon Y \to X$.

9. Show that inverses are unique, i.e. if $g\colon Y \to X$ and $h\colon Y \to X$ are inverses of $f\colon X \to Y$, then they are equal (as functions).

10. Show that if $f\colon X \to Y$ is injective and $g\colon Y \to Z$ is injective, then $f \circ g$ is injective.

11. Show that if $f\colon X \to Y$ is surjective and $g\colon Y \to Z$ is surjective, then $f \circ g$ is surjective.

12. Show that if $f\colon X \to Y$ is bijective and $g\colon Y \to Z$ is bijective, then $f \circ g$ is bijective.

5 Induction and finite disjoint unions

5.1 Induction

- Suppose $P(n)$ is a mathematical statement that depends on a positive integer n (we give several examples below). We often want to prove a mathematical statement of the form "For all positive integers n, $P(n)$." *Mathematical induction* is a fact which is useful for proving these sorts of statements:

- **The principle of mathematical induction:** Let $P(n)$ be a mathematical statement which depends on a positive integer n. IF:

  1. $P(1)$ is true. AND
  2. For all positive integers n, if $P(n)$ is true, then $P(n+1)$ is true.

  THEN: for all positive integers n, $P(n)$ is true.

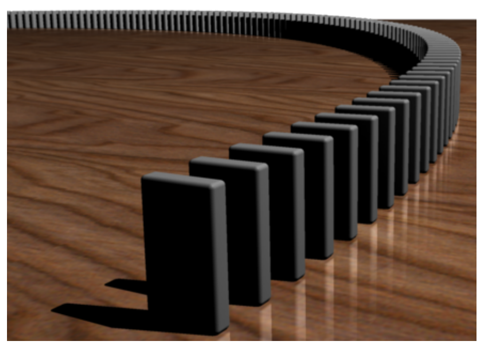

Figure 3: Source: wikipedia

- One nice analogy for induction is to imagine the positive integers as an infinite sequence of dominos, arranged as in the picture so that when the nth domino falls over, it also knocks down the $n+1$st domino (induction step). Thus, as long as we knock down the first domino (the base case), then we will have knocked down all the dominos[1].

- To prove a statement $P(n)$ is true for all n using induction, we simply need to check that both 1 and 2 are true. 1 is called the *base case*. 2 is called the *induction step*.

Proposition 5.1. *For all positive integers n,*

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}. \tag{1}$$

[1]The dominos referenced in this analogy are unrelated to the dominos used in tiling $1 \times n$ boards

Proof. We give a proof using induction.

Base case ($n = 1$): In this case, we verify (1) by direct calculation:

$$\sum_{k=1}^{1} k = 1 = \frac{1(2)}{2}.$$

Induction step: Assume we know (1) holds for some positive integer $n \geq 1$ (this assumption is called the *induction hypothesis*). We prove it holds for $n+1$. Using the induction hypothesis, we have

$$\begin{aligned} \sum_{k=1}^{n+1} k &= \left(\sum_{k=1}^{n} k\right) + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) \\ &= \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)((n+1)+1)}{2}. \end{aligned} \tag{2}$$

So the statement holds for $n + 1$.

Thus, we have shown by mathematical induction that (1) holds for all positive integers $n \geq 1$. $\square$

Here is another example. Recall that the Fibonacci numbers are the sequence F_n, $n = 0, 1, 2, 3, ...$ defined using the recurrence relation:

$$F_0 = 0, \quad F_1 = 1, \quad \text{For all } n \geq 2,\ F_n = F_{n-1} + F_{n-2}.$$

Proposition 5.2. *For all non-negative integers n,*

$$\sum_{k=0}^{n} F_{2k+1} = F_{2n+2} \tag{3}$$

Proof. We give a proof using induction.

Base case ($n = 0$): We verify (3) by direct computation:

$$F_1 = 1 = F_2.$$

Induction step: Assume we know (3) holds for some non-negative integer $n \geq 0$. We prove it holds for $n + 1$. Using the induction hypothesis and the recurrence relation for the Fibonnaci numbers, we have

$$\begin{aligned} \sum_{k=0}^{n+1} F_{2k+1} &= \left(\sum_{k=0}^{n} F_{2k+1}\right) + F_{2(n+1)+1} \\ &= F_{2n+2} + F_{2n+3} \\ &= F_{2n+4} \\ &= F_{2(n+1)+2}. \end{aligned} \tag{4}$$

So the statement holds for $n+1$.

Thus, we have shown by mathematical induction that (3) holds for all positive integers n. $\square$

5.2 Disjoint unions

- Let $n \geq 2$ be a positive integer. Given sets $A_1, \ldots, A_n$, denote

  $$\bigcup_{k=1}^{n} A_k = A_1 \cup \cdots \cup A_n,$$

  the union of all the sets $A_1, \ldots, A_n$.

- Sets $A_1, \ldots, A_n$ are *mutually disjoint* if $A_i \cap A_j = \emptyset$ for all $1 \leq i, j \leq n$, $i \neq j$.

The following proposition is proven using induction. See the textbook.

Proposition 5.3 (How to Count, Theorem 2.2.1, p. 29)**.** *For any integer $n \geq 2$, if $A_1, \ldots, A_n$ are mutually disjoint finite sets, then*

$$|\cup_{k=1}^{n} A_k| = \sum_{k=1}^{n} |A_k|.$$

- Generalizing the terminology from last week: if $S = \cup_{k=1}^{n} A_k$ and $A_1, \ldots, A_n$ are mutually disjoint, then we say that $A_1, \ldots, A_n$ form a *partition* of S.

5.3 Another domino example

- We want to give a combinatorial proof of Proposition 5.2.

- Recall that we defined $t(n)$ to be the number of tilings of a $1 \times n$ board using square and domino shaped tiles.

- Recall that for $n \geq 2$, $t(n-1) = F_n$ (see the practice problems from lecture 2). For this reason, it is useful to define $t(0) = 1$ and $t(-1) = 0$ by convention so that

  $$\text{For all } n \geq 0,\ t(n-1) = F_n.$$

  (One can say that $t(0)$ counts the 1 possible way to tile the 1×0 board and $t(-1)$ counts the 0 ways to tile the empty board. Ultimately, this is just a convenient convention.)

- Thus, we can prove Proposition 5.2 by proving the following equivalent identity about the tiling numbers $t(n)$:

  $$\text{For all } n \geq 0, \quad \sum_{k=0}^{n} t(2k) = t(2n+1).$$

Proof. By definition, the RHS counts the number of ways to tile the $1 \times (2n+1)$ board (i.e. the number of elements in the set T_{2n+1}). We need to show that the LHS counts the same thing.

For $0 \leq k \leq n$, let A_k be the set of tilings of the $1 \times (2n+1)$ board such that the position of the last square tile is $2n+1-2k$. Since the last $2k+1$ tiles are fixed, a tiling of the $2n+1$ board whose last square tile is $2n+1-2k$ is the same as a tiling of the $2n-2k$ board, so

$$|A_k| = t(2n-2k).$$

The sets A_k are mutually disjoint and every tiling of the $1 \times (2n+1)$ board is contained in one of the A_k. In other words, the sets $A_0, \ldots, A_n$ form a partition of T_{2n+1}, so

$$|T_{2n+1}| = \sum_{k=0}^{n} |A_k| = \sum_{k=0}^{n} t(2n-2k) = \sum_{k=0}^{n} t(2k).$$

□

- As we see, the idea of the proof is to partition the set T_{2n+1} into $n+1$ mutually disjoint subsets. Each subset is described by a different value of a property (in this case, the position of the last square in the tiling). Partitioning a set in this way is called *conditioning* the set on that property.
- Comparing our two proofs of Proposition 5.2, we note that both rely on induction. In the combinatorial proof we used Proposition 5.3 whose proof relies on induction.

5.4 Practice problems

1. The following proof of Proposition 5.1 is often attributed to a young Gauss.

   Proof. Let n be a positive integer. Using algebra and properties of finite sums,

$$\begin{aligned} 2\sum_{k=1}^{n} k &= \sum_{k=1}^{n} k + \sum_{k=1}^{n} k \\ &= \sum_{k=1}^{n} k + \sum_{k=1}^{n} (n+1-k) \\ &= \sum_{k=1}^{n} (k+n+1-k) \\ &= \sum_{k=1}^{n} (n+1) \\ &= n(n+1). \end{aligned} \tag{5}$$

   Dividing both sides by 2 completes the proof. □

Superficially, it looks like this proof does not rely on mathematical induction, but it does. Explain how.

2. Any set of straight lines in the plane divides the plane into regions. For any integer $n \in \mathbb{N}$, what is the maximum number of regions formed by n lines?

3.
   - Mathematical induction (as stated above) is sometimes awkward for proving a statements holds for all positive integers. Another form of induction which is sometimes more suitable is "strong induction." Despite the name, strong induction is not mathematically stronger than ordinary induction (in fact, the two are equivalent).
   - **The principle of strong mathematical induction:** Let $P(n)$ be a mathematical statement which depends on a positive integer n. IF:
     (a) $P(1)$ is true. AND
     (b) For all positive integers n, if $P(1), P(2), \ldots, P(n)$ is true, then $P(n+1)$ is true.

     THEN: for all positive integers n, $P(n)$ is true.
   - Recall that a triangulation of a convex n-gon is a set of diagonals which divide the n-gon into triangular regions.
   - Use strong induction to prove the following proposition.

   Proposition 5.4. *For every positive integer $n \geq 3$, the number of diagonals in any triangulation of a convex n-gon is $n-3$.*

4. (How To Count, Exercise 1.5.7) Give a proof by induction that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

5. Show that for all integers $n \geq 2$,

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

6. (How To Count, Exercise 1.5.8) Suppose that the seqeunce R_n satisfies $R_n = 5R_{n-1} - 6R_{n-2}$ with $R_0 = 0$ and $R_1 = 1$. Give a proof by induction that $R_n = 3^n - 2^n$.

7. The following is a famous example of a false proof by induction. What is the gap in this proof?

   Proposition 5.5. *All horses are the same colour.*

Proof. We give a proof using induction on the total number of horses, n.

Base case ($n = 1$): If there is only one horse, then the statement is true.

Induction step: Let n be a positive integer. Suppose we know that if there is n horses in total, then they are all the same colour. We want to prove that if there is $n + 1$ horses, then they all have the same colour.

So, suppose there is $n+1$ horses. Label them $1, \ldots, n+1$. By the induction hypothesis, the first n horses, $1, \ldots, n$ all have the same colour. Similarly, by the induction hypotheses, the last n horses, $2, \ldots, n + 1$ all have the same colour. These two sets of horses (the first n horses and the last n horses) have at least one horse in common, so they must all have the same colour. This proves the induction step.

Since we have proven the statement for any possible total number of horses, we know that it holds, in particular, for the total number of horses in the world. $\square$

8. Recall that $t(n)$ is the number of tilings of a $1 \times n$ board by square and domino shaped tiles. Give a combinatorial proof of the following identity. (hint: condition on the position of the last domino)

$$\text{For all } n \geq 0, \quad \sum_{k=0}^{n} t(k) = t(n+2) - 1.$$

   (Bonus: give a non-combinatorial proof of the same identity by using induction.)

9. In chess, a rook is *attacking* if it is on the same row or column as another piece. For any positive integer n, let $p(n)$ be the number of ways of arranging n identical rooks on a $n \times n$ chessboard so that no rook is attacking. Use a combinatorial proof to derive a recurrence relation for $p(n)$. Use this recurrence relation to give an exact formula for $p(n)$.

6 Cartesian products and multiplication

- Date: September 18, 2020
- Reading: How To Count, section 1.3, 2.1.

6.1 Cartesian product

- The *Cartesian product* of two sets A and B is the set $A \times B$ of all ordered pairs (a, b), $a \in A$, $b \in B$:

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

- Example: The plane $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ is the set of ordered pairs (x, y), $x, y \in \mathbb{R}$.
- The *Cartesian product* of sets $A_1, \ldots, A_n$ is the set $A_1 \times \cdots \times A_n$ of n-tuples $(a_1, \ldots, a_n)$, $a_i \in A_i$:

$$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) \mid a_i \in A_i,\, i = 1, \ldots n\}$$

- A bit more precisely, the products $A_1 \times \cdots \times A_n$ are defined inductively as $(A_1 \times \cdots \times A_{n-1}) \times A_n$.
- One can use induction to prove the following.

Proposition 6.1 (How To Count, Proposition 1.3.10)**.** *For any integer $n \geq 2$, if $A_1, \ldots, A_n$ are finite sets, then*

$$|A_1 \times \cdots \times A_n| = \prod_{k=1}^{n} |A_i|.$$

6.2 The multiplication principle

Proposition 6.2 (The multiplication principle, How To Count, Theorem 2.1.1)**.** *Suppose that there are n sets denoted $A_1, \ldots, A_n$. If elements are selected from each set independently, then the number of ways to select one element from each set $A_1, \ldots, A_n$ is $\prod_{k=1}^{n} |A_i|$.*

Proof. Independently selecting one element a_i from each set A_i, $i = 1, \ldots, n$ is the same as selecting a tuple $(a_1, \ldots, a_n)$ from the product $A_1 \times \cdots \times A_n$. Thus the total number of choices equals the cardinality of the product, which is $\prod_{k=1}^{n} |A_i|$. □

Example 6.3. There is a lunch combo at the sushi restaurant. There are four choices of side, 12 choices of roll, and two choices of drink. Every combo comes with two rolls: roll 1 and roll 2. How many possible ways to order a lunch combo are there?[2] What if the two rolls must be different?

[2] In this question we are imagining that there is an order to the way the rolls are ordered. i.e. if someone orders avocado for roll 1 and tuna for roll 2, then that is different from tuna for roll 1 and avocado for roll 2. Later on we will deal with questions where the order is not important.

Solution. The first half of the question is describing independent choice from four sets: A the set of sides, B the set of rolls to choose for roll 1, B (again) the set of rolls to choose for roll 2, and C the set of drinks. Thus the number of possible ways to order lunch is $|A| \cdot |B| \cdot |B| \cdot |C| = 4 \cdot 12 \cdot 12 \cdot 2$.

The second half of the question is not describing independent choice since the way that you choose your second roll depends on how you chose your first roll. Nonetheless, this problem is simple enough that we can solve it with the techniques we have. Since the choice of drinks and sides is independent from how we chose our rolls, we know that the number of ways to order lunch will be

$$|A| \cdot |C| \cdot (\text{the number of ways to choose roll 1 and roll 2 so they are different.}).$$

It remains to explain how many ways there are to choose roll 1 and roll 2 so they are different. First we choose roll 1: there are 12 ways to do this. Then we choose roll 2: since we cannot choose the same as roll 1, there are 11 ways to do this. The total number of choices is $12 \cdot 11$. □

Example 6.4. The Krusty Krab restaurant has three employees: Spongebob, Patrick, and Squidward. Every month Mr. Krabs gives one of them an employee of the month award. What is the total number of ways for Mr. Krabs to distribute the awards in a year?

Solution. There are 12 months. For each month there are 3 choices for employee of the month. Since the choice of employee of the month is independent (e.g. employee of the month in September does not depend in any way on who was employee of the month in October), the total number of choices is 3^{12}. □

Example 6.5. Suppose $\mathcal{A}$ is an alphabet containing n letters. How many words of length k can be formed using the alphabet $\mathcal{A}$?

Solution. This question is very similar to the previous one. Since the choices of each letter in a word are indpendent, the total number of choices is n^k. □

- The previous two examples are examples of *sampling with replacement*, which means that we choose k elements from a set of cardinality n, and we can choose a given element more than once. If the choices are independent and the order of the choices matters, then there are n^k ways to choose k objects with replacement from a set of size n.

Example 6.6. Suppose A and B are finite sets. How many functions $f\colon A \to B$ are there?

Solution. Each function $f\colon A \to B$ is a way of sampling $|A|$ elements from B with replacement: if we enumerate the elements of A by $a_1, \ldots, a_n$, then the function f corresponds to choosing the elements $f(a_1), \ldots, f(a_n)$ from B. The choice is independent since there are no restrictions on the functions. The choice depends on order, since we have ordered the elements of A.

Thus, the total number of functions $f\colon A \to B$ equals $|B|^{|A|}$. □

6.3 Practice problems

1. All the exercises from the end of How To Count, section 2.1.
2. Suppose A and B are sets (not necessarily finite). Show that the function

$$f\colon A \times B \to B \times A, \quad f(a, b) = (b, a)$$

is a bijection.

3. Suppose A is a set. What is $\emptyset \times A$?

7 Permutations

7.1 Non-attacking rooks

- Recall the following from a practice problem in lecture 6:

  > In chess, a rook is *attacking* if it is on the same row or column as another piece. What is the number of ways to arrange n identical rooks on a $n \times n$ chessboard so that no rook is attacking?

- Now we introduce a similar problem:

  > A *permutation of* n is a bijection $f\colon [n] \to [n]$. How many permutations of n are there?

- The following is a bijective proof that the answers to both problems are the same:
  - Index the rows and columns of the chessboard with $1, \ldots, n$. By position (i, j) we mean column i and row j.
  - If $f\colon [n] \to [n]$ is a bijection, then rooks at positions $(1, f(1)), \ldots, (n, f(n))$ are non-attacking: no two rooks are in the same column and, since f is injective, no two rooks are in the same row.
  - On the other hand, suppose n rooks are arranged so they are non-attacking. Let $f(i)$ be the row of the rook in column i. Since the rooks are non-attacking and the number of rooks equals the number of columns, each column contains exactly one rook, so this defines a function $f\colon [n] \to [n]$. Since no two rooks are in the same row, f is injective. Since f is injective, $|\operatorname{Im}(f)| = n$. Thus f is surjective. Since f is both injective and surjective, it is bijective.
- Thus, permutations have a nice combinatorial interpretation as arrangements of non-attacking rooks.
- We will solve both counting problems simultaneously at the end of the lecture.

7.2 Notation for permutations

A permutation $f\colon [n] \to [n]$ can be represented as a $2 \times n$ matrix:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ f(1) & f(2) & \ldots & f(n) \end{pmatrix}.$$

It can also be represented by listing the values in order: $f(1)f(2)\ldots f(n)$.

7.3 Counting Permutations

For $n \geq 1$, $n! = \prod_{k=1}^{n} k$ (pronounced "n factorial"). The sequence $n!$ is defined by its initial value, $1! = 1$, and recurrence relation, $n! = n(n-1)!$.

Proposition 7.1. *The number of permutations of n is $n!$.*

Proof. Let

$$P_n = \{f\colon [n] \to [n] \mid f \text{ is a bijection}\}$$

be the set of all permutations of n and let $p(n) = |P_n|$. We show that $p(n) = n!$ by induction.

Base case ($n = 1$): Since there is one bijection $f\colon [1] \to [1]$, $p(1) = 1 = 1!$.

Induction step: Let $n \geq 2$ and assume we have proven that $p(n-1) = (n-1)!$. Partition P_n by the subsets

$$P_n(k) = \{f \in P_n \mid f(n) = k\}.$$

We claim that for each k, $|P_n(k)| = p(n-1)$ (we will prove this in a moment). Thus,

$$p(n) = \sum_{k=1}^{n} |P_n(k)| = \sum_{k=1}^{n} p(n-1) = n \cdot p(n-1) = n \cdot (n-1)! = n!.$$

This completes the proof, except for our claim that $|P_n(k)| = p(n-1)$.

Proof that $|P_n(k)| = p(n-1)$: It remains to show that $|P_n(k)| = p(n-1)$ for all $1 \leq k \leq n$. We consider two cases:

1. **($k = n$)** Restricting the domain and codomain of $f \in P_n(n)$ to the subset $[n-1]$ produces an element of P_{n-1}. This defines a bijection $P_n(n) \to P_{n-1}$, $f \mapsto f|_{[n-1]}$. Thus $|P_n(n)| = p(n-1)$.

2. **($k < n$)** Let $g\colon [n] \to [n]$ be the permutation of n that interchanges k and n, i.e.

$$g(i) = \begin{cases} i & \text{if } i \neq k, n, \\ k & \text{if } i = n \\ n & \text{if } i = k \end{cases}.$$

   Then $P_n(k) \to P_n(n)$, $f \mapsto g \circ f$, is a bijection. Thus $|P_n(k)| = |P_n(n)| = p(n-1)$.

Since this is all possible cases, we have shown that $|P_n(k)| = p(n-1)$ for all $1 \leq k \leq n$. This completes the proof. □

- The permutation g used in the proof above is an example of a *transposition*: a permutation of n that swaps two elements and leaves the others the same. Transpositions are quite useful.

7.4 Practice problems

1. How To Count, Exercises 2.3.5 to 2.3.14.

2. A permutation $f\colon [n] \to [n]$ is an *involution* if $f \circ f = id_{[n]}$, where $id_{[n]}\colon [n] \to [n]$ is the identity function.

   (a) Show that if a permutation f is a transposition, then it is an involution.

   (b) Which non-attacking arrangements of n rooks on a $n \times n$ chessboard correspond to transpositions?

3. In our derivation of the recurrence relation for $p(n)$, we used two maps which we claimed were bijections:

   - $P_n(n) \to P_{n-1}$, $f \mapsto f|_{[n-1]}$,
   - $P_n(k) \to P_n(n)$, $f \mapsto g \circ f$.

   Check that these maps are bijections.

4. Give both an inductive proof and a combinatorial proof of the following identity:

$$\text{For all integers } n \geq 1, \quad \sum_{k=1}^{n-1} k \cdot k! = n! - 1.$$

8 Ordered subsets of $[n]$

8.1 Permutations of $[n]$

Let's recap some of what we learned about permutations last week.

- Permutations of n are bijections $f\colon [n] \to [n]$.
- In other words, a permutation is an ordering of the numbers 1 though n.
- If the first i numbers in the order have been chosen, then there are $n-i$ numbers remaining to choose the number for the $i+1$-st position. So there are
$$\prod_{i=0}^{n-1}(n-i) = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 = n!$$
ways to order the numbers 1 through n.
- By convention, $0! = 1$.

8.2 Seating problems

Before moving on to ordered subsets of $[n]$, let's go over two examples from section 2.3 about permutations.

Example 8.1 (How To Count, Example 2.3.2)**.** Find the number of ways to seat n Vulcans and n Klingons around a circular table with $2n$ seats labelled $1, 2, \dots, 2n$ so that the alien species alternate.

Example 8.2 (How To Count, Example 2.3.3)**.** Suppose we have $n \geq 3$ people to be seated around a circular table with n seats labelled $1, \dots, n$. Three of the guests, Alice, Bob, and Chad, must be seated together. How many ways can it be done?

8.3 Ordered subsets of $[n]$

- Let $P(n,k)$ (pronounced "n place k") denote the number of ordered ways to choose k elements from $[n]$.
- This can be thought of as *ordered sampling without replacement.* (Recall the number of ways to do an ordered sampling with replacement was n^k.)
- On a starship with n crew members, $P(n,k)$ is the number of ways to choose k commanders (with distinct ranks).
- Equivalently, $P(n,k)$ denotes the number of injective functions $f\colon [k] \to [n]$.
- If $k > n$, then $P(n,k) = 0$.
- $P(n,1) = n$. There are n ways to choose one element of $[n]$.

- $P(n,n) = n!$. An ordered choice of n elements from $[n]$ is the same as a permutation of n.
- $P(n, n-1) = P(n,n)$. If we choose $n-1$ elements, then there is only one choice remaining for the nth choice.
- If the first i numbers been chosen, then there are $n-i$ numbers remaining to choose the number for the $i+1$-st position. Thus, for $1 \le k \le n$,

$$P(n,k) = \prod_{i=0}^{k-1} (n-i) = n \cdot (n-1) \cdot \dots \cdot (n-k+2) \cdot (n-k+1) \; = \frac{n!}{(n-k)!}.$$

- By convention, $P(n,0) = 1$ for any n. In particular, $P(0,0) = 1$.

Example 8.3. Consider the case $n = 4$ and $k = 2$. We enumerate all the ordered ways to choose 2 elements from $[4]$:

$$12,\ 13,\ 14,\ 21,\ 23,\ 24,\ 31,\ 32,\ 34,\ 41,\ 42,\ 43.$$

There are $12 = 4!/2!$ ways in total.

Example 8.4 (How To Count, Example 2.5.3)**.** On the last mission, all the commanders on the starship Enterprise were taken hostage. The remaining crew of 15 people needs to choose 5 new commanders (along with their ranks). Alice is quite stubborn and refuses to be a commander unless she is captain. How many ways are there for the crew to pick their new commanders?

8.4 An identity for n place k

Theorem 8.5 (How To Count, Theorem 2.5.4)**.** *For all integers $n, k \ge 1$,*

$$P(n,k) = P(n-1,k) + kP(n-1,k-1).$$

Proof. We show that both sides count the number of ways to choose k commanders with distinct ranks from the crew of a starship with n crew members.

LHS The number of ways to choose k commanders with distinct ranks from the crew of a starship with n crew members is $P(n,k)$ by definition.

RHS Suppose there is a distinguished crew member of the club (let's call him Bob). We partition the set of all possible ways to choose k commanders into two subsets: those where Bob is a commander and those where Bob is not a commander.

If Bob is not a commander, then the k commanders must be chosen from the remaining $n-1$ crew members. Thus, the number of ways to choose k commanders without choosing Bob is $P(n-1,k)$.

If Bob is a commander, then there are k possible ways to assign a rank to commander Bob. There are $P(n-1, k-1)$ ways to assign the remaining $k-1$ ranks to the remaining $n-1$ crew members. Thus, by the multiplication principle there are $kP(n-1, k-1)$ ways to assign ranks so that Bob is a commander.

Finally, by the addition principle, the total number of ways to choose k commanders with distinct ranks from the crew of a starship with n crew members is

$$P(n-1, k) + kP(n-1, k-1). \qquad \square$$

8.5 Practice problems

1. Explain in your own words why $P(n, k)$ equals the number of injective functions $f\colon [k] \to [n]$.
2. (How To Count, Exercise 2.5.6) In chess a rook is attacking if it is in the same row or column of another piece. Find the number of ways to place k non-attacking rooks on a $n \times k$ chessboard. Assume $n \geq k$.
3. Re-write the combinatorial proof of the Theorem from the lecture using arrangements of k non-attacking rooks on a $n \times k$ chessboard.
4. Give an algebraic proof of the Theorem from the lecture using the formula for $P(n, k)$.
5. How To Count, Exercises 2.5.7, 2.5.8, 2.5.9.

9 Cycles and Seatings

9.1 Cycles of a permutation

- For example, consider the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 1 & 7 & 9 & 3 & 4 & 2 & 8 & 6 \end{pmatrix}.$$

- If we illustrate this permutation with directed arrows (from each integer to its image under π), we see that the resulting diagram (directed graph) has three components.
- Each component is a *cycle*.
- The length of a cycle is the number of arrows. The cycles for the permutation π have lengths 5, 3, and 1.
- A cycle of length k is sometimes called a k-cycle.
- The lengths of the cycles formed from a permutation π, listed in descending order is called the *cycle type* of π.
- The permutation in the example has cycle type $[5, 3, 1]$.

9.2 Cycle notation

- Another way to represent a permutation (different from the two notations we saw previously) is with *cycle notation*.
- The cycle notation for π is given by listing the numbers in each cycle in the order which they appear. It's simply a compact way of representing the diagram in the previous section without the arrows.
- The cycle notation for the example from the previous section is

$$\pi = (1, 5, 3, 7, 2)(4, 9, 6).$$

- We don't include 1-cycles in this notation. If a number does not appear in cycle notation, then it is fixed by the permutation (for example, the number 8 does not appear in the example above because $\pi(8) = 8$). The identity permutation, which would have cycle notation $(1)(2)\dots(n)$ if we included 1-cycles in our cycle notation, is usually represented by the number 1, or the notation $id_{[n]}$.
- Cycle notation is compact, but it is not unique! The permutation from our example can also be written as

$$\pi = (4, 9, 6)(1, 5, 3, 7, 2) \quad \text{or} \quad \pi = (2, 1, 5, 3, 7)(9, 6, 4).$$

  Thus, we get the same permutation even if we shift the cycles.

9.3 Examples

Example 9.1. Count the number of permutations of 8 with cycle type [8].

Solution.

$$8!/8 = 7!.$$

□

Example 9.2. Count the number of permutations of 8 with cycle type $[5, 1, 1, 1]$.

Solution.

$$P(8,5)/5 = \frac{8!}{(8-5)!}\frac{1}{5} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{5}.$$

□

Example 9.3. Count the number of permutations of 8 with cycle type $[5, 3]$.

Solution. We can solve this two ways:

$$P(8,5)/5 \cdot P(3,3)/3 = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{5}\frac{3 \cdot 2 \cdot 1}{3}.$$
$$P(8,3)/3 \cdot P(5,5)/5 = \frac{8 \cdot 7 \cdot 6}{3}\frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5}.$$

□

Example 9.4. Count the number of permutations of 8 with cycle type $[4, 4]$.

Solution.

$$P(8,4)/4 \cdot P(4,4)/4 \cdot (1/2) = \frac{8!}{4! \cdot 4}\frac{4!}{4}\frac{1}{2}.$$

□

Example 9.5. Count the number of permutations of 8 with cycle type $[2, 2, 2, 2]$.

Solution.

$$P(8,2)/2 \cdot P(6,2)/2 \cdot P(4,2)/2 \cdot P(2,2)/2 \cdot (1/4!).$$

□

Example 9.6. Count the number of permutations of 8 with cycle type $[1, 1, 1, 1, 1, 1, 1, 1]$

Solution. If we do the same approach as above:

$$\prod_{k=0}^{7} P(n-k,1) \cdot (1/8!) = 8! \cdot (1/8!) = 1.$$

Thinking a bit more carefully, we know the answer must be 1 because a cycle with cycle type $[1, 1, 1, 1, 1, 1, 1, 1]$ must fix every element of [8], so it is the identity. □

9.4 Seating problems

In the previous section we saw examples of the following counting problem:

> Count all the permutations of n of a given cycle type, $[n_1, \ldots, n_k]$.

This problem has the following combinatorial interpretation:

> Count the number of ways to seat n people at k unlabelled round tables with unlabelled seats so that the number of people at each table is $n_1, \ldots, n_k$.

In this version of the problem, we must note a few details:

- The seats are unlabelled so for each table all we care about is the relative position of everyone at the table, i.e. who is sitting to a persons right and left.

- The tables are round, so any seating can be rotated clockwise or counterclockwise. Since the seats are unlabelled, any rotation of a seating is considered the same (however, a reflection of the seating is different). For example, the seating (Alice, Bob, Joe) is the same as (Bob, Joe, Alice) but different from (Bob, Alice, Joe).

- The tables are unlabelled, so each tables can only be distinguished by the number of people sitting at it. For example, if we have (Alice, Bob, Joe)(Carl, Natasha, Alex) it is considered the same as (Carl, Natasha, Alex)(Alice, Bob, Joe).

Example 9.7. Count the number of ways to seat 8 people at a round table with unlabelled seats.

Example 9.8. Count the number of ways to seat 8 people at 4 unlabelled round tables with unlabelled seats so that one table seats 5 people.

Example 9.9. Count the number of ways to seat 8 people at 2 unlabelled round tables with unlabelled seats so that one table seats 5 people.

Example 9.10. Count the number of ways to seat 8 people at 2 unlabelled round tables with unlabelled seats with 4 people at each table.

Example 9.11. Count the number of ways to seat 8 people at unlabelled tables in pairs.

Example 9.12. Count the number of ways to seat 8 people separately at unlabelled tables.

We can also consider more elaborate seating problems by adding conditions. There are several examples of this in the textbook (How To Count, Examples 2.7.5, 2.7.6). It's not difficult to imagine many similar types of problems. We present a similar example:

Example 9.13. Count the number of ways to seat 8 people at 2 unlabelled round tables with unlabelled seats so that one table seats 5 people. Frank and Elvis demand to sit next to each other.

Solution: There are two cases:

- **Case 1 (Frank and Elvis sit at the table with 5 people):** There are 2 ways to seat Frank and Elvis relative to each other at the table for 5 people. Once they are seated, there are $P(6,3)$ ways to fill the remaining 3 seats. Then there are $P(3,3)/3$ ways to seat the table of three people. In total, there are

  $$2 \cdot P(6,3) \cdot P(3,3)/3 = 2 \cdot (6 \cdot 5 \cdot 4)\frac{3 \cdot 2 \cdot 1}{3}$$

  ways to seat everyone.

- **Case 1 (Frank and Elvis sit at the table with 3 people):** There are 2 ways to seat Frank and Elvis relative to each other at the table for 3 people. Once they are seated, there are $P(6,1)$ ways to fill the remaining seat. Then there are $P(5,5)/5$ ways to seat the table of three people. In total, there are

  $$2 \cdot P(6,1) \cdot P(5,5)/5 = 2 \cdot 6 \cdot \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5}$$

  ways to seat everyone.

By the addition principle, the total number of ways is

$$2 \cdot (6 \cdot 5 \cdot 4) \cdot 2 + 2 \cdot 6 \cdot (4 \cdot 3 \cdot 2) = .$$

□

9.5 Practice problems

1. In HW2 problem 2, we are asked to discuss two proofs of the following theorem.

   Theorem 9.14. *For all integers $n \geq 1$, the number of ways to partition the set $[2n] = \{1, 2, \ldots, 2n\}$ by subsets of cardinality 2 equals*

   $$\prod_{k=1}^{n}(2k-1) = 1 \cdot 3 \cdot 5 \cdots (2n-1).$$

   Give an alternative proof of this theorem by interpreting it as a seating problem for unlabelled round tables with unlabelled seats and using the techniques discussed in this lecture.

2. How To Count, Exercises 2.7.13, 2.7.14, 2.7.15, 2.7.16, 2.7.17.

10 Stirling numbers of the first kind

10.1 Examples

Example 10.1. Count the number of ways to seat n people at a round table.

Solution: The answer is $(n-1)!$. There are two different ways to see this:

- **Solution 1:** Suppose one of the people is named Bob. Seat Bob at the table and then seat the rest of the table in a clockwise manner. There are $(n-1)$ choices for the person sitting to the left of Bob. There are $(n-2)$ choices for the person sitting to the left of that person, and so on. Continuing in this fashion, we see there are

  $$(n-1)(n-2)\dots 2\cdot 1=(n-1)!$$

  ways to seat everyone.

- **Solution 2:** Choose the first seat at the table to put someone in then seat everyone in a clockwise manner. There are n ways to put someone in that seat. There are $(n-1)$ ways to put someone in the seat to the left. Continuing in this way, we see there are

  $$P(n,n)=n(n-1)(n-2)\dots 2\cdot 1=n!$$

  ways to seat everyone. However, by choosing the first seat to put someone in and then seating in a clockwise manner we have ordered the seats. But the seats were supposed to be unlabelled, so we have counted each seating exactly n times. Thus, we must divide our result by n to get the correct number of seatings:

  $$\frac{P(n,n)}{n}=\frac{n!}{n}=(n-1)!.$$

□

Example 10.2. Given n people, count the number of ways to seat k of them at a round table.

Solution: The answer is

$$\frac{P(n,k)}{k}=\frac{n!}{k(n-k)!}.$$

There are two different ways to see this:

- **Solution 1:** Choose the first seat at the table to put someone in and then seat people in a clockwise fashion. There are $P(n,k)$ ways to do this. However, because we have ordered the seats, we have counted each seating exactly k times. The number of seatings is thus

  $$\frac{P(n,k)}{k}.$$

- **Solution 2:** From the n people, choose k to be seated at the table. There are $\binom{n}{k}$ ways to do this. By the previous example, there are $(k-1)!$ ways to seat these people at the table. By the multiplication principle, there are
$$\binom{n}{k} \cdot (k-1)!$$
ways in total. We have not yet seen these numbers $\binom{n}{k}$ (we will learn about them next week). Once we do, we will see that, in fact,
$$\binom{n}{k} \cdot (k-1)! = \frac{n!}{k(n-k)!}.$$

□

Example 10.3. Count the number of ways to seat 4 people at two unlabelled round tables so that at least one person sits at each table.

Solution: There are two cases:

- **Case 1 ($[2,2]$):** Following the same argument as in the previous lecture, there are
$$\frac{P(4,2)}{2}\frac{P(2,2)}{2}\frac{1}{2} = 3$$
ways to seat two people at each table. We can see this another way: suppose one of the 4 people is named Bob. There are 3 ways to choose who sits at the table with Bob. Since the remaining two people sit at the other table together (and there is only one possible cyclic ordering of 2 elements), this is the total number of choices.

- **Case 2 ($[3,1]$):** There are
$$\frac{P(4,3)}{3} = 8$$
ways to seat three people at the same table. Another way to see this: There is 4 ways to choose the person who sits alone and $2 = (3-1)!$ ways to seat the remaining 3 people at a table together. Thus there are $8 = 4 \cdot 2$ ways in total.

The total number of ways is $8 + 3 = 11$. □

Example 10.4. Count the number of ways to seat 5 people at two unlabelled round tables so that at least one person sits at each table.

Solution 1: There are two cases:

- **Case 1 ($[4,1]$):** There are 5 ways to choose the person who sits alone and $(4-1)! = 3! = 6$ ways to seat the remaining 4 people together for a total of $5 \cdot 6 = 30$ ways.

- **Case 2 ($[3,2]$):** There are

$$\frac{P(5,3)}{3} = 20$$

ways to choose three people and seat them together. Once this is done there is exactly one way to seat the remaining 2 people together.

In total, there are $30 + 20 = 50$ ways to seat everyone together. $\square$

Solution 2: Let's suppose one person is named Bob. There are two cases:

- **Bob sits alone:** In this case, the remaining 4 people sit at a table together. There are $\left[{4 \atop 1}\right] = 3! = 6$ ways to do this.

- **Bob does not sit alone:** Seat the other 4 people at the two tables so that each table has at least one person sitting at it. There are $\left[{4 \atop 2}\right] = 11$ ways to do this. Then, seat Bob at one of the two tables. There are exactly 4 places where Bob can sit. By the multiplication principle, there are $4 \cdot 11 = 44$ ways in total.

Since these are the only two possible cases, by the addition principle $6 + 44 = 50$. $\square$

10.2 Stirling numbers of the first kind

- The *cycle index* of a permutation is the number of cycles in the permutation.

- For example: The permutations $(1,2,3)(4,5)$ and $(1,2,4,5)$ both have cycle index 2.

- The *(unsigned) Stirling numbers of the first kind*, denoted $\left[{n \atop k}\right]$, count the number of permutations of n with cycle index k. (Your textbook uses the notation $s(n,k)$.)

- Equivalently, $\left[{n \atop k}\right]$ is the number of ways to seat n people around k circular unlabelled tables with unlabelled seats, where each table must seat at least one person.

- For $n \geq 1$, $\left[{n \atop 0}\right] = 0$.

- By convention, $\left[{0 \atop 0}\right] = 1$.

- $\left[{n \atop 1}\right] = (n-1)!$ since this is the number of ways to seat n people at a round table.

- $\left[{n \atop n}\right] = 1$ since this is the number of ways to seat n people at n unlabelled tables so that each table has at least one person. Equivalently, the only cycle type of index n is $[1,1,1,\ldots,1]$.

10.3 An identity for Stirling numbers of the first kind

Theorem 10.5. *For all integers $1 \leq k < n$,*

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = n \begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}.$$

The idea of the proof is the same as the second solution to Example 10.4.

Proof. We need to show that the RHS counts the number of ways to seat $n+1$ people at k unlabelled round tables with unlabelled seats so that each table seats at least one person. Let's suppose one person is named Bob. There are two cases:

- **Bob sits alone:** In this case the remaining n people sit at the remaining $k-1$ tables together. There are $\left[\begin{smallmatrix} n \\ k-1 \end{smallmatrix}\right]$ ways to do this.
- **Bob does not sit alone:** First, seat the other n people at the k tables so that each table has at least one person sitting at it. There are $\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]$ ways to do this. Then, seat Bob at one of the two tables. There are exactly 4 places where Bob can sit. By the multiplication principle, there are $4 \cdot \left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix}\right]$ ways in total.

Since these are the only two possible cases, the identity holds by the addition principle. □

10.4 Practice problems

1. How To Count, Exercises 2.7.18 – 2.7.23.
2. Recall that the nth Harmonic number is

$$H_n = \sum_{k=1}^{n} \frac{1}{k}.$$

   (a) Compute the harmonic numbers for $n = 1, \ldots, 5$.

   (b) Compute the numbers

$$\frac{\left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix}\right]}{n!}$$

   for $n = 1, \ldots, 5$.

   (c) (challenging) Prove the two sequences are the same.

11 Binomial coefficients, part 1

11.1 Subsets

- B and A are sets.
- B is a *subset* of A if every element of B is an element of A. We write $B \subset A$.
- The *power set* of a set A is the set of all subsets of A,

$$P(A) = \{B \mid B \subset A\}.$$

- Note that $\emptyset, A \in P(A)$.

Proposition 11.1 (How To Count, 2.1.10)**.** *If A is a finite set, then $|P(A)| = 2^{|A|}$.*

Proof. Let $n = |A|$. Enumerate the elements $A = \{a_1, \ldots, a_n\}$. A subset $B \subset A$ is the same as choosing, for each $1 \leq i \leq n$, whether $a_i \in B$ or $a_i \notin B$. There are 2 choices for each $1 \leq i \leq n$, and these choices are independent, so by the multiplication principle there are 2^n possible ways to choose a subset of A. $\square$

11.2 Binomial coefficients

- Let n and k be non-negative integers.
- $\binom{n}{k}$ is the number of subsets of $[n]$ containing k elements.
- Equivalently, $\binom{n}{k}$ is the number of ways to choose a committee of k people from a group of n people.
- If $k > n$, then $\binom{n}{k} = 0$.
- $\binom{n}{0} = 1$. In particular, $\binom{0}{0} = 1$ by convention.
- $\binom{n}{n} = 1$.
- $\binom{n}{1} = n$.
- $\binom{n}{k} = \binom{n}{n-k}$.

Proposition 11.2 (How To Count, 3.1.1)**.** *For all $n \geq k \geq 0$,*

$$\binom{n}{k} = \frac{P(n,k)}{k!} = \frac{n!}{k!(n-k)!}.$$

See the textbook for a proof.

Proposition 11.3 (How To Count, Corollary 3.3.3)**.**

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n.$$

The textbook derives this identity from the Binomial Theorem which we will discuss in the next lecture. We give a combinatorial proof.

Proof. We show that both sides count the number of ways to form a committee from a group of n people (where the committee can have any number of people, including 0).

RHS: Forming a committee from a group of n people is the same as choosing a subset of $[n]$. We showed that the number of subsets of $[n]$ equals 2^n in the previous section.

LHS: Condition on the number of people in the committee. There are $\binom{n}{k}$ ways to choose a committee with k people. Thus, by the addition principle, there are

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n-1} + \binom{n}{n}$$

ways in total. $\square$

11.3 An identity relating binomial coefficients and tiling numbers

Recall that the tiling number $t(n)$ counts the number of ways to tile an n-board with square and domino tiles.

Proposition 11.4.

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

Proof. We need to show that the LHS counts the number of ways to tile an n-board with square and domino tiles. Condition on the number of dominos tiles, k, in a tiling. If there are k dominos, then the tiling has $n-k$ tiles in total. Ordering the tiles from left to right, we see that there are $\binom{n-k}{k}$ ways to choose k of the $n-k$ tiles to be dominos. $\square$

11.4 Practice problems

1. Given a subset $B \subset A$, one can define the *indicator function of* B,

$$\chi_B : A \to \{0,1\}, \quad \chi_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{else.} \end{cases}$$

Let $\{0,1\}^A$ denote the set of functions with domain A and codomain $\{0,1\}$, i.e.

$$\{0,1\}^A = \{f : A \to \{0,1\} \text{ a function}\}.$$

- Show that the map

$$P(A) \to \{0,1\}^A, \quad B \mapsto \chi_B$$

  is a bijection.

- Use the previous item to give an alternate proof of (How To Count, 2.1.10). (This will be a short proof, you need to use a fact from a previous lecture.)

2. How many ways are there to arrange k non-attacking rooks on a $n \times n$ chessboard?

12 Binomial coefficients, part 2

12.1 A generalization of a familiar identity

Recall:

Proposition 12.1.

$$\sum_{j=0}^{n} j = \frac{n(n+1)}{2}.$$

We can now give this identity a combinatorial proof.

Proof. We count the number of ways to choose two numbers from the set $\{0, 1, \ldots, n\}$.

RHS: Using the formula for binomial coefficients from the previous lecture, notice that

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

By definition of the binomial coefficient $\binom{n+1}{2}$, this is precisely the number of ways to choose two numbers from a set of $n+1$ elements.

LHS: Condition on the larger of the two elements. If the larger element is j, then there are exactly j ways to choose the other element from $\{0, \ldots, j-1\}$. By the addition principle, the total number of ways to choose two elements from $\{0, \ldots, n\}$ is

$$\sum_{j=1}^{n} j. \qquad \square$$

In fact, the same argument can be used to prove a more general form of this identity:

Proposition 12.2.

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

Proof. We need to show that the LHS counts the number of ways to choose $k+1$ numbers from the set $\{0, 1, \ldots, n\}$.

LHS: Condition on the largest of the $k+1$ elements. If the largest element is j, then there are exactly $\binom{j}{k}$ ways to choose the remaining k elements from $\{0, \ldots, j-1\}$. By the addition principle, the total number of ways to choose two elements from $\{0, \ldots, n\}$ is

$$\sum_{j=0}^{n} \binom{j}{k}. \qquad \square$$

12.2 Pascal's recurrence

Recall that we saw the Stirling numbers of the first kind satisfy a recurrence relation,

$$\begin{bmatrix} n \\ k \end{bmatrix} = (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}.$$

The binomial coefficients satisfy a very similar recurrence. In fact, the combinatorial proof is also very similar: we condition on what happens to a dude named Bob.

Theorem 12.3 (How To Count, 3.1.2)**.**

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

Proof. We show that the RHS counts the number of ways to choose a committee of k people from a group of n people. Suppose one of the people is named Bob. There are two cases: the committee contains Bob, and the committee does not contain Bob. If the committee contains Bob, then there are $\binom{n-1}{k-1}$ ways to choose the remaining $k-1$ committee members. If the committee does not contain Bob, then there are $\binom{n-1}{k}$ ways to choose all k committee members from the other $n-1$ people. Since these are the only two cases, the result follows by the addition principle. □

12.3 Binomial theorem

Theorem 12.4 (How To Count, 3.3.1)**.**

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

We can derive identities for the binomial coefficients from the Binomial Theorem. One way to do this is by substituting values for x and y. For example, setting $x = y = 1$ recovers

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

Setting $x = -1$ and $y = 1$ recovers

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots$$

Another thing we can do is apply calculus operations such as differentiation and integration. For example: Differentiate with respect to x and then set $x = y = 1$, we recover

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

It's worth mentioning at this point that the Stirling numbers of the first kind satisfy a polynomial equation as well:

$$x(x+1)(x+2)\dots(x+n-1) = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k.$$

This can be used in a similar manner to derive identities for the Stirling numbers of the first kind. For example, setting $x = 1$ recovers

$$n! = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix}.$$

12.4 Practice problems

1. Give a combinatorial proof of the identity

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots$$

   (Hint: one counts ways to choose committees with an even number of people while the other counts ways to choose committees with an odd number of people.)

2. Give a combinatorial proof of the identity

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

3. Give a combinatorial proof of the identity

$$n! = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix}.$$

   (Hint: count the number of ways to seat n people at n round tables where some tables are allowed to be empty.)

4. (challenge) Use induction to show that the Stirling numbers of the first kind satisfy the identity

$$x(x+1)(x+2)\dots(x+n-1) = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k.$$

5. How To Count, Exercises 3.4.6, 3.4.8, 3.4.9.

13 Binomial coefficients, part 3

13.1 Vandermonde identity

Example 13.1. There are 3 red candies and 5 blue candies. All the candies are distinct. How many ways are there to eat 4 of them?

A1: $\binom{3+5}{4} = 70$.

A2: Condition on the number of red candies which are eaten. Then by the addition and multiplication principle there are

$$\binom{3}{0}\binom{5}{4} + \binom{3}{1}\binom{5}{3} + \binom{3}{2}\binom{5}{2} + \binom{3}{3}\binom{5}{1}.$$

More generally:

Theorem 13.2. *For all positive integers $n, m, k \geq 0$,*

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

Sketch of the proof. We give a combinatorial proof. Suppose there are n red candies and m blue canides, all distinct. By definition of the binomial coefficient, there are $\binom{n+m}{k}$ ways to eat k of the candies (this is the LHS). Alternatively, we can condition on the number of red candies which are eaten (call this i). By the addition and multiplication principle, we recover the RHS. $\square$

13.2 Poker

- In a deck of poker cards there are 4 suits (hearts, diamonds, clubs, and spades) and 13 cards per suit (A, 2,3,4,5,6,7,8,9,10,J,Q,K). There are 52 cards in total.
- A hand in poker consists of 5 cards. There are $\binom{52}{5}$ possible hands.
- A flush is a hand where all cards have the same colour. There are 4 ways to choose the suit of a flush and $\binom{13}{5}$ ways to choose the 5 cards from that suit for a total of $4\binom{13}{5}$ ways to get a flush (by the multiplication principle).

13.3 Stars and bars

- We have two symbols: $*$ (star) and $|$ (bar). All stars are identical and all bars are identical.
- How many ways are there to order n stars and k bars. i.e. how many words of length $n+k$ are there which use exactly n stars and k bars.

- Example: $(n = 2,\ k = 2)$. There are 6 possibilities:

$$||**, \quad |*|*, \quad *||*, \quad *|*|, \quad **||, \quad |**|.$$

- The answer to the general problem is $\binom{n+k}{k} = \binom{n+k}{n}$ because we just need to choose the k positions which are to be bars, or, equivalently, we just need to choose the n positions which will be stars.

- There are many versions of this problem.

13.4 A distribution problem

Here is an important variant of stars and bars. (see How To Count, Theorem 4.2.6)

- Suppose we have n identical balls and k distinct boxes. How many ways are there to place the balls into the boxes?

- Since the boxes are distinct, we may consider them as being ordered. Represent this on a line by writing the boxes in order, left to right.

- Represent ways of putting balls into boxes by writing all the balls that go into a box to the left of that box.

- If the boxes are represented as bars, and the balls are represented as stars, then we end up with a word of length $n + k$ consisting of bars and stars.

- Since every ball must go in a box, the last letter in this word must be a bar, so the only choice we have is how the first $n + k - 1$ symbols are arranged. These $n+k-1$ symbols consist of n stars (balls) and $k-1$ bars (boxes). Since we have reduced to stars and bars, the solution is

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

- One can interpret this distribution in the following equivalent way: how many multisets of size k can be formed using the integers $1, \ldots, n$ (or any n distinct objects). If we label the boxes $1, \ldots, n$, then in the distribution problem above, the number of balls in box i tells us the number of times i occurs in the multiset.

- Because of this equivalent formulation, the solution to this problem is often called "k multichoose n" and denoted

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

13.5 Lattice paths

Here is another variant.

- Suppose we have a $n \times k$ grid representing the streets in a city with very uninspired urban planning.
- How many ways are there to travel from the point $A = (0, 0)$ to $B = (n, k)$? We do not go backwards, so we only walk to the North and East.
- Every possible path has $n + k$ steps, n of which are in the E direction and k of which are in the N direction.
- Thus, we see that the problem is the same as counting the number of words of length $n + k$ using the letter N exactly k times and the letter E exactly n times.
- This is precisely stars and bars (except stars and bars are replaced by E's and N's). Thus, the solution is

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

  As a special case, if it is a $n \times n$ square grid then the answer is

  $$\binom{2n}{n}.$$

13.6 Catalan numbers

One way which Catalan numbers can be represented is with lattice paths. Let's use this interpretation to derive the formula for the Catalan numbers.

- Recall the setting from the previous subsection, except we work with a $n \times n$ grid.
- How many ways are there to travel from the point $A = (0, 0)$ to $B = (n, n)$ without going above the diagonal?
- We will show the answer is $\frac{1}{n+1}\binom{2n}{n}$. To do this, we derive a formula for the number of paths which DO go above the diagonal. Then we subtract the result from the total number of paths, $\binom{2n}{n}$.
- In order to count the number of paths which go above the diagonal, we employ a clever trick to construct a bijection with another set of paths which are easier to count (the following is easiest to make sense of by working out an example as we did in the lecture):
- Suppose we have a path which goes above the diagonal. Let $i = 2j + 1$ be the first step where the path goes above the diagonal. At the end of that step, we will be at position $(i, i + 1)$. Thus, we have travelled East i times and North $i + 1$ times. In the rest of the path, we must travel East $n - i$ times and North $n - i - 1$ times.

- Take the remainder of the path and reflect it by switching the East and North steps. The resulting reflected path goes from $(0,0)$ to $(n-1,n+1)$.
- The procedure for reflecting paths described in the previous items defines a bijection between the set of all paths from $(0,0)$ to (n,n) which go above the diagonal, and the set of all paths from $(0,0)$ to $(n-1,n+1)$. (It is worth checking this for yourself)
- By the previous section, the number of paths from $(0,0)$ to $(n-1,n+1)$ equals
$$\binom{2n}{n+1}$$
- Thus, the number of ways to go from $A=(0,0)$ to $B=(n,n)$ without going above the diagonal equals
$$\binom{2n}{n}-\binom{2n}{n+1}=\frac{1}{n+1}\binom{2n}{n}.$$

13.7 Practice problems

1. How To Count, Exercises 3.5.5, 3.5.6, 3.5.7, 3.5.8, 3.5.9, 3.5.11, 3.5.12.

14 Multinomial coefficients

14.1 Example: distributing candies

- Bob has 6 friends and 3 types of candy (blue Whales, fuzzy Peaches, and gummy Bears). Candies of the same type are identical (unlabelled).
- How many ways are there for Bob to distribute 1 piece of candy to each of his friends so that he gives away 2 blue whales and 4 fuzzy peaches?
- A1: There are

  $$\binom{6}{4}$$

  ways to choose the friends who get fuzzy peaches.
- A2: Let's over-count by labelling the candies of each type. Now all 6 candies are distinct so there are 6! ways to distribute them. There are $2! \cdot 4!$ ways to relabel the candies, so we have over-counted each way of distributing the candies $2! \cdot 4!$ times. Thus, the number of ways to distribute the candies (without labels) equals

  $$\frac{6!}{2!4!}.$$

We can ask another version of this question:

- How many ways are there for Bob to distribute 1 piece of candy to each of his friends so that he gives away 2 blue whales, 2 fuzzy peaches, and 2 gummy bears?
- A1: Let's over-count by labelling the candies of each type. Now all 6 candies are distinct so there are 6! ways to distribute them. There are $2! \cdot 2! \cdot 2!$ ways to relabel the candies, so we have over-counted each way of distributing the candies $2! \cdot 2! \cdot 2!$ times. Thus, the number of ways to distribute the candies (without labels) equals

  $$\frac{6!}{2!2!2!}.$$

Let's consider a general version of the problem:

- Bob has $k \geq 0$ friends and $n \geq 1$ types of candy.
- Suppose $k_1, \ldots, k_n$ are non-negative integers with the property that $k_1 + \cdots + k_n = k$.
- How many ways are there for Bob to distribute 1 piece of candy to each of his friends so that k_1 friends get candies of type 1, k_2 friends get candies of type 2, and so on?

- A: Let's over-count by labelling the candies of each type. Now all k candies are distinct so there are $6k!$ ways to distribute them to the k friends. There are $k_1! \cdots k_n!$ ways to relabel the candies, so we have over-counted each way of distributing the candies $k_1! \cdots k_n!$ times. Thus, the number of ways to distribute the candies (without labels) equals

$$\frac{k!}{k_1! \cdots k_n!}.$$

14.2 Multinomial coefficients

- Let k and n be integers with $k \geq 0$ and $n \geq 1$.
- Suppose $k_1, \ldots, k_n$ are non-negative integers with the property that $k_1 + \cdots + k_n = k$.
- (Combinatorial definition) The multinomial coefficient

$$\binom{k}{k_1, \ldots, k_n}$$

is the number of ways to distribute 1 piece of candy to each of n friends so that k_j friends get candies of type j, for $j = 1, \ldots n$?

- (Algebraic definition) The multinomial coefficient is

$$\binom{k}{k_1, \ldots, k_n} = \frac{k!}{k_1! \cdots k_n!}.$$

Some quick observations about multinomial coefficients:

- The order of the numbers $k_1, \ldots, k_n$ is not important. For example:

$$\binom{6}{1,2,3} = \binom{6}{3,2,1} = \frac{6!}{1!2!3!}.$$

- When $n = 2$, multinomial coefficients are binomial coefficients:

$$\binom{k}{k_1, k_2} = \binom{k}{k_1} = \binom{k}{k_2}.$$

- Every multinomial coefficient can be computed as a product of binomial coefficients:

$$\binom{k}{k_1, \ldots, k_n} = \binom{k}{k_1}\binom{k-k_1}{k_2} \cdots \binom{k - k_1 \cdots - k_{n-1}}{k_n}.$$

This can be seen two ways: algebraically by expanding both sides, or combinatorially by giving an alternate answer to the candy distribution problem.

- A2: There are $\binom{k}{k_1}$ ways to choose the k_1 friends who get candies of type 1. Then there are $\binom{k-k_1}{k_2}$ ways to choose the remaining $k - k_1$ friends who get candies of type 2, and so on. The total number of ways then equals the product of binomial coefficients above, by the multiplication principle,

14.3 Generalized stars and bars

- Let k and n be integers with $k \geq 0$ and $n \geq 1$.
- There are n^k words of length k in the alphabet $x_1, \ldots, x_n$.
- Suppose $k_1, \ldots, k_n$ are non-negative integers with the property that $k_1 + \cdots + k_n = k$.
- The number of words of length k in the alphabet $x_1, \ldots, x_n$ which contain the letter x_j exactly k_j times equals the multinomial coefficient
$$\binom{k}{k_1, \ldots, k_n}.$$
- One can interpret this as a generalized version of stars and bars. In stars and bars we have an alphabet with two letters: star and bar. Now we have an alphabet with n letters. We could say we have n types of stars.

Example 14.1. How many anagrams are there for the word
$$MATHEMATICS.$$
Equivalently, what is the number of words of length 11 which contain 2 each of the letters M,A,T, and 1 each of the letters H,E,I,C,S. The answer is the multinomial coefficient
$$\binom{11}{2,2,2,1,1,1,1,1} = \frac{11!}{2!2!2!} = 4989600.$$

Example 14.2. Enumerate all words in the alphabet x_1, x_2, x_3 corresponding to the multinomial coefficient
$$\binom{4}{2,1,1} = \frac{4!}{2!1!1!} = 12.$$
We want to enumerate words of length 4 so that x_1 appears twice, x_2 appears once, and x_3 appears once. There are twelve such words, so we should be systematic in our enumeration (else it might get messy). There are $6 = \binom{4}{2}$ ways to choose the two letters in the word which are x_1:
$$x_1x_1--, \quad x_1-x_1- \quad ,-x_1x_1- \quad ,-x_1-x_1 \quad ,x_1--x_1 \quad ,--x_1x_1.$$
For each of these, there are two ways to choose which of the remaining spaces are x_2:
$$x_1x_1x_2x_3, \quad x_1x_2x_1x_3 \quad ,x_2x_1x_1x_3 \quad ,x_2x_1x_3x_1 \quad ,x_1x_2x_3x_1 \quad ,x_2x_3x_1x_1.$$
$$x_1x_1x_3x_2, \quad x_1x_3x_1x_2 \quad ,x_3x_1x_1x_2 \quad ,x_3x_1x_2x_1 \quad ,x_1x_3x_2x_1 \quad ,x_3x_2x_1x_1.$$
If you think about how we have chosen to enumerate these words, you will see that we are using the combinatorial interpretation of the identity
$$\binom{4}{2,1,1} = \binom{4}{2}\binom{2}{1}.$$

Another combinatorial application of multinomial coefficients is counting lattice paths. As we saw, binomial coefficients can be used to count paths in a 2 dimensional lattice. Multinomial coefficients can be used to generalize this result to arbitrary dimensions.

Example 14.3. Consider the lattice of points in $\mathbb{R}^n$ with integer coordinates. We want to count the number of lattice paths from $(0, \dots, 0)$ to $(k_1, \dots, k_n)$. We only allow steps in our lattice paths to be in the direction of standard basis vectors,

$$e_1 = (1, \dots, 0), \dots, e_n = (0, \dots, 1).$$

Every such path must contain k_j steps in the direction e_j, for each $j = 1, \dots n$. Writing down the steps in a lattice path, we obtain a word of length $k = k_1 + \cdots + k_n$ in the alphabet $e_1, \dots, e_n$. This is a bijection between the set of all such lattice paths and the set of all words of length k in the $e_1, \dots, e_n$ such that e_j occurs k_j times for each $j = 1, \dots n$. Thus, the total number of lattice paths is

$$\binom{k}{k_1, \dots, k_n}.$$

Example 14.4. Consider a $3 \times 3 \times 3$ rubics cube shaped grid in $\mathbb{R}^3$. How many lattice paths are there between the opposite vertices of the cube, $(0, 0, 0)$ and $(3, 3, 3)$? The answer is

$$\binom{3+3+3}{3, 3, 3} = \frac{9!}{3!3!3!} = 1680.$$

14.4 Pascal's recurrence

Recall that binomial coefficients satisfy the recurrence relation

$$\binom{k}{k_1} = \binom{k-1}{k_1-1} + \binom{k-1}{k_1}.$$

This can be re-written using notation for multinomial coefficients as

$$\binom{k}{k_1, k_2} = \binom{k-1}{k_1-1, k_2} + \binom{k-1}{k_1, k_2-1}$$

where $k_2 = k - k_1$. The generalization of Pascal's recurrence for arbitrary multinomial coefficients has a similar form:

$$\binom{k}{k_1, \dots, k_n} = \binom{k-1}{k_1-1, \dots, k_n} + \cdots + \binom{k-1}{k_1, \dots, k_n-1}.$$

One could write this with sigma notation as

$$\binom{k}{k_1, \dots, k_n} = \sum_{j=1}^{n} \binom{k-1}{k_1, \dots, k_j-1, \dots, k_n}.$$

For trinomial coefficients ($n = 3$) this identity is

$$\binom{k}{k_1, k_2, k_3} = \binom{k-1}{k_1 - 1, k_2, k_3} + \binom{k-1}{k_1, k_2 - 1, k_3} + \binom{k-1}{k_1, k_2, k_3 - 1}.$$

How do we prove this identity? The proof is very similar to the combinatorial proof of Pascal's recurrence for binomial coefficients. We consider what happens to a dude named Bob.

Proof. Let's use the candy distribution interpretation of multinomial coefficients. The multinomial coefficient

$$\binom{k}{k_1, \ldots, k_n}$$

on the LHS of Pascal's recurrence counts the number of ways to distribute n types of candy to k friends so that each friend gets one piece each and the number of friends who get type j is k_j (for $j = 1, \ldots n$).

For the RHS of Pascal's recurrence, suppose one of the friends is named Bob. Condition on the type of candy given to Bob. If Bob gets candy j then there are

$$\binom{k-1}{k_1, \ldots, k_j - 1, \ldots, k_n}$$

ways to distribute the candy to the remaining friends. The theorem follows by the addition principle. $\square$

The same proof could be described using the interpretation counting words of length k in an alphabet with n letters.

14.5 The multinomial theorem

The binomial theorem is a special case of the multinomial theorem.

Theorem 14.5. *Let n, k be integers with $n \geq 1$ and $k \geq 0$. Then for any numbers $x_1, \ldots, x_n$,*

$$(x_1 + \cdots + x_n)^k = \sum_{k_1 + \cdots + k_n = k} \binom{k}{k_1, \ldots, k_n} x_1^{k_1} \ldots x_n^{k_n}.$$

The sigma notation used in the statement of the multinomial theorem may appear somewhat confusing. For $n = 3$ and $k = 2$:

$$\begin{aligned}(x_1 + x_2 + x_3)^2 &= \binom{2}{2,0,0} x_1^2 + \binom{2}{0,2,0} x_2^2 + \binom{2}{0,0,2} x_3^2 \\ &\quad + \binom{2}{1,1,0} x_1 x_2 + \binom{2}{1,0,1} x_1 x_3 + \binom{2}{0,1,1} x_2 x_3.\end{aligned}$$

Proof. The proof has two steps.

Step 1 (expand): Expanding the product on the LHS, we have

$$\begin{aligned}(x_1 + \cdots + x_n)^k &= (x_1 + \cdots + x_n) \ldots (x_1 + \cdots + x_n) \\ &= \sum_{i_1,\ldots,i_k \in [n]} x_{i_1} \ldots x_{i_k}.\end{aligned}$$

The resulting sum consists of all words of length k in the alphabet $x_1, \ldots, x_n$.

Step 2 (group): Grouping like terms, we have that

$$\sum_{i_1,\ldots,i_k \in [n]} x_{i_1} \ldots x_{i_k} = \sum_{k_1+\cdots+k_n=k} C^k_{k_1,\ldots,k_n} x_1^{k_1} \ldots x_n^{k_n}$$

for some coefficients $C^k_{k_1,\ldots,k_n}$. We just need to show that

$$C^k_{k_1,\ldots,k_n} = \binom{k}{k_1, \ldots, k_n}.$$

The coefficient of $C^k_{k_1,\ldots,k_n}$ is the number of terms in the sum

$$\sum_{i_1,\ldots,i_k \in [n]} x_{i_1} \ldots x_{i_k}$$

which equal $x_1^{k_1} \ldots x_n^{k_n}$. In other words, $C^k_{k_1,\ldots,k_n}$ is the number of anagrams of the word

$$x_1^{k_1} \ldots x_n^{k_n} = x_1 \ldots x_1 \ldots x_n \ldots x_n.$$

The number of anagrams of this word is precisely the multinomial coefficient

$$\binom{k}{k_1, \ldots, k_n}. \qquad \square$$

Corollary 14.6.

$$n^k = \sum_{k_1+\cdots+k_n=k} \binom{k}{k_1, \ldots, k_n}$$

Proof 1. Set $x_1 = \cdots = x_n = 1$ in the multinomial theorem. $\square$

Proof 2. The LHS equals the number of words of length k in the alphabet $\{x_1, \ldots, x_n\}$. It remains to show that the RHS also equals the number of such words.

Condition the set of all such words on the number of times each letter occurs. The number of words which contain x_j exactly k_j times, $j \in [n]$ is the binomial coefficient

$$\binom{k}{k_1, \ldots, k_n}.$$

The result follows by the addition principle. $\square$

14.6 Practice problems

1. How To Count, Exercises 3.6.16, 3.6.19, 3.6.20, 3.6.21.

15 Principle of Inclusion and Exclusion

15.1 Principle of Inclusion and Exclusion, version 1

All sets are finite.

- Recall the subtraction principle tells us that if $B \subset A$, then $|B| = |A| - |A \setminus B|$.
- Example: The number of permutations of n such that 1 is not sent to 1 is equal to
$$n! - (n-1)! = (n-1)(n-1)!$$
since there are $n!$ permutations of n and there are $(n-1)!$ permutations of n which send 1 to 1.
- The following theorem generalizes the addition principle to the case where the two sets might intersect.

Theorem 15.1 (Principle of Inclusion and Exclusion, version 1)**.** *Let A and B be finite sets. Then,*

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

We can see this two ways. Both are described in How To Count, Section 1.3. We give an example.

Example 15.2. Count lattice paths from $(0,0)$ to $(5,4)$ which pass through $(2,1)$ or $(2,3)$.

The number of lattice paths which pass through $(2,1)$ is

$$\binom{3}{2}\binom{6}{3}.$$

The number of lattice paths which pass through $(2,3)$ is

$$\binom{5}{2}\binom{4}{2}.$$

The number of lattice paths which pass through both $(2,1)$ and $(2,3)$ is

$$\binom{3}{2}\binom{2}{0}\binom{4}{2}.$$

By the principle of inclusion and exclusion (version 1), the total number of paths through $(2,1)$ or $(2,3)$ is

$$\binom{3}{2}\binom{6}{3} + \binom{5}{2}\binom{4}{2} - \binom{3}{2}\binom{2}{0}\binom{4}{2}.$$

15.2 Principle of Inclusion and Exclusion, version 2

What if we want to compute the cardinality of a union of three finite sets, A_1, A_2, A_3? We can do this by repeatedly applying the facts from the previous section.

$$\begin{aligned}
|A_1 \cup A_2 \cup A_3| &= |(A_1 \cup A_2) \cup A_3| \\
&= |A_1 \cup A_2| + |A_3| - |(A_1 \cup A_2) \cap A_3| \\
&= |A_1| + |A_2| - |A_1 \cap A_2| + |A_3| - |(A_1 \cap A_3) \cup (A_2 \cap A_3)| \\
&= |A_1| + |A_2| - |A_1 \cap A_2| + |A_3| - (|A_1 \cap A_3| + |A_2 \cap A_3| - |A_1 \cap A_2 \cap A_3|) \\
&= |A_1| + |A_2| + |A_3| \\
&\quad - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| \\
&\quad + |A_1 \cap A_2 \cap A_3|.
\end{aligned}$$

Thus, we arrive at

Theorem 15.3 (Principle of Inclusion and Exclusion, version 2)**.** *Let A_1, A_2, A_3 be finite sets. Then,*

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$

A slightly different way to see this is to consider how many times both sides of the equation count an element $x \in A_1 \cup A_2 \cup A_3$. If x is contained in exactly one of the A_i, then we get

$$1 = 1 - 0 + 0.$$

If x is contained in two of the A_i, then we get

$$1 = 2 - 1 + 0.$$

If x is contained in all three of the sets A_i, then we get

$$1 = 3 - 3 + 1.$$

Notice anything familiar? We are getting rows of Pascal's triangle!

Suppose A_i are subsets of a set X. Recall the complement of a subset $A \subset X$ is $A^c = X \setminus A$. By DeMorgan's Law and version 2 of P.I.E., we see that

$$\begin{aligned}
|A_1^c \cap A_2^c \cap A_3^c| &= |(A_1 \cup A_2 \cup A_3)^c| \\
&= |X| - |A_1 \cup A_2 \cup A_3| \\
&= |X| - \sum_{i=1}^{3} |A_i| + \sum_{i \neq j} |A_i \cap A_j| - |A_1 \cap A_2 \cap A_3|.
\end{aligned}$$

Example 15.4. How many ways are there to place n labelled balls into 3 labelled boxes so that no box is empty?

In other words, how many surjective functions $f\colon [n] \to [3]$ are there?

Let X be the set of all possible ways of putting the balls into the boxes (including cases where some boxes are empty). Then $|X| = 3^n$.

For $i = 1, 2, 3$, let A_i be the set of all ways such that box i is empty. Then $|A_i| = 2^n$.

The intersections $A_i \cap A_j$, $i \neq j$, consist of ways where all the balls go into the remaining box. Thus, $|A_i \cap A_j| = 1$.

The intersection $A_1 \cap A_2 \cap A_3$ is empty.

The set of all ways where none of the boxes are empty is $A_1^c \cap A_2^c \cap A_3^c$. Vombining with the formula above, we see that

$$|A_1^c \cap A_2^c \cap A_3^c| = 3^n - 3 \cdot 2^n + 3 \cdot 1 - 0.$$

15.3 Principle of Inclusion and Exclusion, version 3

Given finite sets $A_1, \ldots, A_n$, let

$$c_k = \sum_{\{i_1,\ldots,i_k\} \subset [n]} |A_{i_1} \cap \cdots \cap A_{i_k}|.$$

Theorem 15.5. *Let $A_1, \ldots, A_n$ be finite sets. Then,*

$$|A_1 \cup \cdots \cup A_n| = c_1 - c_2 + \cdots + (-1)^{n-1} c_n.$$

We can prove this two ways. One would be a proof by induction which uses steps somewhat similar to how we prove version 2 using version 1. Let's look at a more combinatorial way to see this.

Proof. We want to show that the RHS counts each element $x \in A_1 \cup \cdots \cup A_n$ exactly once. The idea is going to be similar to the second proof which we saw for version 2.

Suppose that the number of A_i containing x is m. Then for each $1 \leq k \leq n$, the sum c_k counts x exactly $\binom{m}{k}$ times. The number of times the RHS counts x is given by the formula

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

Using one of the identities from last week, we see that

$$1 = \binom{m}{0} = \sum_{k=1}^{m} (-1)^{k+1} \binom{m}{k}.$$

(notice here that we are again getting rows of Pascal's triangle, as in the case of version 2.) $\square$

Similar to before, we also have by DeMorgan's Law that

$$\begin{aligned}|A_1^c \cap \cdots \cap A_n^c| &= |(A_1 \cup \cdots \cup A_n)^c| \\ &= |X| - |A_1 \cup \cdots \cup A_n| \\ &= |X| - (c_1 - c_2 + \cdots + (-1)^{n-1} c_n) \\ &= |X| - c_1 + c_2 - \cdots + (-1)^n c_n.\end{aligned}$$

We revisit the example of counting ways to put labelled balls into labelled boxes so that no box is empty.

Theorem 15.6. *The number of surjective functions $f\colon [n] \to [m]$ equals*

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

Sketch of proof: Let X be the set of all functions $f\colon [n] \to [m]$. Then $|X| = m^n$.

Let A_i be the set of functions so that i is not in the image. The set we are interested in is $A_1^c \cap \cdots \cap A_n^c$.

All k-fold intersections have

$$|A_{i_1} \cap \cdots \cap A_{i_k}| = (m-k)^n.$$

Since there are $\binom{m}{k}$ possible k-fold intersections,

$$c_k = \sum_{\{i_1,\ldots,i_k\} \subset [n]} |A_{i_1} \cap \cdots \cap A_{i_k}| = \sum_{\{i_1,\ldots,i_k\} \subset [n]} (m-k)^n = \binom{m}{k}(m-k)^n.$$

Plugging this into the formula,

$$\begin{aligned}|A_1^c \cap \cdots \cap A_n^c| &= |X| - c_1 + c_2 - \cdots + (-1)^n c_n \\ &= m^n - \sum_{k=1}^{m} (-1)^{k+1} \binom{m}{k} (m-k)^n \\ &= m^n + \sum_{k=1}^{m} (-1)^{k} \binom{m}{k} (m-k)^n \\ &= \sum_{k=0}^{m} (-1)^{k} \binom{m}{k} (m-k)^n.\end{aligned}$$

□

15.4 practice problems

1. How to count, exercises 1.3.14, 1.3.16, 7.2.6, 7.2.7, 7.3.5, 7.3.6.

16 More combinatorial numbers (and related distribution problems)

16.1 Stirling numbers of the second kind

- Last time we derived a formula for the number of ways to place n labelled balls into k labelled boxes so that no box is empty.
- The Stirling number of the second kind $\left\{ {n \atop k} \right\}$ counts the number of ways to place n labelled balls into k *unlabelled* boxes so that no box is empty. The textbook uses the notation $S(n,k)$.
- For $n \geq 1$, $\left\{ {n \atop n} \right\} = 1$, $\left\{ {n \atop 0} \right\} = 0$, $\left\{ {0 \atop 0} \right\} = 1$, and $\left\{ {n \atop k} \right\} = 0$ if $k > n$.
- Since there are $k!$ ways to label k boxes, it follows from the formula from the previous lecture that

$$k!\left\{ {n \atop k} \right\} = \sum_{i=0}^{k} (-1)^k \binom{k}{i} (k-i)^n.$$

- Stirling numbers of the second kind satisfy the recurrence

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

  To prove this, choose a favourite ball and condition on whether it is placed in a box with other balls or by itself.
- Another way to think about Stirling numbers of the second kind: they count the number of ways to partition a set of n elements into k subsets.

16.2 Bell numbers

- The Bell number $B(n)$ counts the number of ways to place n labelled balls into n unlabelled boxes. Note that some boxes may be empty.
- Equivalently, $B(n)$ counts the number of ways to partition a set of n elements.
- The number of ways to factor a product of n distinct primes $p_1 \cdots p_n$ is $B(n)$.
- The number of possible rhyming schemes for a poem with n lines is $B(n)$.
- $B(0) = 1$, $B(1) = 1$, $B(2) = 2$, $B(3) = 5$.
- If we condition on the number of non-empty boxes, then we see that

$$B(n) = \sum_{k=1}^{n} \left\{ {n \atop k} \right\}.$$

Thus, the Bell numbers are just the sums of rows in the triangle for Stirling numbers of the second kind.

- The Bell numbers satisfy the recurrence

$$B(n) = \sum_{k=0}^{n-1} \binom{n-1}{k} B(k) \quad \text{for all } n > 0.$$

To see this, condition on the number k of balls not put into the same box as ball 1. Note that $0 \le k \le n-1$. The number of ways to choose the k balls not in the same box as ball 1 is $\binom{n-1}{k}$. The number of ways to place those balls into boxes of their own is $B(k)$. By the multiplication principle, the total number of ways with k balls not in the same box as ball 1 is $\binom{n-1}{k}B(k)$.

16.3 Partition numbers

- Recall that a multiset is like a set but elements are not required to be distinct. e.g. $\{1, 1\}$ is a multiset but not a set.
- A *partition of* n *into* k *parts* is a multiset of k positive integers whose sum equals n. The number of partitions of n into k parts is denoted $p(n, k)$.
- $p(5, 3) = 2$. The two partitions are $\{1, 1, 3\}$ and $\{2, 2, 1\}$.
- As with sets, multi-sets do not have an order. If we care about order in our partitions then the answer is completely different. In the previous example we would consider $1 + 1 + 3 = 5$ as different from $1 + 3 + 1 = 5$. The number of ordered partitions n into k parts is $\binom{n-1}{k}$ (it's just a version of stars and bars).
- Note that $p(0, 0) = 1$ by convention. For $n \ge 1$, $p(n, 0) = 0$, $p(n, 1) = 1$, $p(n, n) = 1$, and $p(n, k) = 0$ for $k > n$.
- Equivalently, $p(n, k)$ is the number of ways to put n unlabelled balls into k unlabelled boxes so that no box is empty. Indeed, since the balls and boxes are unlabelled, all that matters is the multiset of numbers of balls in each box.
- The partition numbers satisfy the recurrence

$$p(n, k) = p(n-1, k-1) + p(n-k, k).$$

To see this, condition on whether a partition contains 1. If it does, then there are $p(n-1, k-1)$ ways to complete the partition $\{1, , \ldots, \}$ to a partition of n. If it does not, then subtracting 1 from each number in the partition produces a partition of $n-k$ into k parts, the total number of which is $p(n-k, k)$.

- The number of ways to place n unlabelled balls into k boxes (some boxes may be empty) is
$$\sum_{i=1}^{k} p(n,i).$$
- In particular, the number of ways to place n unlabelled balls into n unlabelled boxes is
$$p(n) = \sum_{i=1}^{n} p(n,i).$$
Equivalently, this is the total number of partitions of n. The first few values of $p(n)$, beginning with $p(0)$, are
$$1, 1, 2, 3, 5, 7, 11, 15, 22, 30, 42, 56, 77, 101, \ldots$$

16.4 Practice problems

1. How to count, exercises: 4.3.18,4.3.19, 4.3.20.

17 Derangements and the gift exchange problem

17.1 The gift exchange problem

- n people want to have a gift exchange.
- Each person gives a gift to 1 person.
- Each person receives exactly 1 gift.
- No one receives their own gift.

17.2 Derangements

- We can model a gift exchange as a function $f\colon [n] \to [n]$ where $f(i) = j$ if person i gives their gift to person j.
- Since each person receives exactly one gift, this function is a bijection. So it is a permutation of n.
- The condition that no one receives their own give translates to the condition that $f(i) \neq i$ for all $i \in [n]$.
- We call a permutation with this property a *derangement*.
- Some more terminology: if $f(i) = i$, then i is said to be a *fixed point* of the permutation f. Thus, derangements are permutations which have no fixed points.

17.3 Counting derangements

- Let A_i be the set of all permutations of n with $f(i) = i$.
- The set of derangements of n is $A_1^c \cap \dots A_n^c$.
- Thus, we count derangements using the principle of inclusion and exclusion:
- For any k, the cardinality of a k-fold intersection $A_{i_1} \cap \cdots \cap A_{i_k}$ is the set of all permutations of n which fix $i_1, \dots, i_k$. Thus, $|A_{i_1} \cap \cdots \cap A_{i_k}| = (n-k)!$.
- Since there are $\binom{n}{k}$ possible k-fold intersections,

$$c_k = \sum |A_{i_1} \cap \cdots \cap A_{i_k}| = \sum (n-k)! = \binom{n}{k}(n-k)! = \frac{n!}{k!}.$$

- Plugging this into the principle of inclusion and exclusion, we have that the number of derangements of n is

$$
\begin{aligned}
|A_1^c \cap \dots A_n^c| &= |X| + \sum_{k=1}^{n} (-1)^k c_k \\
&= n! + \sum_{k=1}^{n} (-1)^k \frac{n!}{k!} \\
&= \sum_{k=0}^{n} (-1)^k \frac{n!}{k!}.
\end{aligned}
$$

17.4 Practice problems

1. How to count, exercises: 2.3.5, 7.3.4, 7.3.7.

18 Introduction to generating functions

18.1 Introduction to generating functions

If $(a_n)_{n=0}^{\infty}$ is a sequence of numbers, then the power series

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

is the *(ordinary) generating function* of $(a_n)_{n=0}^{\infty}$. (The word ordinary will be used to distinguish from other types of generating function which we will see later.)

Example 18.1. Recall the Fibonacci numbers are the sequence $(F_n)_{n=0}^{\infty}$ defined by the recurrence $F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$. The generating function of the Fibonacci numbers is

$$\sum_{n=0}^{\infty} F_n x^n.$$

Example 18.2. Let $(p_n)_{n=1}^{\infty}$ denote the sequence of prime numbers $2, 3, 5, 7, 11, \dots$ The generating function of the prime numbers is

$$\sum_{n=1}^{\infty} p_n x^n.$$

Example 18.3. A variation on the problem of counting partitions: we can restrict the numbers which can be used in the partition. For example:

- Let a_n denote the number of partitions of n using only 2. Then $a_n = 1$ if n is even and $a_n = 0$ otherwise.
- Let b_n denote the number of partitions of n using only 3. Then $b_n = 1$ if n is a multiple of 3, otherwise $b_n = 0$..

The generating functions of these sequences are

$$\sum_{n=0} x^{2n} = 1 + x^2 + x^4 + \dots, \quad \sum_{n=0} x^{3n} = 1 + x^3 + x^6 + \dots.$$

How many ways are there to partition n using 2 and 3? What's remarkable about generating functions is that if we expand the product

$$\sum_{n=0}^{\infty} d_n x^n = \left(\sum_{n=0} x^{2n}\right)\left(\sum_{n=0} x^{3n}\right)$$

$$= (1 + x^2 + x^4 + x^6 + x^8 + \dots)(1 + x^3 + x^6 + x^9 + \dots)$$
$$= 1 + x^2 + x^3 + x^4 + x^5 + 2x^6 + 2x^6 + x^7 + 2x^8 + \dots$$

then the coefficient of x^n in the resulting power series equals the number of partitions of n using 2 and 3. For example, if $n = 6$, then there are two partitions: 2+2+2 and 3+3. Correspondingly, the term x^6 above has coefficient 2 since $x^6 = x^{2+2+2}$ and $x^6 = x^{3+3}$. For another example, there are two partitions of 8: $2+2+2+2$ and $2+3+3$. Correspondingly, the coefficient of x^8 above is 2 since $x^8 = x^{2+2+2+2}$ and $x^8 = x^{2+3+3}$.

We can also use calculus to study generating functions. Recall that

$$\sum_{n=0} x^n = \frac{1}{1-x}$$

(we will discuss radius of convergence later). It follows that

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

Thus, we can express our generating function as a bonafide function

$$f(x) = \frac{1}{1-x^2}\frac{1}{1-x^3}.$$

In particular, the sequence can be recovered from f by differentiating:

$$d_n = \frac{1}{n!}f^{(n)}(0).$$

Thus, generating functions give us solutions to complex combinatorial problems using algebra and calculus. This is one of the motivations for studying generating functions.

18.2 Practice problems

1. Write the generating function for each of the following counting problems.
   (a) Count the number of partitions of n using $1, 2, 3$.
   (b) Count the number of partitions of n using the positive integers 1 to 10.
   (c) Count the number of ways to make change for n cents using pennies (1 cent), nickles (5 cents), dimes (10 cents), and quarters (25 cents).

19 Power series

19.1 Formal power series

A *formal power series* is an expression

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \dots$$

where $a_0, a_1, a_2, \dots$ is a sequence of numbers and x is a variable. In the terminology of the previous lecture, it is the generating function of the sequence $(a_n)_{n=0}^{\infty}$.

Formal power series are algebraic objects. Let

$$A = \sum_{n=0}^{\infty} a_n x^n, \quad B = \sum_{n=0}^{\infty} b_n x^n$$

be formal power series. We say that $A = B$ if $a_n = b_n$ for all $n \in \mathbb{N}$.

We define the sum and product of formal power series to be

$$A + B := \sum_{n=0}^{\infty} (a_n + b_n) x^n,$$

$$A \cdot B := \sum_{n=0}^{\infty} c_n x^n, \quad c_n := \sum_{k=0}^{n} a_k b_{n-k}.$$

The sequence $(c_n)_0^{\infty}$ is the *convolution* of $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$. We can also formally differentiate (and integrate) formal power series:

$$\frac{d}{dx} A := \sum_{n=1}^{\infty} a_n n x^{n-1} = \sum_{n=0}^{\infty} a_{n+1} (n+1) x^n.$$

Although these operations are purely algebraic, they have all the properties one expects. For example, one can show that differentiation of formal power series satisfies a product rule (see practice problems).

Example 19.1. Consider the geometric series

$$G = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \dots$$

The formal derivative of G is

$$\frac{d}{dx} G = \sum_{n=1}^{\infty} n x^{n-1} = \sum_{n=0}^{\infty} (n+1) x^n.$$

The convolution of the sequence $1, 1, 1, 1, \dots$ with itself is

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

Thus, the formal product of G with itself is

$$G^2 = G \cdot G = \sum_{n=0}^{\infty} (n+1)x^n.$$

In particular, the geometric series satisfies the following identity:

$$\frac{d}{dx} G = G^2.$$

19.2 Power series

In calculus, the expression

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

has an analytic meaning: it is the limit $\lim_{N\to\infty} S_N(x)$ of the partial sums

$$S_N(x) = \sum_{n=0}^{N} a_n x^n = a_0 + a_1 x + \cdots + a_N x^N.$$

It is a theorem of calculus that there exists a number $r \geq 0$ such that

$$\lim_{N\to\infty} S_N(x) = \begin{cases} \text{exists} & \text{if } |x| < r, \text{ and} \\ \text{D.N.E.} & \text{if } |x| > r. \end{cases}$$

The number r is the *radius of convergence* and the interval $(-r, r)$ is the *interval of convergence*. On the interval of convergence, this limit defines a function

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad x \in (-r, r).$$

In this case,

$$a_n = \frac{f^{(n)}(0)}{n!}.$$

In other words, a_n are the Taylor coefficients of f at $x = 0$ and $\sum_{n=0}^{\infty} a_n x^n$ is the Taylor series of f at $x = 0$ (also known as the MacLaurin series).

Example 19.2. The radius of convergence of

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

is $r = 0$ (one can check this with the ratio test). Thus, it does not define a function. Still, it is a formal power series/generating function.

Example 19.3. Consider the geometric series

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \dots$$

Its partial sums are

$$S_N(x) = \sum_{n=0}^{N} x^n = 1 + x + \dots + x^N.$$

Observe that

$$\begin{aligned} S_N(x) - xS_N(x) &= 1 - x^{N+1} \\ (1-x)S_N(x) &= 1 - x^{N+1} \\ S_N(x) &= \frac{1 - x^{N+1}}{1-x}. \end{aligned}$$

The radius of convergence is $r = 1$ since

$$\lim_{N\to\infty} \frac{1 - x^{N+1}}{1-x} = \begin{cases} \dfrac{1}{1-x} & \text{if } |x| < 1, \text{ and} \\ \text{D.N.E.} & \text{if } |x| > 1. \end{cases}$$

Thus, the geometric series defines the function

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \qquad x \in (-1, 1).$$

In conclusion, we can do operations on power series two ways: as formal power series, or (provided $r > 0$) as functions. In a calculus course, one learns that operations on functions and operations on formal power series coincide (on the interval of convergence). As we will see, this allows us to use techniques from calculus to study generating functions.

19.3 Examples

Example 19.4. A polynomial is an expression of the form

$$p(x) = a_0 + a_1 x + \dots + m x^m.$$

For example, the generating function of the binomial coefficients is the polynomial

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

The radius of convergence of any polynomial is $r = \infty$, so manipulating polynomials as functions or as formal power series is the same.

Example 19.5. Recall that the number of ways to distribute n identical balls into k distinct boxes is

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

We can consider the generating functions of these numbers for fixed values of k. For $k = 1$,

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

For $k = 2$,

$$\sum_{n=0}^{\infty} \left(\!\!\binom{2}{n}\!\!\right) x^n = \sum_{n=0}^{\infty} \binom{n+1}{n} x^n = \sum_{n=0}^{\infty} (n+1) x^n = \frac{1}{(1-x)^2}.$$

In general,

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

Outline of proof. One can prove this by induction on k. The base case $k = 1$ is already given above.
Induction step: Suppose we know that

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

Differentiating both sides, we get

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

On the RHS, we can shift the indices, apply an identity for binomial coefficients (we leave this identity as an exercise):

$$\begin{aligned}
\sum_{n=1}^{\infty} \binom{k+n-1}{n} n x^{n-1} &= \sum_{n=0}^{\infty} \binom{k+n}{n+1} (n+1) x^n \\
&= \sum_{n=0}^{\infty} \binom{k+n}{k} k x^n \\
&= \sum_{n=0}^{\infty} \left(\!\!\binom{k+1}{n}\!\!\right) k x^n
\end{aligned}$$

Thus,

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

Cancelling the factor of k on both sides completes the proof. $\square$

19.4 Practice problems

1. Show that differentiation of formal power series satisfies the product rule. i.e. if
$$A = \sum_{n=0}^{\infty} a_n x^n, \quad B = \sum_{n=0}^{\infty} b_n x^n$$
are formal power series, then
$$\frac{d}{dx}(AB) = \left(\frac{d}{dx}A\right)B + A\left(\frac{d}{dx}B\right).$$

2. Prove the following identity for binomial coefficients,
$$\binom{k+n}{k}k = \binom{k+n}{n+1}(n+1).$$
Challenge: can you give a combinatorial proof?

3. Use generating functions for binomial coefficients to prove Vandermonde's identity,
$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}.$$

4. The power A^k of a formal power series is defined by multiplying A by itself n times. Use the definition of multiplication of formal power series and induction to show that
$$A^k = \sum_{n=0}^{\infty} \left(\sum_{i_1+\cdots+i_k=n} a_{i_1}\cdots a_{i_k} \right) x^n.$$

5. Pick up a calculus textbook and review the MacLaurin series of common functions.

19.5 Bonus practice problems

Prove the following useful facts.

1. For any formal power series $A = \sum_{n=0}^{\infty} a_n x^n$,
$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k \right) x^n = A \cdot G$$
where G is the geometric series.

2. For any formal power series $A = \sum_{n=0}^{\infty} a_n x^n$,
$$\sum_{n=0}^{\infty} d_n x^n = (1-x) \cdot A$$
where $d_0 = a_0$ and $d_n = a_n - a_{n-1}$ for $n \geq 1$.

3. The harmonic numbers are $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$. Show that

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

Hint: a) What is the MacLaurin series of $\ln\left(\frac{1}{1-x}\right)$*? b) Try to use one of the earlier practice problems.*

20 Products of generating functions

20.1 Multiplying functions and power series

Last time we recalled how operations on power series and on functions are the same (provided we work on the interval of convergence). One of the most useful instances of this in combinatorics is multiplication: If

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad g(x) = \sum_{n=0}^{\infty} b_n x^n,$$

then

$$(f \cdot g)(x) = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) x^n.$$

In particular, the Taylor coefficients of the product of two functions are

$$\frac{(f \cdot g)^{(n)}(0)}{n!} = \sum_{k=0}^{n} a_k b_{n-k}.$$

It is a nice calculus exercise to check this for yourself if you haven't seen a proof before.

20.2 Proving identities by multiplying generating functions

We give two examples of combinatorial identities (one of which we already proved combinatorially) which can be proven very easily by multiplying generating functions (and applying the theorem mentioned above).

Theorem 20.1 (Vandermonde's identity)**.**

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

Proof. Using the generating function for the the coefficients $\binom{m+p}{n}$ (which we derived in the previous lecture),

$$\begin{aligned}
\sum_{n=0}^{\infty} \binom{m+p}{n} x^n &= (1+x)^{m+p} \\
&= (1+x)^m (1+x)^p \\
&= \left(\sum_{n=0}^{\infty} \binom{m}{n} x^n\right)\left(\sum_{n=0}^{\infty} \binom{p}{n} x^n\right) \\
&= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} \binom{m}{k}\binom{p}{n-k}\right) x^n.
\end{aligned}$$

The identity follows by comparing coefficients. □

Recall the coefficients $\left(\!\binom{m}{n}\!\right)$ from last lecture (and from our discussion of stars and bars).

Theorem 20.2.

$$\left(\!\!\binom{m}{n}\!\!\right) = \sum_{k=0}^{n} \left(\!\!\binom{m-1}{k}\!\!\right)$$

Proof. Using the generating function for the coefficients $\left(\!\binom{m}{n}\!\right)$ (which we derived in the previous lecture),

$$\begin{aligned}
\sum_{n=0}^{\infty} \left(\!\!\binom{m}{n}\!\!\right) x^n &= \frac{1}{(1-x)^m} \\
&= \frac{1}{(1-x)^{m-1}} \frac{1}{1-x} \\
&= \left(\sum_{n=0}^{\infty} \left(\!\!\binom{m-1}{n}\!\!\right) x^n \right) \left(\sum_{n=0}^{\infty} x^n \right) \\
&= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \left(\!\!\binom{m-1}{k}\!\!\right) \right) x^n.
\end{aligned}$$

The identity follows by comparing coefficients. □

20.3 The convolution principle

There is a simple reason why multiplying generating functions can produce proofs of combinatorial identities: the coefficients of the product of two generating functions have a combinatorial interpretation! This interpretation can be summarized as the following principle.[3]

The convolution principle: *If a_n counts the number of ways to select n red objects and b_n counts the number of ways to select n blue objects, then*

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

counts the number of ways to select n red or blue objects. In other words, the generating function for the problem of selecting n red or blue objects equals the product

$$\left(\sum_{n=0}^{\infty} a_n x^n \right) \left(\sum_{n=0}^{\infty} b_n x^n \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) x^n.$$

The proof behind every instance of the convolution principle is the same: If we condition on the number k of red objects that are chosen, then for case k there

[3]Note that we do not state this principle as a theorem (what we mean by "select" in this statement is purposefully vague). Rather we are indicating the pattern of a general argument which works in many cases. This is illustrated in the following discussion.

are a_k ways to select the red objects, b_{n-k} ways to select the blue objects, and, by the multiplication principle, $a_k b_{n-k}$ ways in total. The convolution principle follows by applying the addition principle. For example, this is exactly how we proved Vandermonde's identity (in that case, our mode of selection is choosing without replacement).

This principle also explains the second proof above. In that case our mode of selection is choosing with replacement. Let's discuss this in a bit more detail.

Previously, we defined the coefficient $\left(\!\!\binom{m}{n}\!\!\right)$, m *multichoose* n, as the number of ways to distribute n unlabelled balls into m labelled boxes (with no restrictions on the number of balls in each box). Equivalently, $\left(\!\!\binom{m}{n}\!\!\right)$ is the number of ways to choose n elements from the set $[m]$ with replacement (order is not important): the number of times we choose an element i from the set $[m]$ is simply the number of balls we put in the box labelled i.

If we colour the numbers $1, \ldots, m-1$ red and we colour m blue, then convolution principle tells us that the number of ways to choose n elements with replacement from $[m]$ is

$$\sum_{k=0}^{n} \left(\!\!\binom{m-1}{k}\!\!\right) \left(\!\!\binom{1}{n-k}\!\!\right).$$

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

The utility of the convolution principle goes beyond explaining the logic behind our proofs. As we will see next, it gives us a way to construct generating functions for complicated counting problems.

20.4 Practice problems

1. Prove the identity from the beginning of the lecture by checking that

$$\frac{(f \cdot g)^{(n)}(0)}{n!} = \sum_{k=0}^{n} a_k b_{n-k}.$$

   Hint: use induction and the product rule. (This is a calculus exercise, not a combinatorics exercise.)

2. Prove the following identity by multiplying generating functions.

$$\left(\!\!\binom{m+p}{n}\!\!\right) = \sum_{k=0}^{n} \left(\!\!\binom{m}{k}\!\!\right) \left(\!\!\binom{p}{n-k}\!\!\right).$$

21 Applications of generating functions

21.1 Constructing generating functions for counting problems

In this section we looked at How To Count, Section 5.3, Examples 5.3.4 and 5.3.5.

21.2 The method of generating functions

In this section we looked at the example of deriving a formula for the Fibonacci numbers using the method of generating functions. This example is given in detail in How To Count, Section 6.2.

21.3 Practice problems

1. How To Count, Exercises 5.3.9, 5.3.10, 5.3.11, 5.3.14, 5.3.15, 5.3.16.
2. How To Count, Exercises 5.3.19, 5.3.20, 5.3.21.
3. How To Count, Exercises 6.2.5, 6.2.6, 6.2.9.

22 Exponential generating functions

22.1 Exponential generating functions

- The *exponential generating function* of a sequence a_n is the power series

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

- Example 1: By the binomial theorem,

$$(1+x)^m = \sum_{n=0}^{\infty} \binom{m}{n} x^n = \sum_{n=0}^{\infty} \frac{m!}{n!(m-n)!} x^n = \sum_{n=0}^{\infty} \frac{P(m,n)}{n!} x^n.$$

  Thus $(1+x)^m$ is the exponential generating function of the numbers $P(m,n)$ (m place n).

- Example 2: Recall the geometric series

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

  Thus, $\frac{1}{1-x}$ is the exponential generating function of $a_n = n!$.

- Example 3: Recall from calculus that

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

  Thus e^x is the exponential generating function of the constant sequence $a_n = 1$.

22.2 Rules for exponential generating functions

- Let

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n, \quad g(x) = \sum_{n=0}^{\infty} \frac{b_n}{n!} x^n$$

  be the exponential generating function of sequences a_n and b_n.

- Rule 1:

$$f^{(n)}(0) = a_n.$$

- Rule 2: The derivative $f'(x)$ is the generating function of the shifted sequence a_{n+1}.

Proof.

$$f'(x) = \sum_{n=1}^{\infty} \frac{na_n}{n!} x^{n-1} = \sum_{n=1}^{\infty} \frac{a_n}{(n-1)!} x^{n-1} = \sum_{n=0}^{\infty} \frac{a_{n+1}}{n!} x^n.$$

□

- Rule 3 (the convolution rule): The product $f(x)g(x)$ is the exponential generating function of the sequence

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

Proof.

$$\begin{aligned} f(x)g(x) &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} \frac{a_k b_{n-k}}{k!(n-k)!} \right) x^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{\infty} \frac{n! a_k b_{n-k}}{k!(n-k)!} \right) x^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{\infty} \binom{n}{k} a_k b_{n-k} \right) x^n \end{aligned}$$

□

22.3 Example: Bell numbers

Recall the Bell numbers $B(n)$. They satisfy the recurrence relation

$$B(0) = B(1) = 1, \quad B(n) = \sum_{k=0}^{n-1} \binom{n-1}{k} B(k).$$

It's exponential generating function is

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

We derive a closed form for $f(x)$ as follows. By Rule 2:

$$f'(x) = \sum_{n=0}^{\infty} \frac{B(n+1)}{n!} x^n.$$

Applying the recurrence relation, this becomes:

$$f'(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} B(k) \right) x^n.$$

By Rule 3 and Example 3 above,

$$f'(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} B(k) \right) x^n = f(x)e^x.$$

Thus, the function $f(x)$ satisfies the differential equation

$$f'(x) = f(x)e^x.$$

Solutions of this differential equation are of the form

$$f(x) = ce^{e^x}.$$

It remains to solve for the coefficient c. By Rule 1 (and the initial value for our recurrence relation), we know that

$$1 = B(0) = f(0) = ce^{e^0} = ce \Longrightarrow c = e^{-1}.$$

Thus,

$$f(x) = e^{e^x - 1}.$$

22.4 Practice problems

1. Read the proof of How To Count, Theorem 6.2.4 which derives the formula

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

   for the exponential generating function of the derrangement numbers D_n. (We derived a formula for the numbers D_n earlier in the course using the P.I.E.)

2. Prove the following theorem using exponential generating functions.

   Theorem 22.1 (Binomial inversion)**.** *If two sequences $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are related by the formula*

   $$a_n = \sum_{k=0}^{n} \binom{n}{k} b_k,$$

   then

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

   This theorem is called "inversion" because a formula for a_n in terms of b_n is "inverted" into a formula for the coefficients b_n in terms of the sequence a_n.

23 Solving linear recurrence relations

23.1 Classifying recurrence relations

Many recurrence relations can be solved systematically. We describe which recurrence relations can be solved systematically using which methods by classifying them with various properties.

- Linear: A recurrence relation is *linear* if it can be written in the form

$$c_k R_{n+k} + c_{k-1} R_{n+k-1} + \cdots + c_1 R_{n+1} + c_0 R_n = g(n)$$

  where c_i, $g(n)$ are functions which are allowed to depend on n, but are not allowed to depend on elements R_j of the sequence.

- Homogeneous: A linear recurrence relation is *homogeneous* if $g(n) = 0$.

- Constant coefficients: we say that the recurrence relation has *constant coefficients* if the c_i do not depend on n.

Finally, the *order* of a linear recurrence relation of the form

$$c_k R_{n+k} + c_{k-1} R_{n+k-1} + \cdots + c_1 R_{n+1} + c_0 R_n = g(n)$$

is k. One needs k initial conditions to solve a recurrence relation of order k.

23.2 The method of characteristic polynomials

The method of characteristic polynomials can solve any linear, homogeneous recurrence relation with constant coefficients. The general algorithm for this method is described in How To Count, Theorem 6.3.2. We illustrate this method with an example.

Example 23.1 (see How To Count, Example 6.3.3 for additional details)**.** Solve the recurrence relation

$$R_n = 7R_{n-1} - 16R_{n-2} + 12R_{n-3}, \quad R_0 = 7, R_1 = 16, R_2 = 40.$$

This is a linear, homogeneous recurrence relation with constant coefficients of order 3. First, we rewrite the recurrence in the form

$$R_{n+3} - 7R_{n+2} + 16R_{n+1} - 12R_n = 0.$$

From this form, we write the *characteristic polynomial* of the recurrence relation, which is a polynomial in a variable (we use w) of order 3,

$$w^3 - 7w^2 + 16w - 12 = 0.$$

This polynomial factors as $(w-2)^2(w-3)$. The general method of characteristic polynomials then tells us the general solution of the recurrence relation has the form

$$R_n = A2^n + Bn2^n + C3^n.$$

where A, B, C are some coefficients which we need to solve for. We solve for these coefficients using the initial conditions. Plugging in the initial conditions gives us a system of three linear equations in three unknowns:

$$\begin{aligned} 7 &= A + C \\ 16 &= 2A + 2B + 3C \\ 40 &= 4A + 8B + 9C \end{aligned}$$

Solving this system yields $A = 3$, $B = -1$, and $C = 4$. Thus,

$$R_n = 3 \cdot 2^n - n \cdot 2^n + 4 \cdot 3^n.$$

23.3 Why it works

It's nice to have this general algorithm for solving recurrence relations: if you ever encounter this sort of problem in the wild you can easily get a computer to solve it for you. Of course, as mathematicians, we are more interested in why this method works.

We illustrate why it works using the example from the previous section:

$$R_n = 7R_{n-1} - 16R_{n-2} + 12R_{n-3}, \quad R_0 = 7, R_1 = 16, R_2 = 40.$$

Let

$$y(x) = \sum_{n=0}^{\infty} \frac{R_n}{n!} x^n$$

be the exponential generating function of this sequence. Recall that differentiating an exponential power series results in a shift of the coefficients:

$$y'(x) = \sum_{n=0}^{\infty} \frac{R_{n+1}}{n!} x^n.$$

Combining this fact with the recurrence relation, we get a linear homogeneous differential equation for the function y:

$$y''' - 7y'' + 16y' - 12y = 0.$$

Now, you may have learned in calculus that these sorts of differential equations can be solved by the *method of characteristic polynomials* (huh, sounds familiar!). The characteristic polynomial of this differential equation is

$$w^3 - 7w^2 + 16w - 12 = (w-2)^2(w-3) = 0$$

(the same as the polynomial we got in the example above). The method of characteristic polynomials for differential equations tells us that the general solution has the form

$$y = ae^{2x} + bxe^{2x} + ce^{3x}.$$

where a, b, c are coefficients. Since y is an exponential generating function for the sequence R_n, initial conditions for the recurrence relation give us initial conditions for the differential equation:

$$y(0) = R_0 = 7, \quad y'(0) = R_1 = 16, \quad y''(0) = R_2 = 40.$$

Using these initial conditions, we can solve for the coefficients. Doing this, we get $a = 3$, $b = -2$, $c = 4$. Thus,

$$y(x) = 3e^{2x} - 2xe^{2x} + 4e^{3x}.$$

Ok, so we found a function. Why is this related to the solution of our recurrence relations. Well, if we expand the exponential functions as power series, then we see that:

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{R_n}{n!} x^n &= y(x) \\
&= 3e^{2x} - 2xe^{2x} + 4e^{3x} \\
&= 3\sum_{n=0}^{\infty} \frac{2^n}{n!} x^n - 2x\sum_{n=0}^{\infty} \frac{2^n}{n!} x^n + 4\sum_{n=0}^{\infty} \frac{3^n}{n!} x^n \\
&= \sum_{n=0}^{\infty} \frac{3 \cdot 2^n}{n!} x^n - \sum_{n=0}^{\infty} \frac{2^{n+1}}{n!} x^{n+1} + \sum_{n=0}^{\infty} \frac{4 \cdot 3^n}{n!} x^n \\
&= \sum_{n=0}^{\infty} \frac{3 \cdot 2^n}{n!} x^n - \sum_{n=0}^{\infty} \frac{n 2^n}{n!} x^n + \sum_{n=0}^{\infty} \frac{4 \cdot 3^n}{n!} x^n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(3 \cdot 2^n - n \cdot 2^n + 4 \cdot 3^n\right) x^n
\end{aligned}$$

Equating coefficients, we recover the solution of our recurrence relation:

$$R_n = 3 \cdot 2^n - n \cdot 2^n + 4 \cdot 3^n.$$

This principle is pervasive in many other techniques for solving recurrence relations. The following two sections of the textbook, 6.4 and 6.5, cover additional techniques for solving recurrence relations. These techniques, which we will not go into in the course, also derive from techniques for solving differential equations via exponential generating functions.

23.4 Practice problems

1. How To Count, Exercises 6.3.10, 6.3.11, 6.3.12.

24 Existence and proof by contradiction

24.1 Existence problems in combinatorics

Until this point, we have focused on finding exact solutions to counting problems. For the remainder of the course we will focus on a new type of problem: existence problems. An existence problem is a question of the form "Does there exist X" where X is a mathematical object of some kind, often with a list of constraints (if this sounds vague bear with me, we will see some concrete examples in a moment).

In general, the phrase "there exists" means "there is at least 1". This is in contrast with questions we dealt with until now which ask "how many are there."

As we will see in a moment, there are two common strategies for solving existence problems:

- Constructive proofs.
- Proof by contradiction.

These are new to us. In the course of exploring existence problems, we will also explore these two proof strategies.

24.2 Example: perfect covers

A *perfect cover* of a chessboard is a tiling by dominos. There is a lot of flexibility in studying perfect covers: one can make the chessboard whatever shape they like. For each shape, the most elementary question one can ask is "Does there exist a perfect cover?"

Example 24.1. Just to make sure we understand what a perfect cover is, observe that there are exactly 2 perfect covers of the 2×2 chessboard: one with vertical dominos and one with horizontal dominos.

Example 24.2. Does there exist a perfect cover of the standard 8×8 chessboard.

Solution: The answer is yes. The question is: how do we prove it? In order to show there is at least one perfect cover, all we need to do is produce an example. (note that this does not tell us how many perfect covers there are, only that there is at least one). Let's describe how to produce a perfect cover in words (in the lecture we drew a picture, which is easier for this particular problem).

Each column is a 8×1 chessboard. Each column has a perfect cover using 4 dominos (it's easy to build this perfect cover yourself). Tiling all the columns in this way produces a perfect cover of the 8×8 chessboard. $\square$

The solution above is an example of a constructive proof: to show that a perfect cover exists, we build, or construct, an example. That's all there is to it. Now let's consider a slightly different problem.

Example 24.3. Does there exist a perfect cover of the 9×9 chessboard.

Solution: The answer is no. How do we prove it? Observe that the number of squares on a 9×9 chessboard is $9^2 = 81$, an odd number. However, in a perfect cover each domino covers 2 squares. Thus, if there was a perfect cover, then the number of squares would have to be even. This is false, so there does not exist a perfect cover. $\square$

The solution above is an example of a proof by contradiction. We will discuss this strategy in more detail later in the lecture. For now, let's see more examples:

Example 24.4. Does there exist a perfect cover of the 8×8 chessboard if 1 square is deleted.

Solution: No. If we delete a square then the resulting board has $8^2 - 1 = 63$ squares, an odd number. By the same logic as the previous example, there cannot be a perfect cover. $\square$

Example 24.5. Does there exist a perfect cover of the 8×8 chessboard if 2 squares are deleted.

Solution: Now it depends. For example, suppose we delete two adjacent corner squares. Then it is possible to construct a perfect cover (see the lecture video for an example, or try to construct one for yourself).

On the other hand, if we delete two opposite corner squares, then the answer is "no." We leave this case as a problem to think about before the next lecture.

$\square$

24.3 Proof by contradiction

Proof by contradiction is a fundamental strategy for existence problems. Every proof by contradiction has the following format:

Theorem 24.6. *P is true.*

Proof. We give a proof by contradiction. Assume that P is false.[4]

... (Use the assumption to prove something we know is impossible. This is called a contradiction.)

Thus, our assumption was false. So P is true. $\square$

The middle part is where the work happens. We illustrate this with a theorem.

Theorem 24.7. *Let n be an odd number. There does not exist a perfect cover of the $n \times n$ chessboard.*

[4] Another popular way to open a proof by contradiction is to say "Assume for the sake of contradiction that P is false."

Our statement P is "There does not exist a perfect cover of the $n \times n$ chessboard." It's negation is the statement "There exists a perfect cover of the $n \times n$ chessboard."

Proof. Assume for the sake of contradiction that there exists a perfect cover of the $n \times n$ chessboard.

Let N be the number of dominos in a perfect cover. Then the number of squares is $2N$.

$$\begin{aligned} &\Rightarrow n^2 = 2N \\ &\Rightarrow n^2 \text{ is even} \\ &\Rightarrow n \text{ is even.} \end{aligned}$$

But n is odd, so this is a contradiction. Thus, our assumption was false. There does not exist a perfect cover of the $n \times n$ chessboard. □

24.4 Practice problems

1. There are two very famous examples of proof by contradiction that everyone with an undergraduate math degree should know: the proof that $\sqrt{2}$ is irrational and Euler's proof that there are infinitely many primes. These proofs are given in your textbook (Proposition 1.2.3 and 1.2.4). Read both proofs and answer the consider the following questions:

   0

   - What is the existence problem being answered in each proof? (phrasing the latter as an existence problem, "there is at least one", as discussed in the lecture is slightly more difficult)
   - Are the answers to the two existence problems the same or different?

2. How To Count, Exercises 1.2.8, 1.2.9, 1.2.10.

3. Does there exist a perfect cover of the 8×8 chessboard if two opposite corners have been deleted? Justify your answer. *Hint: squares on a chessboard are often coloured black and white.*

25 The pigeonhole principle

25.1 The pigeonhole principle

We covered Theorem 1.5.1 (the pigeonhole principle), Proposition 1.5.4, Theorem 1.5.5 (the generalized pigeonhole principle), and Proposition 1.5.6.

25.2 Practice problems

1. How To Count, Exercises 1.5.7 - 1.5.10.

26 Incidence structures and block designs

26.1 Incidence structures

An *incidence structure* consists of two sets, X and Y and an incidence relation $\in$ on the Cartesian product $X \times Y$. It is often useful to think of the elements y of an incidence relation as subsets of X and $x \in y$ as the relation "x is contained in y." Incidence structures are also commonly known as *hypergraphs*.

Example 26.1. Let $X = \{x_1, x_2, x_3, x_4, x_5\}$ and $Y = \{y_1, y_2, y_3, y_4\}$ where

$$y_1 = \{x_1, x_2, x_4\}, \quad y_2 = \{x_1, x_3, x_5\}, \quad y_3 = \{x_1, x_4\}, \quad y_4 = \{x_3, x_5\}.$$

One convenient way to represent the data of an incidence structure on finite sets X and Y is with an *incidence matrix*, a $v \times b$ binary matrix I with entries

$$a_{i,j} = \begin{cases} 1 & \text{if } x_i \in y_j \\ 0 & \text{else.} \end{cases}$$

For example, the incidence matrix of the incidence structure above is,

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \end{array}$$

Although an incidence structure does not depend on an ordering of X and Y, an incidence matrix does depend on a choice of order. A different indexing of X and Y produces a different incidence matrix which differs by permuting the rows and columns.

Example 26.2. Let X be the set of points in the plane $\mathbb{R}^2$, let Y be the set of lines in $\mathbb{R}^2$ and define $x \in \ell$ if the point x is contained in the line ℓ.

26.2 Regular block designs

Block designs are another word for incidence relations that satisfy certain additional constraints. In the language of block designs, the elements of X are *treatments* or *varieties* and the elements of Y are *blocks*. To my knowledge, most of the naming conventions for block designs originate in experimental design.

The *size* of a block y the number of varieties x with $x \in y$. We assume all the blocks in a block design have a fixed size, k. In other words, the columns of the incidence relation all sum to k. For example, a *graph* is just a block design with block size $k = 2$, e.g.

$$\begin{array}{c|c} & \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \end{array}$$

(When talking about graphs we typically call the varieties *vertices* and the blocks *edges*.)

The degree of a variety x is the number of blocks y with $x \in y$. The degree of x_i is the sum of the entries in row i of the incidence matrix,

$$deg(x_i) = \sum_{j=1}^{b} a_{i,j}.$$

where $I = (a_{i,j})$ is the incidence matrix.

Theorem 26.3. *For every block design,*

$$bk = \sum_{i=1}^{v} deg(x_i).$$

Proof. The proof follows by observing that the entries of the incidence matrix can be summed by rows or by columns:

$$bk = \sum_{j=1}^{b} k = \sum_{j=1}^{b} \sum_{i=1}^{v} a_{i,j} = \sum_{i=1}^{v} \sum_{j=1}^{b} a_{i,j} = \sum_{i=1}^{v} deg(x_i). \qquad \square$$

The sum of the entries of the incidence matrix equals the number of pairs $(x, y) \in X \times Y$ such that $x \in y$. The proof above is counting all such pairs two ways: by counting x incident to each y and by counting y incident to each x.

For example, this recovers a well-known fact about graphs.

Corollary 26.4. *For every graph,*

$$2b = \sum_{i=1}^{v} deg(x_i).$$

A block design is *regular* (or *r-regular*) if there is a number r with such that every x has $deg(x) = r$.

Corollary 26.5. *For every r-regular block design,*

$$bk = vr.$$

In particular, every r-regular graph with v vertices and b edges has $2b = vr$.

Example 26.6 (How To Count, Example 10.2.2)**.** Does there exist a regular block design with $v = 10$, $r = 1$, $b = 5$, $k = 2$? If so, represent it as a graph.

Call the varieties $A, B, C, D, E, F, G, H, I, J$. Then

$$AB, CD, EF, GH, IJ$$

is a block design with $v = 10$, $r = 1$, $b = 5$, $k = 2$.

Example 26.7 (How To Count, Example 10.2.2)**.** Does there exist a regular block design with $v = 5$, $r = 6$, $b = 10$, $k = 3$?

Example 26.8. Does there exist a regular block design with $v = 5$, $r = 6$, $b = 11$, $k = 4$? No. If there was, then we would have that

$$44 = bk = vr = 30.$$

But this is false, so there does not exist a block design with these parameters.

27 Balanced Incomplete Block Designs (BIBDs)

A block design is *balanced* (or λ*-balanced*) if every pair of varieties x, x' occurs in the same block exactly λ times. A block design is *complete* if all v varieties occur together in a block. Regular complete block design have $k = v$, $b = r$ and are quite uninteresting. Thus we are interested in block designs which are not complete, or *incomplete*.

Definition 27.1. A (v, b, r, k, λ)*-BIBD* is a block design which is:

- r-regular,
- λ-balanced, and
- incomplete, with
- v varieties, b blocks, and block size k.

Theorem 27.2. *Every* (v, b, r, k, λ)*-BIBD has*

$$r(k - 1) = (v - 1)\lambda.$$

Proof. Form a submatrix of the incidence matrix as follows. First, delete all the columns whose last entry is 0. Then delete the last row. Call the resulting matrix S. We make several observations:

- S has $v - 1$ rows, $1, \ldots, v - 1$.
- Since $deg(x_v) = r$, the matrix S has r columns, $j_1, \ldots, j_r$.
- The sum of each column of S is $k - 1$.

- The sum of each row of S equals λ.

Putting these observations together,

$$r(k-1) = \sum_{k=1}^{r}(k-1) = \sum_{k=1}^{r}\sum_{i=1}^{v-1} a_{i,j_k} = \sum_{i=1}^{v-1}\sum_{k=1}^{r} a_{i,j_k} = \sum_{i=1}^{v-1}\lambda = (v-1)\lambda. \quad \square$$

Example 27.3. There does not exist a $(v, b, 3, 4, \lambda)$-BIBD for any v odd. If there did, then we would have that

$$9 = r(k-1) = \lambda(v-1).$$

If v is odd, then $v-1$ is even, so this equation implies that $9 = \lambda(v-1)$ is even. This is impossible.

Note that this is also a double counting proof in disguise. We have fixed x_v and are counting all pairs (x, y) such that $x \in y$, $x_v \in y$, and $x \neq x_v$.

It follows from this theorem (and the theorem from the previous lecture) that for every (v, b, r, k, λ)-BIBD,

$$r = \lambda\frac{v-1}{k-1} \quad b = \lambda\frac{v(v-1)}{k(k-1)}.$$

As a result, it is common to call them (v, k, λ)-BIBD's.

The two conditions derive in the theorem above are necessary for the existence of a BIBD but not sufficient! There exist examples of tuples (v, k, λ) for which there exists no (v, k, λ)-BIBD. According to Wikipedia, one such example is $(v = 43, k = 7, \lambda = 1)$.

27.1 Practice problems

1. How To Count, Exercises 10.2.4 – 10.2.7.
2. Fisher's inequality says that for every (v, b, r, k, λ)-BIBD, $b \geq v$. Can you prove it?

28 Steiner Triple Systems

28.1 Kirkman's schoolgirls

In 1850, the British mathematician Thomas Kirkman posed the following problem in The Lady's and Gentleman's diary:

> Fifteen schoolgirls walk to school each day in 5 groups of 3. Can they be arranged so that each pair of girls walks together exactly once?

The problem is asking about the existence of a block design with $v = 15$, $k = 3$, and $\lambda = 1$ with some additional constraints. Each girl walks with two other girls every day and has 14 other girls with whom she must walk. Thus, the problem must be solved in 7 days. The total number of blocks is therefore

$$b = 7 \cdot 5 = 35.$$

Since every girl walks to school each day, the block design must be regular with $r = 7$.

Rather than solve this existence problem right now, we first solve some similar but simpler examples:

Example 28.1. Nine schoolgirls walk to school each day in 3 groups of 3. Can they be arranged so that each pair of girls walks together exactly once?

By the same logic as above, the number of days required is 4 and the total number of blocks is thus $b = 12$.

Thinking carefully, one can construct the following example (the letters A–I represent schoolgirls).

Example 28.2. Six schoolgirls walk to school each day in 2 groups of 3. Can they be arranged so that each pair of girls walks together exactly once?

No solution exists. If the first day is

$$ABC \quad DEF,$$

then there are no options for the second day: A, B and C must walk in separate groups, but there are only two groups to walk in.

28.2 Steiner Triple Systems

Definition 28.3. A *Steiner Triple System* (or $STS(v)$) is a λ-balanced incomplete block design with fixed block size $k = 3$ and $\lambda = 1$.

For example, a solution to Kirkman's schoolgirl problem, if one exists, is a $STS(15)$. The following lemma tells us that every $STS(v)$ is a $(v, 3, 1)$-BIBD. It follows that

$$r = \frac{v-1}{2}, \quad b = \frac{v(v-1)}{6}.$$

Lemma 28.4. *Every $STS(v)$ is r-regular with $r = \frac{v-1}{2}$. In particular, every $STS(v)$ is a $(v, 3, 1)$-BIBD.*

Proof. Let x be a variety in a $STS(v)$. Since $\lambda = 1$, all the other $v - 1$ varieties occur in a block with x exactly once. Since $k = 3$, each block containing x contains 2 other varieties. Thus,

$$deg(x) = \frac{v-1}{2}.$$

Thus the $STS(v)$ is r-regular with $r = \frac{v-1}{2}$. $\square$

So what are the possible values of v such that there exists a $STS(v)$? The previous theorem allows us to eliminate some possiblities.

- Since $r = \frac{v-1}{2}$ is an integer, v must be odd.
- Since $b = \frac{v(v-1)}{6}$ is an integer, $v(v-1)$ must be divisible by 6. So $v = 0, 1, 3,$ or 4 mod 6. Since v must be odd by the previous item, we have that $v = 1$ or 3 mod 6.

Thus, the only possibilities are:

$$3, 7, 9, 13, 15, ...$$

It's easy to see there is a $STS(3)$. We already saw an example of a $STS(9)$ above.

Example 28.5. The Fano plane is a $STS(7)$.

The next possible values of v are 13 and 15. There do exist $STS(13)$ and $STS(15)$ (How To Count, Examples 10.3.2, 10.3.4). In fact, one can prove the following

Theorem 28.6 (How To Count, 10.3.5)**.** *For all $v \geq 3$, there exists a $STS(v)$ if and only if $v = 1$ or 3 mod 6.*

The proof is constructive, but a bit complicated. We won't cover it, but you can take a look at it in all its glory in the textbook.

28.3 Practice problems

1. How To Count, Exercises 10.1.4, 10.1.5, 10.1.6.

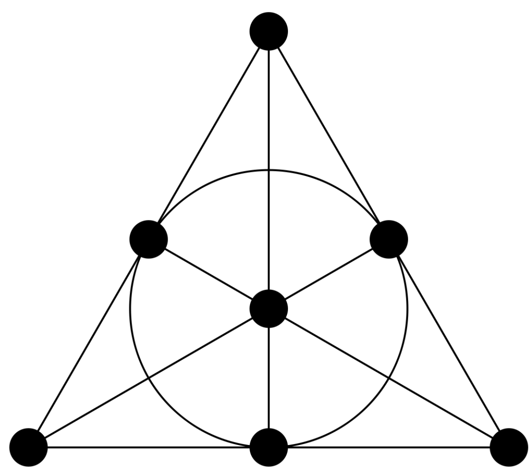

Figure 4: The Fano Plane. Source: Wikipedia.

$2j+1, i+j+11k+2]$, $[i, i+2j+2, i+j+8k+2]$, $[i, i+2j+3k+2, i+j+9k+2]$, and $[i, i+2j+3k+1, i+j+6k+1]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24k+2$.

Suppose that $v \equiv 7 \pmod{24}$. Thus, there exists $k \in \mathbb{N}$ such that $v = 24k+7$. We use the blocks $[i, i+2k+1, i+8k+3]$ for $i = 0, 1, \ldots, 24k+6$. If $k \geq 1$, then we use the additional blocks $[i, i+2j+1, i+j+11k+4]$, $[i, i+2j+2, i+j+8k+4]$, $[i, i+2j+3k+3, i+j+9k+4]$, and $[i, i+2j+3k+2, i+j+6k+3]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24+6$.

Suppose that $v \equiv 9 \pmod{24}$. Thus, there exists $k \in \mathbb{N}$ such that $v = 24k+9$. The case where $v = 9$ is done in Example 10.3.3. Hence, we can assume that $k \geq 1$. For $\ell = 0, 1, \ldots, 8k+2$, use the blocks $[\ell, \ell+8k+3, \ell+16k+6]$. For $k \geq 1$, we use the additional blocks $[i, i+2k-1, i+5k+2]$, $[i, i+3k, i+12k+3]$, $[i, i+3k+1, i+12k+5]$, $[i, i+2j+3k+2, i+j+9k+5]$, and $[i, i+2j+3k+5, i+j+6k+4]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24k+8$. If $k \geq 2$, then we additionally use the blocks $[i, i+2j+1, i+j+11k+4]$ and $[i, i+2j+2, i+j+8k+4]$ for $j = 0, 1, \ldots, k-2$ and $i = 0, 1, \ldots, 24k+8$.

Suppose that $v \equiv 13 \pmod{24}$. Thus, there exists $k \in \mathbb{N}$ such that $v = 24k+13$. If $k \geq 0$, then we use the blocks $[i, i+2k+1, i+8k+4]$ and $[i, i+3k+2, i+12k+7]$ for $i = 0, 1, \ldots, 24k+12$. If $k \geq 1$, then we use the additional blocks $[i, i+2j+1, i+j+11k+6]$, $[i, i+2j+2, i+j+8k+5]$, $[i, i+2j+3k+4, i+j+9k+6]$, and $[i, i+2j+3k+3, i+j+6k+4]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24k+12$.

Suppose that $v \equiv 15 \pmod{24}$. Thus, there exists $k \in \mathbb{N}$ such that $v = 24k+15$. For $\ell = 0, 1, \ldots, 8k+4$, use the blocks $[\ell, \ell+8k+5, \ell+16k+10]$. Additionally, we use the blocks $[i, i+3k+2, i+12k+8]$ and $[i, i+2j+3k+3, i+j+6k+4]$ for $j = 0, 1, \ldots, k$ and $i = 0, 1, \ldots, 24k+14$. If $k \geq 1$, then we use the additional blocks $[i, i+2j+1, i+j+11k+7]$, $[i, i+2j+2, i+j+8k+6]$, and $[i, i+2j+3k+4, i+j+9k+7]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24k+14$.

Suppose that $v \equiv 19 \pmod{24}$. Thus, there exists $k \in \mathbb{N}$ such that $v = 24k+19$. If $k \geq 0$, then we use the blocks $[i, i+2k+1, i+8k+5]$, $[i, i+3k+2, i+12k+8]$, and $[i, i+3k+3, i+12k+10]$ for $i = 0, 1, \ldots, 24k+18$. If $k \geq 1$, then we use the additional blocks $[i, i+2j+1, i+j+11k+8]$, $[i, i+2j+2, i+j+8k+6]$, $[i, i+2j+3k+5, i+j+9k+8]$, and $[i, i+2j+3k+4, i+j+6k+5]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24k+18$.

Suppose that $v \equiv 21 \pmod{24}$. Thus, there exists $k \in \mathbb{N}$ such that $v = 24k+21$. For $\ell = 0, 1, \ldots, 8k+6$, use the blocks $[\ell, \ell+8k+7, \ell+16k+14]$. Additionally, use the blocks $[i, i+2j+1, i+j+11k+10]$, $[i, i+2j+3k+3, i+j+9k+8]$, and $[i, i+2j+3k+4, i+j+6k+6]$ for $j = 0, 1, \ldots, k$ and $i = 0, 1, \ldots, 24k+20$. If $k \geq 1$, then use the additional blocks $[i, i+2j+2, i+j+8k+8]$ for $j = 0, 1, \ldots, k-1$ and $i = 0, 1, \ldots, 24k+20$.

∎

Figure 5: Excerpt from the proof in How To Count.

29 Finite projective planes, part I

29.1 Finite projective planes

We followed the material from section 10.4 very closely, up to the end of the proof of Theorem 10.4.3 (the duality principle).

29.2 Practice problems

1. Find 4 points in the Fano plane that satisfy axiom 3 of finite projective planes.

2. Find 4 lines in the Fano plane, no three of which contain the same point.

3. Let P, L be the set of points and lines in the Fano plane. Define a new incidence structure as follows: $P' = L$, $L' = P$ and $\ell \in p$ if $p \in \ell$. Illustrate the resulting incidence structure using Figure 28.5: draw elements of P' at points and elements of L' as lines so that the new incidence structure is encoded in the picture.

30 Finite projective planes, part II

30.1 Finite projective planes

After discussing the game Spot it! we followed the remainder of section 10.4 from the textbook.

Figure 6: Cards from the game Spot it! form lines in a projective plane of order 7. Each card (line) contains 8 pictures, which represent points in the projective plane. Rules of the game are based on the principle that any two cards have exactly 1 picture in common.

For more information about Spot it!, check out this fun article on the web (which includes a youtube video explaining the game Spot It! for those who are unfamiliar): https://www.smithsonianmag.com/science-nature/math-card-game-spot-it-180970873/

30.2 Practice problems

1. How To Count, Exercises 10.4.12, 10.4.14.

2. Write down the details of the proofs that the correspondences described in the proofs of Theorem 10.4.4 and Theorem 10.4.5 are bijections.

3. Read Theorem 10.4.10 and Example 10.4.11 in the textbook, then try to solve Exercise 10.4.15.

4. A much more tedious, unofficial way to play Spot It! is the following. Place all the cards face up on the table. Pick two pictures at random from the list of 57 pictures. The first person to find the card with both pictures on it wins.

   (a) For each pair of pictures there is exactly one card with both on it. Explain why this is using the axioms of finite projective planes.

   (b) In fact, this version of the game does not quite work because a pack of Spot it! cards only contains 55 cards, not 57. Explain why this is an issue for our new unofficial version of the game.

   (c) How come the official way to play Spot it! works, despite there being 2 cards missing? What if we removed more cards from the deck? Would the game still work?