Math 222 Fall 2022, Lecture 12: Introduction

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

1. Introduction (cont'd)

1.6. Interchange of summations

1.6.2. The Fubini principle with a predicate (cont'd)

Last time we stated a "triangular Fubini principle" for interchanging non-independent summation signs. Here is another such principle:

Corollary 1.6.8. Let $n \in \mathbb{N}$. For each pair $(x, y) \in [n] \times [n]$ satisfying $x \leq y$, let $a_{(x,y)}$ be a number. Then,

$$\sum_{x=1}^{n} \sum_{y=x}^{n} a_{(x,y)} = \sum_{\substack{(x,y) \in [n] \times [n]; \\ x \leq y}} a_{(x,y)} = \sum_{y=1}^{n} \sum_{x=1}^{y} a_{(x,y)}.$$

Visually speaking, this is saying that the entries of a triangular table of the form

$$\begin{array}{ccccc} a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,n-1)} & a_{(1,n)} \\ & a_{(2,2)} & \cdots & a_{(2,n-1)} & a_{(2,n)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{(n-1,n-1)} & a_{(n-1,n)} \\ & & & & a_{(n,n)} \end{array}$$

can be summed in any order (row by row, column by column, or just arbitrarily). This allows us to interchange summation signs if the index of the outer sum appears as a bound (either lower or upper) of the inner sum.

Again, Corollary 1.6.8 follows easily from the finite Fubini principle with a predicate.

Last time, we saw an application of one of our triangular Fubini principles. See §1.6 in the 2019 notes for two other such applications.

1.6.3. Interchange of infinite sums

Time for a cautionary tale.

We have mostly been talking about finite sums so far. However, for convenience, we have sometimes used infinite sums as well. Specifically, we have used **essentially finite sums** – i.e., sums that have only finitely many nonzero addends.[1]

Almost all rules of summation that apply to finite sums can be mindlessly applied to essentially finite sums as well. For instance, the rule

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

holds no matter whether the set S is finite or infinite, as long as the two sums $\sum\limits_{s\in S}a_s$ and $\sum\limits_{s\in S}b_s$ are essentially finite.[2]

There is one exception: interchange of summations. Let us see what happens when we try to interchange two essentially finite sums.

For any $x,y\in\mathbb{N}$, let

$$a_{(x,y)}=[y=x]-[y=x+1].$$

Then, the numbers $a_{(x,y)}$ can be arranged in an infinite table, the first few rows and columns of which look as follows:

	$y=0$	$y=1$	$y=2$	$y=3$	$\cdots$
$x=0$	1	-1			$\cdots$
$x=1$		1	-1		$\cdots$
$x=2$			1	-1	$\cdots$
$x=3$				1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

(the "missing" entries are 0's).

Thus,

$$\begin{aligned}\sum_{x\in\mathbb{N}}\sum_{y\in\mathbb{N}}a_{(x,y)}&=\underbrace{\sum_{y\in\mathbb{N}}a_{(0,y)}}_{\substack{=1+(-1)+0+0+0+\cdots\\=0}}+\underbrace{\sum_{y\in\mathbb{N}}a_{(1,y)}}_{\substack{=0+1+(-1)+0+0+\cdots\\=0}}+\underbrace{\sum_{y\in\mathbb{N}}a_{(2,y)}}_{\substack{=0+0+1+(-1)+0+\cdots\\=0}}+\cdots\\&=0+0+0+\cdots=0,\end{aligned}$$

but

$$\begin{aligned}\sum_{y\in\mathbb{N}}\sum_{x\in\mathbb{N}}a_{(x,y)}&=\underbrace{\sum_{x\in\mathbb{N}}a_{(x,0)}}_{\substack{=1+0+0+0+0+\cdots\\=1}}+\underbrace{\sum_{x\in\mathbb{N}}a_{(x,1)}}_{\substack{=(-1)+1+0+0+0+0+\cdots\\=0}}+\underbrace{\sum_{x\in\mathbb{N}}a_{(x,2)}}_{\substack{=0+(-1)+1+0+0+0+\cdots\\=0}}+\cdots\\&=1+0+0+\cdots=1.\end{aligned}$$

[1]For example, we have used such sums in the proof of the binomial formula.
[2]We have actually used this in our proof of the binomial formula.

All the sums involved here are essentially finite (each having at most two nonzero addends!), but the results are not the same! So we cannot interchange the $\sum\limits_{x\in\mathbb{N}}$ and $\sum\limits_{y\in\mathbb{N}}$ signs in this example.

This is scary. Can we ever interchange two infinite sums?

Yes, we can, but only if a certain extra requirement is satisfied. Namely, the finite Fubini principle can be extended to infinite sums as long as the "**total sum**" $\sum\limits_{(x,y)\in X\times Y} a_{(x,y)}$ is essentially finite (i.e., as long as only finitely many **pairs** $(x,y) \in X \times Y$ satisfy $a_{(x,y)} \neq 0$). In other words, we have the following:

Theorem 1.6.9 (essentially finite Fubini principle)**.** Let X and Y be two (not necessarily finite!) sets. Let $a_{(x,y)}$ be a number for every $(x,y) \in X \times Y$. Assume that only finitely many **pairs** $(x,y) \in X \times Y$ satisfy $a_{(x,y)} \neq 0$. Then,

$$\sum_{x\in X} \sum_{y\in Y} a_{(x,y)} = \sum_{(x,y)\in X\times Y} a_{(x,y)} = \sum_{y\in Y} \sum_{x\in X} a_{(x,y)}.$$

So, if the MHS[3] of this equality is well-defined, then the LHS[4] and the RHS[5] are equal to it. However, if the MHS is not well-defined, then the LHS and RHS need not be equal, even if they themselves happen to be well-defined! This is the situation that we saw in the above scary example.

The upshot is: We can interchange two infinite summation signs $\sum\limits_{x\in X}$ and $\sum\limits_{y\in Y}$ provided that the corresponding $\sum\limits_{(x,y)\in X\times Y}$ sum is essentially finite.

1.7. Counting permutations: an introduction

Now, we return to counting maps. This time, we shall count **bijective** maps from a set to itself; these are also known as *permutations*.

1.7.1. Permutations and derangements

There are two different things called "permutations". One is a kind of mappings ("active permutations"), and the other is a kind of lists ("passive permutations"). We shall here only refer to the former as "permutations". (The latter will later be called "anagrams".) Here is the formal definition:

[3]"MHS" is short for "middle hand side". This is how we refer to the b part in an equality chain of the form $a = b = c$.

[4]This is short for "left hand side".

[5]Guess what this is short for.

Definition 1.7.1. A **permutation** of a set X means a bijection from X to X.

Recall that "bijection" is short for "bijective map" (i.e., a map that is both surjective and injective), which is also known as "one-to-one correspondence". A bijective map is furthermore the same as an invertible map (i.e., a map that has an inverse).

We shall use the two-line notation for maps. (That is, we shall encode a map from a finite set X to a set Y by a two-row table, where the top row has the elements of X listed in it, and the bottom row has their corresponding images under the map.)

Thus, for example,

$$\begin{pmatrix} 0 & 1 & 5 & 7 & 9 \\ 1 & 7 & 9 & 0 & 5 \end{pmatrix}$$

denotes the map from $\{0,1,5,7,9\}$ to $\{0,1,5,7,9\}$ that sends $0,1,5,7,9$ to $1,7,9,0,5$, respectively. This map is a permutation of $\{0,1,5,7,9\}$, since it is bijective. Meanwhile, the map

$$\begin{pmatrix} 0 & 1 & 5 & 7 & 9 \\ 1 & 7 & 7 & 0 & 5 \end{pmatrix}$$

is not a permutation (worse yet, it is neither injective nor surjective, since 9 is not in its image, while 7 appears as an image twice).

The most obvious counting question that we can ask has a simple answer:

Theorem 1.7.2. Let $n \in \mathbb{N}$. Let X be an n-element set. Then,

$$(\#\text{ of permutations of } X) = n!.$$

We shall prove this later, when we have more convenient terminology for the proof. (The proof is easy to explain by waving one's hands, but formalizing it takes some more work.)

Example 1.7.3. The 3-element set $[3]$ has $3! = 6$ permutations, whose two-line notations are

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Let us next focus on a specific kind of permutations:

Definition 1.7.4. Let X be a set.

(a) If $f : X \to X$ is a map, then a **fixed point** of f means an element $x \in X$ such that $f(x) = x$.

(b) A **derangement** of X means a permutation of X that has no fixed points.

Example 1.7.5. **(a)** The map $\begin{pmatrix} 0 & 1 & 5 & 7 & 9 \\ 7 & 1 & 5 & 0 & 9 \end{pmatrix}$ has fixed points 1, 5 and 9. So it is not a derangement.

(b) The map $\begin{pmatrix} 0 & 1 & 5 & 7 & 9 \\ 1 & 7 & 9 & 0 & 5 \end{pmatrix}$ has no fixed points and is a permutation. So it is a derangement.

(c) Among the 6 permutations of $[3]$, only two are derangements, namely

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

1.7.2. Only the size counts

How many derangements does a given n-element set have?

A first simplification is to restrict our focus to a specific n-element set, namely $[n]$, because all n-element sets should be interchangeable as far as their # of derangements is concerned. This relies on the following lemma:

Lemma 1.7.6. Let $n \in \mathbb{N}$. Let X be any n-element set. Then,

$$(\text{\# of derangements of } X) = (\text{\# of derangements of } [n]).$$

Proof. Informally, this is clear: The derangements of an n-element set should not depend on the "private life" of the elements of this set. In other words, relabelling the set should just relabel the derangements, without affecting their number.

Can we turn this into a formal proof? Yes, but surprisingly this takes some work (mostly writing). I will restrict myself to an outline here, but the details can be found in §1.7.2 of the 2019 notes. Essentially, we need to fix a way to "relabel" the n elements of our set X as the numbers $1, 2, \ldots, n$, and then use this "relabelling" to turn the derangements of X into the derangements of $[n]$.

The first of these two steps is easy. To "relabel" the elements of X as the numbers $1, 2, \ldots, n$ means (in rigorous mathematical language) to choose a bijection $\phi : X \to [n]$. Such a bijection ϕ exists because the two sets X and $[n]$ have the same size (since both X and $[n]$ are n-element sets). Generally, there are many such bijections ϕ, and there is no reason to prefer one over another. We just pick one such ϕ and stick with it. Thus, we can think of each element $x \in X$ being "relabelled" as the number $\phi(x) \in [n]$.

Now to the second step: We have to use this bijection ϕ to turn each derangement ω of X into a derangement ω' of $[n]$. This needs to be done in a bijective manner (so that each ω gives rise to exactly one ω' and vice versa). There is an easy way to do this, using our "relabelling" bijection ϕ: Namely, if

$$\omega = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} : X \to X$$

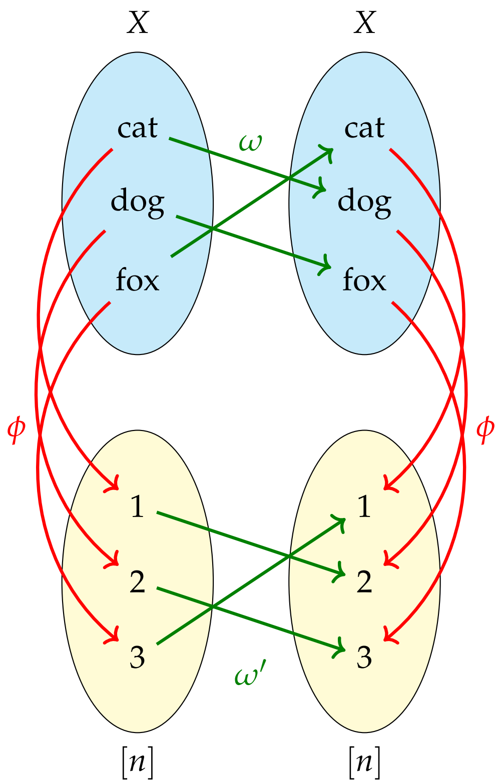
is a derangement of X (written in two-line notation), then we set

$$\omega' := \begin{pmatrix} \phi(x_1) & \phi(x_2) & \cdots & \phi(x_n) \\ \phi(y_1) & \phi(y_2) & \cdots & \phi(y_n) \end{pmatrix} : [n] \to [n]$$

(again, written in two-line notation), which is easily seen to be a derangement of $[n]$. Notice that this ω' can be written more compactly as follows:

$$\omega' = \phi \circ \omega \circ \phi^{-1}.$$

You can convince yourself of this using the following picture (which shows a situation in which $n = 3$ and $X = \{\text{cat, dog, fox}\}$, and where we choose our "relabelling" bijection ϕ to relabel "cat", "dog" and "fox" as 1, 2 and 3, respectively):



Thus, to each derangement ω of X, we have assigned a "corresponding" derangement $\omega' = \phi \circ \omega \circ \phi^{-1}$ of $[n]$. This assignment is bijective – i.e., each ω

corresponds to a unique ω' and vice versa. In other words, we have obtained a bijection

$$\begin{aligned}\{\text{derangements of } X\} &\to \{\text{derangements of } [n]\}, \\ \omega &\mapsto \phi \circ \omega \circ \phi^{-1}.\end{aligned}$$

Indeed, this map is easily seen to be well-defined (you just need to show that $\phi \circ \omega \circ \phi^{-1}$ is a derangement of $[n]$ whenever ω is a derangement of X), and its bijectivity follows from observing that its inverse map is

$$\begin{aligned}\{\text{derangements of } [n]\} &\to \{\text{derangements of } X\}, \\ \omega' &\mapsto \phi^{-1} \circ \omega' \circ \phi.\end{aligned}$$

Thus, by the bijection principle, we get

$$|\{\text{derangements of } X\}| = |\{\text{derangements of } [n]\}|.$$

In other words,

$$(\text{\# of derangements of } X) = (\text{\# of derangements of } [n]).$$

This proves Lemma 1.7.6. □

Lemma 1.7.6 is not really specific to derangements. It is merely an instance of a general principle, which I shall now state in a semi-formal form:

> **"Isomorphism principle" (informal):** Assume that we are given a way to assign a number $f(S)$ to any finite set S, and this way does not depend on what the elements of S are (i.e., its definition works the same no matter whether the elements of S are numbers or tuples or maps or anything else). Let $n \in \mathbb{N}$. Let X be an n-element set. Then,
> $$f(X) = f([n]).$$

Lemma 1.7.6 is the particular case of this principle for

$$f(S) = (\text{\# of derangements of } S).$$

Similarly, we can apply the same principle to

$$\begin{aligned}f(S) &= (\text{\# of subsets of } S), && \text{or} \\ f(S) &= (\text{\# of 5-element subsets of } S), && \text{or} \\ f(S) &= (\text{\# of permutations of } S), && \text{or} \\ f(S) &= (\text{\# of surjective maps } S \to [3]), && \text{or} \\ f(S) &= \sum_{T \subseteq S} |T|.\end{aligned}$$

But we cannot apply it to

$$f(S) = (\text{\# of lacunar subsets of } S),$$

since the lacunarity of a set depends on the elements of this set (and, no surprise, $f(\{1,3\})$ is different from $f([2])$).

I will not formally state the isomorphism principle, but every time I will apply it, I hope it will be clear how it could be proved formally if one wants to.

1.7.3. Intermezzo: OEIS

So, in order to count derangements of any n-element set, it suffices to count derangements of $[n]$. Let us try to do it.

Definition 1.7.7. For each $n \in \mathbb{N}$, let

$$D_n := (\text{\# of derangements of } [n]).$$

Here is a table of the first ten D_n's:

n	0	1	2	3	4	5	6	7	8	9
D_n	1	0	1	2	9	44	265	1854	14833	133496

What now? Is there any simple formula for D_n ?

Here is a very modern method for solving counting problems, and more generally for dealing with sequences of integers: You use the OEIS (short for "Online Encyclopedia of Integer Sequences"), a search engine for integer sequences. It is available at `https://oeis.org` and knows many thousands of sequences along with their properties, formulas and references. When you have a sequence of integers you want to recognize (and learn more about), you compute its first few entries, then plug them into the OEIS, and finally check whether any of the results fits your definition (or, barring that, if it still seems to result in the same sequence). In our case, we quickly figure out that $(D_0, D_1, D_2, \ldots)$ is the sequence A000166, and several formulas are known for it:

$$D_n = n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!};$$

$$D_n = \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor = \operatorname{round} \frac{n!}{e} \qquad \text{for all } n \geq 1 \quad \left(\text{where } e = \sum_{k=0}^{\infty} \frac{1}{k!} \right);$$

$$D_n = nD_{n-1} + (-1)^n \qquad \text{for all } n \geq 1;$$

$$D_n = (n-1)(D_{n-1} + D_{n-2}) \qquad \text{for all } n \geq 2;$$

$$\ldots$$

We will eventually (in Lecture 20, Example 2) prove the first two of these formulas.