Math 222 Fall 2022, Lecture 13: Introduction

website: `https://www.cip.ifi.lmu.de/~grinberg/t/22fco`

1. Introduction (cont'd)

1.7. Counting permutations: an introduction

1.7.4. The one-line notation

If σ is a permutation of the set $[n]$ (for some $n \in \mathbb{N}$), then we can write σ in two-line notation as follows:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

But the top row here carries no useful information: We already know that the elements of $[n]$ are $1, 2, \ldots, n$, and it is not exactly unreasonable to list them in increasing order. Thus, we can omit it, and write only the bottom row, i.e., the n-tuple

$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) \in [n]^n.$$

This is the so-called **one-line notation** of σ:

Definition 1.7.8. Let $n \in \mathbb{N}$. Let σ be a permutation of $[n]$. Then, the **one-line notation** of σ means the n-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$.

For instance, the six permutations of $[3]$ have one-line notations

$$(1,2,3), \quad (1,3,2), \quad (2,1,3), \quad (2,3,1), \quad (3,1,2), \quad (3,2,1).$$

These are precisely all possible triples (i.e., 3-tuples) of distinct elements of $[3]$ that contain each element of $[3]$. This pattern persists for "higher values of 3":

Proposition 1.7.9. Let $n \in \mathbb{N}$. Then:

(a) If σ is any permutation of $[n]$, then the one-line notation of σ is an n-tuple of distinct elements of $[n]$ that contains each element of $[n]$.

(b) If $(i_1, i_2, \ldots, i_n)$ is an n-tuple of distinct elements of $[n]$ that contains each element of $[n]$, then there exists a unique permutation σ of $[n]$ such that $(i_1, i_2, \ldots, i_n)$ is the one-line notation of σ.

(c) Let

$$T_n = \{n\text{-tuples of distinct elements of } [n] \text{ that contain each element of } [n]\}.$$

Then, the map

$$\{\text{permutations of } [n]\} \to T_n,$$
$$\sigma \mapsto (\text{one-line notation of } \sigma) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$$

is well-defined and a bijection.

Proof. This is an easy exercise in understanding the definitions (and basic set theory). Most importantly, a map $\sigma : [n] \to [n]$ is uniquely determined by its list of values $(\sigma(1), \sigma(2), \ldots, \sigma(n))$, and furthermore this map σ is

- injective if and only if the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ is a list of **distinct** elements;
- surjective if and only if the list $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ contains each element of $[n]$.

See the 2019 notes (Proposition 1.7.12) for details. □

The main use of the one-line notation (besides being short) is that it allows us to manipulate permutations as tuples. We will eventually see some instances of that (e.g., in Lecture 28, proof of Proposition 4.4.5).

1.7.5. Short-legged permutations

Now let us count a few more classes of permutations.

Definition 1.7.10. Let $n \in \mathbb{N}$. A permutation σ of $[n]$ will be called **short-legged** if each $i \in [n]$ satisfies $|\sigma(i) - i| \leq 1$.

How many such permutations does $[n]$ have?
For example, the short-legged permutations of $[3]$ are (in one-line notation)

$$(1,2,3), \quad (1,3,2), \quad (2,1,3).$$

So there are 3 of them. Based on experiments, we suspect:

Proposition 1.7.11. Let $n \in \mathbb{N}$. Then,

$$(\text{\# of short-legged permutations of } [n]) = f_{n+1}.$$

(Again, the Fibonacci numbers.)

Proof. See the 2019 notes (§1.7.5) or (more fun!) figure it out yourself (what is the relation between short-legged permutations of $[n]$ and other combinatorial objects that are counted by Fibonacci numbers?). □

Exercise 1. How many short-legged derangements does $[n]$ have?

1.7.6. Long-legged permutations

Definition 1.7.12. Let $n \in \mathbb{N}$. A permutation σ of $[n]$ will be called **long-legged** if each $i \in [n]$ satisfies $|\sigma(i) - i| > 1$.

Again, we can ask: How many are there? Here is a table:

n	0	1	2	3	4	5	6	7	8	9
ℓ_n	1	0	0	0	1	4	29	206	1708	15702

,

where ℓ_n denotes the # of long-legged permutations of $[n]$.

On the OEIS, the sequence $(\ell_0, \ell_1, \ell_2, \ldots)$ is Sequence A001883, and you can find two recurrent equations for it, one of which is

$$\begin{aligned}\ell_n = n\ell_{n-1} + 4\ell_{n-2} - 3(n-3)\ell_{n-3} + (n-4)\ell_{n-4} \\ + 2(n-5)\ell_{n-5} - (n-7)\ell_{n-6} - \ell_{n-7} \qquad \text{for all } n \geq 7\end{aligned}$$

(not the nicest formula, but beats checking all the $n!$ permutations of $[n]$ for sure!). This seems to be the most direct expression we have for ℓ_n. No closed-form formula is known (or expected to exist). Not every counting problem has a good answer!

2. Binomial coefficients

Let us continue our exploration of BCs (binomial coefficients) and get a little bit more systematic about proving identities. In particular, we shall (later in this chapter) prove the Chu–Vandermonde identity[1], and derive various consequences from it. We will also learn some methods for proving identities.

2.1. Revisiting the alternating sum of a row in Pascal's triangle

Recall the following fact (Proposition 1.3.23 in Lecture 7):

Proposition 2.1.1. We have

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = [n = 0] \qquad \text{for each } n \in \mathbb{N}.$$

We already proved this (by applying the binomial formula to 1 and -1). We shall now give two more proofs of this proposition. Each proof will exemplify an important technique.

[1]Note: you are already free to use it on homeworks and midterms.

2.1.1. Telescoping sums

The first new proof of Proposition 2.1.1 relies on the **telescoping sum principle** (or, for short, the **telescope principle**):

> **Theorem 2.1.2** (telescoping sum principle)**.** Let u and v be two integers with $u \leq v+1$. Let $a_u, a_{u+1}, \ldots, a_v, a_{v+1}$ be any numbers. Then,
>
> $$\sum_{j=u}^{v} \left(a_{j+1} - a_j\right) = a_{v+1} - a_u.$$

Proof. First of all, the claim is obvious when $u = v+1$, because both sides are 0 in this case. Thus, WLOG assume that $u < v+1$, so the sum on the left is nonempty. Then,[2]

$$\begin{aligned}
&\sum_{j=u}^{v} \left(a_{j+1} - a_j\right) \\
&= \left(a_{u+1} - a_u\right) + \left(a_{u+2} - a_{u+1}\right) + \left(a_{u+3} - a_{u+2}\right) + \cdots + \left(a_{v+1} - a_v\right) \\
&= \left(a_{v+1} - a_v\right) + \left(a_v - a_{v-1}\right) + \left(a_{v-1} - a_{v-2}\right) + \left(a_{v-2} - a_{v-3}\right) + \cdots + \left(a_{u+1} - a_u\right) \\
&\qquad\qquad \text{(here, we have flipped the order of the summands)} \\
&= a_{v+1} + \underbrace{\left(-a_v + a_v\right)}_{=0} + \underbrace{\left(-a_{v-1} + a_{v-1}\right)}_{=0} + \underbrace{\left(-a_{v-2} + a_{v-2}\right)}_{=0} + \cdots + \underbrace{\left(-a_{u+1} - a_{u+1}\right)}_{=0} - a_u \\
&= a_{v+1} - a_u.
\end{aligned}$$

This computation (with the many cancellations involved in it) is reminiscent of a telescope being contracted; thus the name of the theorem. □

Note that the telescoping sum principle appears in many equivalent forms in the literature, e.g., in the forms

$$\sum_{j=u}^{v} \left(a_j - a_{j-1}\right) = a_v - a_{u-1} \qquad \text{and}$$
$$\sum_{j=u}^{v} \left(a_j - a_{j+1}\right) = a_u - a_{v+1}.$$

All such forms can be easily derived from Theorem 2.1.2 (e.g., by substituting a_{j-1} for a_j, or by substituting $-a_j$ for a_j); alternatively, they can all be proved in similar ways.

Note also the similarity between Theorem 2.1.2 and the Fundamental Theorem of Calculus:

$$\int_u^v F'\left(x\right) dx = F\left(v\right) - F\left(u\right).$$

[2] This is slightly informal. See the 2019 notes (§2.1.1) for a formalization of this argument, as well as for a different proof.

In a way, Theorem 2.1.2 is a discrete form of the Fundamental Theorem of Calculus.

Now, we can reprove Proposition 2.1.1:

Proof of Proposition 2.1.1. The claim of Proposition 2.1.1 is clear if $n = 0$ (since both sides are equal to 1 in this case). Thus, we WLOG assume that $n \neq 0$. Then,

$$\begin{aligned}
\sum_{k=0}^{n} (-1)^k \underbrace{\binom{n}{k}}_{\substack{=\binom{n-1}{k-1}+\binom{n-1}{k} \\ \text{(by Pascal's recurrence)}}} &= \sum_{k=0}^{n} (-1)^k \left(\binom{n-1}{k-1} + \binom{n-1}{k} \right) \\
&= \sum_{k=0}^{n} \left(\underbrace{(-1)^k}_{=-(-1)^{k-1}} \binom{n-1}{k-1} + (-1)^k \binom{n-1}{k} \right) \\
&= \sum_{k=0}^{n} \left(-(-1)^{k-1} \binom{n-1}{k-1} + (-1)^k \binom{n-1}{k} \right) \\
&= \sum_{k=0}^{n} \left(\underbrace{(-1)^k \binom{n-1}{k}}_{\text{let's call this } a_k} - \underbrace{(-1)^{k-1} \binom{n-1}{k-1}}_{\text{then this is } a_{k-1}} \right) \\
&= \sum_{k=0}^{n} \left(a_k - a_{k-1} \right) \\
&= a_n - a_{-1} \qquad \text{(by the telescoping sum principle)} \\
&= \underbrace{\binom{n-1}{n}}_{\substack{=0 \\ \text{(since } n-1\in\mathbb{N} \\ \text{and } n-1<n)}} - \underbrace{\binom{n-1}{-1}}_{\substack{=0 \\ \text{(by definition)}}} = 0 - 0 \\
&= 0 = [n = 0] \qquad (\text{since } n \neq 0).
\end{aligned}$$

Thus, Proposition 2.1.1 is proved again. □

Using this second proof, we can show a generalization of Proposition 2.1.1:

Exercise 2 ("reverse hockey-stick identity")**.** Let $n \in \mathbb{R}$. Let $m \in \mathbb{N}$. Prove that

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

(This is Exercise 2.1.1 in the 2019 notes.)

2.1.2. A war between the odd and the even (aka sign-reversing involutions)

Now comes the perhaps nicest proof of Proposition 2.1.1.

Third proof of Proposition 2.1.1. As before, we can easily dispatch the case when $n = 0$. Thus, we assume that $n \neq 0$. Hence, the claim that we must prove simplifies to[3]

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

Can we prove this combinatorially? Not directly, because the LHS[4] includes negative numbers, and negative numbers don't count anything. However, we can resolve this issue in the simplest possible way: We can bring the negative numbers to the RHS. So, the equality (1) (which we intend to prove) gets rewritten as follows:

$$\sum_{k \text{ even}} \binom{n}{k} \stackrel{?}{=} \sum_{k \text{ odd}} \binom{n}{k} \tag{2}$$

(where the summation signs "$\sum\limits_{k \text{ even}}$" and "$\sum\limits_{k \text{ odd}}$" are shorthand for "$\sum\limits_{\substack{k \in \mathbb{N};\\ k \text{ is even}}}$" and "$\sum\limits_{\substack{k \in \mathbb{N};\\ k \text{ is odd}}}$", respectively). Equivalently, it gets rewritten as[5]

$$(\text{\# of even-size subsets of } [n]) \stackrel{?}{=} (\text{\# of odd-size subsets of } [n])$$

(since $\sum\limits_{k \text{ even}} \binom{n}{k}$ is the # of even-size subsets of $[n]$ [6], whereas $\sum\limits_{k \text{ odd}} \binom{n}{k}$ is the # of odd-size subsets of $[n]$).

Now, how can we prove such an equality? One way would be to find a bijection between $\{\text{even-size subsets of } [n]\}$ and $\{\text{odd-size subsets of } [n]\}$. Such a bijection can indeed be found: The map

$$\{\text{even-size subsets of } [n]\} \to \{\text{odd-size subsets of } [n]\},$$
$$A \mapsto \begin{cases} A \setminus \{1\}, & \text{if } 1 \in A; \\ A \cup \{1\}, & \text{if } 1 \notin A \end{cases}$$

(that is, the map that removes the element 1 from any subset that contains 1,

[3]The question mark about the equality sign signifies that this equality has yet to be proved.
[4]"LHS" = "left hand side".
[5]An "**even-size**" set means a finite set of even size. An "**odd-size**" set means a finite set of odd size.
[6]Here, we are using the combinatorial interpretation of BCs.

and inserts 1 into any subset that does not)[7] is well-defined[8] and bijective[9]. Thus, this map is a bijection. Hence, the bijection principle yields that

$$(\text{\# of even-size subsets of } [n]) = (\text{\# of odd-size subsets of } [n]) .$$

As explained above, this equality is just a rewritten form of (2), which in turn is equivalent to (1). Thus, (1) holds, and Proposition 2.1.1 is proved (for the third time). $\square$

The bijection that we used in the above proof is worth saying a few words about. Let me drop the "even-size" and "odd-size" conditions in its definition, and thus extend it to a map

$$\begin{aligned} t : \{\text{subsets of } [n]\} &\to \{\text{subsets of } [n]\} , \\ A &\mapsto \begin{cases} A \setminus \{1\} , & \text{if } 1 \in A; \\ A \cup \{1\} , & \text{if } 1 \notin A. \end{cases} \end{aligned}$$

This map t is called "toggling the element 1", since the element 1 is moved either into or out of the subset. It is not just some bijective map, but actually a map that is inverse to itself (i.e., it satisfies $t \circ t = \mathrm{id}$); such maps are known as **involutions**. It furthermore has the property that it flips the sign $(-1)^{|A|}$ (that is, it satisfies $(-1)^{|t(A)|} = -(-1)^{|A|}$ for every subset A of $[n]$). This makes it a "**sign-reversing involution**". This sign-reversing property makes it suited for canceling addends in alternating sums (i.e., sums in which different addends can have different signs); many alternating-sum identities in combinatorics can be proved using sign-reversing involutions. So the presence of powers of -1 should not scare you away from using combinatorial arguments!

One last remark about the toggling involution t: It is a particular case of a set-theoretical operation called **symmetric difference**. Here its definition:

---

[7] For example, this map sends $\{1,4,5,7\}$ to $\{4,5,7\}$, and sends $\{2,4,5\}$ to $\{1,2,4,5\}$.

[8] Indeed, it changes the size of any subset by 1 upwards or downwards, so an even-size subset will always turn into an odd-size subset.

[9] Indeed, its inverse is the map

$$\begin{aligned} \{\text{odd-size subsets of } [n]\} &\to \{\text{even-size subsets of } [n]\} , \\ A &\mapsto \begin{cases} A \setminus \{1\} , & \text{if } 1 \in A; \\ A \cup \{1\} , & \text{if } 1 \notin A \end{cases} \end{aligned}$$

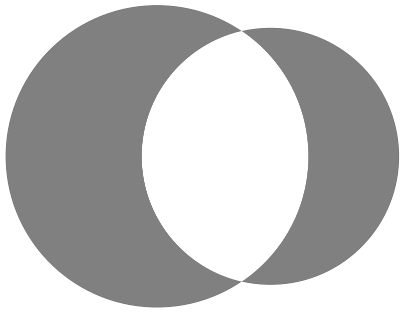
(given by the exact same formula!), because:

- If a subset A of $[n]$ contains 1, then $(A \setminus \{1\}) \cup \{1\} = A$.
- If a subset A of $[n]$ does not contain 1, then $(A \cup \{1\}) \setminus \{1\} = A$.

Definition 2.1.3. Let X and Y be two sets. Then, their **symmetric difference** $X \bigtriangleup Y$ is defined by

$$\begin{aligned} X \bigtriangleup Y &:= \{\text{all elements that belong to } X \text{ or } Y \text{ but not to both}\} \\ &= (X \cup Y) \setminus (X \cap Y) = (X \setminus Y) \cup (Y \setminus X) . \end{aligned}$$

In terms of Venn diagrams, $X \bigtriangleup Y$ is the grey zone in the following Venn diagram (where the two circles are X and Y):



Thus, the definition of our map t can be rewritten as follows:

$$\begin{aligned} t : \{\text{subsets of } [n]\} &\to \{\text{subsets of } [n]\} , \\ A &\mapsto A \bigtriangleup \{1\} . \end{aligned}$$

See the 2019 notes (§2.1.2) for some properties of the symmetric difference, particularly its associative law

$$(A \bigtriangleup B) \bigtriangleup C = A \bigtriangleup (B \bigtriangleup C) .$$