Math 222 Fall 2022, Lecture 4: Introduction

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

1. Introduction (cont'd)

1.2. Sums of powers

1.2.1. The sum $1 + 2 + ... + n$

We now change the topic and recall a famous result (known to the Ancient Greeks, but colloquially known as the "**Little Gauss formula**" due to a famous anecdote about Gauss discovering it in middle school):

Theorem 1.2.1 ("Little Gauss formula")**.** Let $n \in \mathbb{N}$. Then,

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

Keep in mind that $0 \in \mathbb{N}$ by our conventions. So Theorem 1.2.1 applies to $n = 0$ in particular. In this case, the sum $1 + 2 + \cdots + n$ is an empty sum (i.e., it has no addends), and is understood to be 0 by convention. Thus, Theorem 1.2.1 says that $0 = \frac{0(0+1)}{2}$ for $n = 0$.

First proof of Theorem 1.2.1. Induction on n. The details are straightforward and LTTR (= left to the reader). □

Second proof of Theorem 1.2.1. We observe the following general fact: If $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are arbitrary numbers, then

$$\begin{aligned} &(a_1 + a_2 + \cdots + a_n) + (b_1 + b_2 + \cdots + b_n) \\ &= (a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n). \end{aligned}$$

Hence,

$$\begin{aligned} &(1 + 2 + \cdots + n) + (n + (n-1) + \cdots + 1) \\ &= \underbrace{(1+n)}_{=n+1} + \underbrace{(2 + (n-1))}_{=n+1} + \cdots + \underbrace{(n+1)}_{=n+1} \\ &= \underbrace{(n+1) + (n+1) + \cdots + (n+1)}_{n \text{ many } (n+1)\text{'s}} \\ &= n(n+1). \end{aligned}$$

Thus,

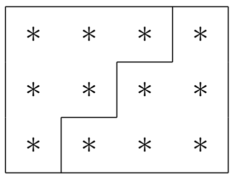
$$\begin{aligned} n(n+1) &= (1+2+\cdots+n) + \underbrace{(n+(n-1)+\cdots+1)}_{=1+2+\cdots+n} \\ &= (1+2+\cdots+n) + (1+2+\cdots+n) \\ &= 2\cdot(1+2+\cdots+n). \end{aligned}$$

Dividing this by 2, we get

$$\frac{n(n+1)}{2} = 1+2+\cdots+n,$$

qed. □

Third proof of Theorem 1.2.1. Here is a picture proof (shown here for the case $n = 3$):



.

This is a 4×3-rectangle[1] (i.e., a rectangle of width 4 and height 3), subdivided into two parts by a broken line which starts in the southwestern corner and winds its way to the northeastern corner, making steps of length 1 eastwards and northwards (by turns). The two parts (the one below and the one above the broken line) have the same area, since they are symmetric to each other with respect to the center of the rectangle. Hence, the area of either part equals half the area of the whole rectangle. Since the area of the whole rectangle is $3 \cdot 4$, we thus conclude that the area of either part equals $\frac{3 \cdot 4}{2}$.

But we can also compute this area in a different way: Let us look at the part below the broken line. This part has 0 squares in the 1-st column[2], 1 square in the 2-nd column, 2 squares in the 3-rd column, and 3 squares in the 4-th column. Hence, in total, it has $0+1+2+3$ many squares. In other words, its area is $0+1+2+3$.

Now we know that the area of the part of the rectangle below the broken line is $0+1+2+3$, but we also know (from before) that it equals $\frac{3 \cdot 4}{2}$. Hence, $0+1+2+3 = \frac{3 \cdot 4}{2}$. In view of $0+1+2+3 = 1+2+3$, this rewrites as $1+2+3 = \frac{3 \cdot 4}{2}$. This is precisely the statement of Theorem 1.2.1 for $n = 3$.

The same argument (but using an $(n+1) \times n$-rectangle instead of a 4×3-rectangle) proves Theorem 1.2.1 for arbitrary n.

[1]Each square of this rectangle is marked with an asterisk.

[2]We count columns from the left.

Now, of course, picture proofs aren't the most convincing thing in mathematics (there is a wonderful 3blue1brown video showing four different fake "picture proofs", with subsequent revelation). Fortunately, the above picture proof is actually a combinatorial proof in disguise, and once we have stripped it down to this combinatorial core, its correctness will be easy to ascertain.

Here is an outline of how to make this picture proof rigorous using combinatorics (see my 2019 notes[3] for more details). As for domino tilings, we think of the $(n+1) \times n$-rectangle as the set $[n+1] \times [n]$ consisting of all pairs (i, j) of integers $i \in [n+1]$ and $j \in [n]$. Set

$$A := \{(i,j) \in [n+1] \times [n] \mid i \leq j\};$$
$$B := \{(i,j) \in [n+1] \times [n] \mid i > j\}.$$

The set A consists of all squares in our rectangle that lie above the broken line on the picture; the set B consists of all squares in our rectangle that lie below that broken line.

The sum rule yields $|[n+1] \times [n]| = |A| + |B|$, since A and B are two disjoint finite sets whose union is $[n+1] \times [n]$. Thus,

$$|A| + |B| = |[n+1] \times [n]| = \underbrace{|[n+1]|}_{=n+1} \cdot \underbrace{|[n]|}_{=n} = (n+1) \cdot n = n(n+1).$$

On the other hand, the set B consists of all pairs $(i, j) \in [n+1] \times [n]$ such that $i > j$. As we said, the elements of B are the squares in our rectangle that lie below the broken line. In our picture proof above, we counted these squares column by column. Combinatorially, this corresponds to counting the pairs $(i, j) \in B$ according to the value of their first entry i (since the square (i, j) lies in the i-th column). In other words, we count the pairs $(1, j) \in B$, the pairs $(2, j) \in B$, the pairs $(3, j) \in B$, and so on, and then we add all the resulting

[3] the Third proof of Theorem 1.2.1 in those notes, in particular

numbers together. Thus, we obtain

$$\begin{aligned}
|B| &= (\text{\# of pairs } (i,j) \in [n+1] \times [n] \text{ such that } i > j) \\
&= \underbrace{(\text{\# of pairs } (1,j) \in [n+1] \times [n] \text{ such that } 1 > j)}_{\substack{=0 \\ \text{(since there are no such pairs)}}} \\
&\qquad + \underbrace{(\text{\# of pairs } (2,j) \in [n+1] \times [n] \text{ such that } 2 > j)}_{\substack{=1 \\ \text{(since the only such pair is (2,1))}}} \\
&\qquad + \underbrace{(\text{\# of pairs } (3,j) \in [n+1] \times [n] \text{ such that } 3 > j)}_{\substack{=2 \\ \text{(since the only such pairs are (3,1) and (3,2))}}} \\
&\qquad + \cdots \\
&\qquad + \underbrace{(\text{\# of pairs } (n+1,j) \in [n+1] \times [n] \text{ such that } n+1 > j)}_{\substack{=n \\ \text{(since all such pairs are } (n+1,1),\ (n+1,2),\ \ldots,\ (n+1,n))}} \\
&= 0 + 1 + 2 + \cdots + n = 1 + 2 + \cdots + n.
\end{aligned}$$

On the other hand, there is a bijection from A to B, which flips a square across the center of the rectangle. Rigorously speaking, it is the map

$$\begin{aligned}
A &\to B, \\
(i,j) &\mapsto (n+2-i,\ n+1-j).
\end{aligned}$$

By the bijection principle, we thus obtain $|A| = |B|$. Hence, $|A| + |B| = |B| + |B| = 2 \cdot |B|$, so that

$$|B| = \frac{|A| + |B|}{2} = \frac{n(n+1)}{2} \qquad (\text{since } |A| + |B| = n(n+1)).$$

Since $|B| = 1 + 2 + \cdots + n$, this yields

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

Thus, Theorem 1.2.1 is proved again. □

1.2.2. Some words on finite sums

The expression $1 + 2 + \cdots + n$ in Theorem 1.2.1 is an example of a finite sum. Rigorously speaking, it is not a-priori clear that such expressions are well-defined; this relies on the theorem that if you have n numbers $a_1, a_2, \ldots, a_n$, then any way of summing them together will produce the same result. For example,

$$((a_1 + a_2) + a_3) + a_4 = a_1 + ((a_2 + a_3) + a_4) = (a_2 + a_4) + (a_1 + a_3) = \cdots.$$

The 2019 notes (§1.2.2) give references to some proofs of this theorem. We will not focus on such foundational matters here.

We will see lots of finite sums over this course; they are one of its main characters. Thus, we introduce a more systematic notation for such sums:

Definition 1.2.2. Let S be a finite set. For each $s \in S$, let a_s be a number. Then,

$$\sum_{s \in S} a_s$$

shall denote the sum of the numbers a_s for all $s \in S$. This notation is pronounced "the sum of the a_s over all $s \in S$" or "the sum of a_s for s ranging over S" or "the sum of a_s for s running through S". The symbol "$\sum$" is called the **summation sign**.

Example 1.2.3. We have

$$\sum_{s \in \{1,2,3,4\}} s = 1 + 2 + 3 + 4 = 10;$$
$$\sum_{s \in \{1,2,3,4\}} s^2 = 1^2 + 2^2 + 3^2 + 4^2 = 30;$$
$$\sum_{s \in \{-3,0,3\}} s^2 = (-3)^2 + 0^2 + 3^2 = 18.$$

A few conventions and observations about summation signs:

- The letter "s" in the notation "$\sum\limits_{s \in S} a_s$" is called the **summation index**. It can be any letter or symbol. For example, you can rewrite $\sum\limits_{s \in \{1,2,3,4\}} s^2$ as $\sum\limits_{i \in \{1,2,3,4\}} i^2$ or $\sum\limits_{\spadesuit \in \{1,2,3,4\}} \spadesuit^2$.
- The set S in "$\sum\limits_{s \in S} a_s$" is called the **indexing set**.

  The numbers a_s are called the **addends** of the sum.

  The whole expression $\sum\limits_{s \in S} a_s$ is called a **finite sum**.
- If $S = \varnothing$, then the sum $\sum\limits_{s \in S} a_s$ is 0 by definition.[4] Such a sum is called an **empty sum**.

[4]The idea is that if you sum no numbers, then you end up with 0.

- The summation index doesn't have to be a single letter. For example,
$$\sum_{(x,y)\in\{1,2\}^2} \frac{x}{y} = \frac{1}{1} + \frac{1}{2} + \frac{2}{1} + \frac{2}{2}.$$

- Let $\mathcal{P}(A)$ denote the powerset of a set A (that is, the set of all subsets of A). Then, the sum of the sizes of all subsets of $\{1,2,3\}$ can be written as
$$\sum_{B\in\mathcal{P}(\{1,2,3\})} |B|$$
$$= |\varnothing| + |\{1\}| + |\{2\}| + |\{3\}| + |\{1,2\}| + |\{2,3\}| + |\{1,3\}| + |\{1,2,3\}|$$
$$= 0+1+1+1+2+2+2+3 = 12.$$

- The expression "$\sum\limits_{B\in\mathcal{P}(A)}$" is often abbreviated "$\sum\limits_{B\subseteq A}$". So we have just seen that $\sum\limits_{B\subseteq\{1,2,3\}} |B| = 12$.

- A summation sign of the form
$$\sum_{s\in\{p,p+1,\ldots,q\}}$$
(where p and q are integers) is usually written as
$$\sum_{s=p}^{q}.$$
When $p > q$, this is always understood to be an empty sum. For example:
$$\sum_{s=4}^{7} s = 4+5+6+7;$$
$$\sum_{i=4}^{4} i = 4;$$
$$\sum_{j=8}^{2} j = (\text{empty sum}) = 0.$$
Note, in particular, that the summation sign $\sum\limits_{i=1}^{n}$ is equivalent to $\sum\limits_{i\in\{1,2,\ldots,n\}}$ or to $\sum\limits_{i\in[n]}$ (since $[n] = \{1,2,\ldots,n\}$).

  In the notation "$\sum\limits_{s=p}^{q}$", the integers p and q are called the **bounds** of the summation.

  A sum $\sum\limits_{s=p}^{q} a_s$ is also written as $a_p + a_{p+1} + \cdots + a_q$.

- There are lots of similarities between definite integrals $\int_p^q f(x)\,dx$ and finite sums $\sum\limits_{x=p}^{q} f(x)$. However, there are also important dissimilarities. An integral $\int_p^p f(x)\,dx$ over a length-0 interval $[p,p]$ is always 0, whereas the sum $\sum\limits_{x=p}^{p} f(x)$ equals $f(p)$. An integral $\int_p^q f(x)\,dx$ with $p > q$ is usually nonzero (and equals $-\int_q^p f(x)\,dx$), whereas such a sum $\sum\limits_{x=p}^{q} f(x)$ equals 0 by definition.

- The summation sign $\sum\limits_{f:A\to B}$ means a sum over all maps f from A to B. It is just shorthand for "$\sum\limits_{f\in\{\text{maps from } A \text{ to } B\}}$". For instance,

$$\sum_{f:\{1,2\}\to\{5,6\}} f(1) = 5+5+6+6$$

  (because there are four maps f from $\{1,2\}$ to $\{5,6\}$, and two of them satisfy $f(1) = 5$ whereas the other two satisfy $f(1) = 6$).

- We can put conditions under the summation sign to restrict the index we are summing over. For example,

$$\sum_{\substack{s\in S;\\ s \text{ is even}}} a_s$$

  means the sum of all a_s where s ranges over all **even** elements of S. For instance,

$$\sum_{\substack{s\in\{1,2,3,4,5,6,7\};\\ s \text{ is even}}} s = 2+4+6;$$
$$\sum_{\substack{s\in\{1,2,3,4,5,6,7\};\\ s \text{ is odd}}} s = 1+3+5+7.$$

  In general, if $\mathcal{A}(s)$ is some logical statement for each $s \in S$, then "$\sum\limits_{\substack{s\in S;\\ \mathcal{A}(s)}}$" is shorthand for "$\sum\limits_{s\in\{t\in S \mid \mathcal{A}(t)\}}$".

More can be found in the 2019 notes.

1.2.3. Some rules for sums

The summation sign satisfies certain laws. A long list of these rules appears in the 2019 notes (§1.2.3 and §1.6). We shall only list a few of them here. All of these rules are intuitively clear (mostly boiling down to the idea that the addends of a sum can be arbitrarily reordered without changing the value of the sum).

- **Splitting-off:** If $t \in S$, then

  $$\sum_{s \in S} a_s = a_t + \sum_{s \in S \setminus \{t\}} a_s. \tag{1}$$

  In other words, we can split off any addend from any finite sum. Formally, this rule is the (recursive) definition of finite sums. (I am skipping the proof that this definition does really give a well-defined result, i.e., that the right hand side of (1) doesn't depend on the choice of t.)

  In particular, (1) yields

  $$\sum_{i=p}^{q} a_i = a_p + \sum_{i=p+1}^{q} a_i = a_q + \sum_{i=p}^{q-1} a_i \qquad \text{if } p \leq q.$$

  (But this does not hold when $p > q$, since we cannot split off an addend from a sum if the addend is not contained in the sum!)

- **Splitting:** If S is the union of two disjoint sets X and Y, then

  $$\sum_{s \in S} a_s = \sum_{s \in X} a_s + \sum_{s \in Y} a_s \tag{2}$$

  (where the right hand side should be understood as $\left(\sum_{s \in X} a_s \right) + \left(\sum_{s \in Y} a_s \right)$; see the 2019 notes for the precise conventions).

  For example, we obtain

  $$\sum_{s \in \{1,2,3,4,5,6,7\}} a_s = \sum_{\substack{s \in \{1,2,3,4,5,6,7\}; \\ s \text{ is even}}} a_s + \sum_{\substack{s \in \{1,2,3,4,5,6,7\}; \\ s \text{ is odd}}} a_s$$

  by applying (2) to

  $$S = \{1, 2, 3, 4, 5, 6, 7\},$$
  $$X = \{\text{all even elements of } \{1, 2, 3, 4, 5, 6, 7\}\},$$
  $$Y = \{\text{all odd elements of } \{1, 2, 3, 4, 5, 6, 7\}\}.$$

- **Splitting an addend:** We have

  $$\sum_{s \in S} (a_s + b_s) = \sum_{s \in S} a_s + \sum_{s \in S} b_s. \tag{3}$$

- **Substituting the index:** Let S and T be finite sets, and let $f : S \to T$ be a bijection. Then,

$$\sum_{t\in T} a_t = \sum_{s\in S} a_{f(s)}. \tag{4}$$

When we use this rule, we say that we are **substituting** $f(s)$ for t in the sum. The idea behind this rule is that the sum on the left hand side of (4) and the sum on the right hand side of (4) have the same addends, just "in a different order" (although strictly speaking, they both come without a definite order).

A particular case of (4) is the rule saying that we can reverse the order of the addends of a sum – i.e., that we have

$$a_1 + a_2 + \cdots + a_n = a_n + a_{n-1} + \cdots + a_1$$

for any numbers $a_1, a_2, \ldots, a_n$. Indeed, this rule can be rewritten as

$$\sum_{t\in[n]} a_t = \sum_{s\in[n]} a_{n+1-s}.$$

And this is an instance of (2), applied to $S = [n]$, $T = [n]$ and $f(s) = n + 1 - s$.

- **Splitting a sum by a value of a function:** Let S and W be two finite sets. Let $f : S \to W$ be a map. Let a_s be a number for each $s \in S$. Then,

$$\sum_{s\in S} a_s = \sum_{w\in W} \sum_{\substack{s\in S;\\ f(s)=w}} a_s \tag{5}$$

(the right hand side is a sum of sums, which is why you are seeing two $\sum$ signs side by side).

This is a bit abstract, but the underlying idea is completely elementary: It says that if we have a "big sum" $\sum\limits_{s\in S} a_s$, and if each $s \in S$ has a "lucky number" $f(s)$ assigned to it (it doesn't have to be an actual number; we're just calling it this way), then we can compute $\sum\limits_{s\in S} a_s$ by first computing the "small sum" $\sum\limits_{\substack{s\in S;\\ f(s)=w}} a_s$ for each w (this sum contains only those a_s whose corresponding s has "lucky number" w), and then adding up all the results. In other words, we can split our "big sum" into little "bunches", then find the sum of each "bunch", then add up these "bunch sums".[5]

For example, let $S = \{-5, -4, \ldots, 4, 5\}$ and $W = \{0, 1, \ldots, 5\}$ and $f(s) = |s|$. Then, (5) yields

$$\sum_{s\in\{-5,-4,\ldots,4,5\}} a_s = \sum_{w\in\{0,1,\ldots,5\}} \sum_{\substack{s\in\{-5,-4,\ldots,4,5\};\\ |s|=w}} a_s.$$

[5]For those who know MapReduce, this is a simple instance of MapReduce.

Written out, this is saying that

$$a_{-5} + a_{-4} + \cdots + a_4 + a_5 = a_0 + (a_{-1} + a_1) + (a_{-2} + a_2) + \cdots + (a_{-5} + a_5)\,.$$