Math 4707 Spring 2018 (Darij Grinberg): homework set 3 with solutions

Contents

Recall the following:

- If $n \in \mathbb{N}$, then $[n]$ denotes the n-element set $\{1, 2, \ldots, n\}$.
- We use the Iverson bracket notation.
- If $a \in \mathbb{N}$ and $b \in \mathbb{N}$, then $\operatorname{sur}(a, b)$ denotes the number of surjective maps from $[a]$ to $[b]$.

Also, here is a collection of identities that we shall use:

- We have
$$\binom{m}{n} = 0 \qquad (1)$$
for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfying $m < n$.
- We have
$$\binom{m}{n} = \binom{m-1}{n-1} + \binom{m-1}{n} \qquad (2)$$
for any $m \in \mathbb{Z}$ and $n \in \mathbb{Z}$. (This is the recurrence relation of the binomial coefficients.)
- We have
$$\binom{m}{n} = \binom{m}{m-n} \qquad (3)$$
for any $m \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfying $m \geq n$.

- We have
$$\binom{m+n}{m} = \binom{m+n}{n} \tag{4}$$
for any $m \in \mathbb{N}$ and $n \in \mathbb{N}$. (This follows by applying (3) to $m+n$ and m instead of m and n.)

- For every $x \in \mathbb{N}$ and $y \in \mathbb{N}$ and $n \in \mathbb{N}$, we have
$$\binom{n+1}{x+y+1} = \sum_{k=0}^{n} \binom{k}{x} \binom{n-k}{y}. \tag{5}$$
(This is Proposition 3.21 in the classwork from 26 February 2018, or [Grinbe16, Proposition 3.32 **(f)**].)

- We have
$$\binom{m}{n} = \frac{m}{n} \binom{m-1}{n-1} \tag{6}$$
for any $m \in \mathbb{Q}$ and $n \in \{1, 2, 3, \ldots\}$. (This is the *absorption identity*, and has been proven in [Grinbe16, Proposition 3.22]. Also, it is very easy to check.)

- Every $n \in \mathbb{N}$ satisfies
$$\sum_{k=0}^{n} \binom{n}{k} = 2^n. \tag{7}$$
(This is Corollary 1.16b in the classwork from 22 January 2018, or [Grinbe16, Proposition 3.39 **(b)**].)

- Every $n \in \mathbb{N}$ satisfies
$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = [n = 0]. \tag{8}$$
(This is Corollary 3.3 in the classwork from 14 February 2018, or [Grinbe16, Proposition 3.39 **(c)**].)

- If $m \in \mathbb{N}$ and $n \in \mathbb{N}$, and if S is an m-element set, then
$$\binom{m}{n} \text{ is the number of all } n\text{-element subsets of } S. \tag{9}$$
(This is the combinatorial interpretation of the binomial coefficients.)

0.1. Another binomial identity

Exercise 1. Let $n \in \mathbb{N}$. Prove that

$$\sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \binom{n}{k} = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}. \qquad (10)$$

[**Hint:** How does the left hand side grow when n is replaced by $n+1$?]

We shall outline a solution of this exercise. (For the missing details, see [Grinbe16, solution to Exercise 3.19].)

Solution to Exercise 1 (sketched). We shall solve Exercise 1 by induction on n:

Induction base: If $n = 0$, then both sides of (10) are empty sums, and thus are equal. Hence, Exercise 1 holds for $n = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that Exercise 1 holds for $n = m$. We must prove that Exercise 1 holds for $n = m+1$.

We have assumed that Exercise 1 holds for $n = m$. In other words, we have

$$\sum_{k=1}^{m} \frac{(-1)^{k-1}}{k} \binom{m}{k} = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{m}. \qquad (11)$$

Now,

$$\begin{aligned}
&\sum_{k=1}^{m+1} \frac{(-1)^{k-1}}{k} \underbrace{\binom{m+1}{k}}_{\substack{=\binom{m}{k-1}+\binom{m}{k} \\ \text{(by the recurrence relation} \\ \text{of the binomial coefficients)}}} \\
&= \sum_{k=1}^{m+1} \underbrace{\frac{(-1)^{k-1}}{k}\binom{m}{k-1}}_{=\frac{(-1)^{k-1}}{m+1}\cdot\frac{m+1}{k}\binom{m}{k-1}} + \underbrace{\sum_{k=1}^{m+1} \frac{(-1)^{k-1}}{k}\binom{m}{k}}_{=\sum_{k=1}^{m}\frac{(-1)^{k-1}}{k}\binom{m}{k}+\frac{(-1)^{(m+1)-1}}{m+1}\binom{m}{m+1}} \\
&= \underbrace{\sum_{k=1}^{m+1} \frac{(-1)^{k-1}}{m+1}\cdot\frac{m+1}{k}\binom{m}{k-1}}_{=\frac{1}{m+1}\sum_{k=1}^{m+1}(-1)^{k-1}\frac{m+1}{k}\binom{m}{k-1}} + \underbrace{\sum_{k=1}^{m} \frac{(-1)^{k-1}}{k}\binom{m}{k}}_{\substack{=\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{m} \\ \text{(by (11))}}} + \frac{(-1)^{(m+1)-1}}{m+1} \underbrace{\binom{m}{m+1}}_{\substack{=0 \\ \text{(by (1))}}} \\
&= \frac{1}{m+1}\sum_{k=1}^{m+1} (-1)^{k-1}\frac{m+1}{k}\binom{m}{k-1} + \left(\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{m}\right).
\end{aligned}$$

In view of

$$\begin{aligned}
&\sum_{k=1}^{m+1} (-1)^{k-1} \underbrace{\frac{m+1}{k}\binom{m}{k-1}}_{\substack{=\binom{m+1}{k} \\ \text{(because (6) (applied to } m+1 \text{ and } k \\ \text{instead of } m \text{ and } n\text{) yields} \\ \binom{m+1}{k} = \frac{m+1}{k}\binom{m}{k-1})}} \\
&= \sum_{k=1}^{m+1} (-1)^{k-1} \binom{m+1}{k} = \sum_{k=0}^{m+1} \underbrace{(-1)^{k-1}}_{=-(-1)^k} \binom{m+1}{k} - \underbrace{(-1)^{0-1}}_{=-1} \underbrace{\binom{m+1}{0}}_{=1} \\
&= -\underbrace{\sum_{k=0}^{m+1} (-1)^k \binom{m+1}{k}}_{\substack{=[m+1=0] \\ \text{(by (8))}}} - (-1) = -\underbrace{[m+1=0]}_{\substack{=0 \\ \text{(since } m+1\neq 0)}} - (-1) = 1,
\end{aligned}$$

this becomes

$$\begin{aligned}
&\sum_{k=1}^{m+1} \frac{(-1)^{k-1}}{k} \binom{m+1}{k} \\
&= \frac{1}{m+1} \underbrace{\sum_{k=1}^{m+1} (-1)^{k-1} \frac{m+1}{k} \binom{m}{k-1}}_{=1} + \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{m}\right) \\
&= \frac{1}{m+1} + \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{m}\right) = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{m+1}.
\end{aligned}$$

In other words, Exercise 1 holds for $n = m+1$. This completes the induction step. Thus, Exercise 1 is solved. □

0.2. More on inclusion/exclusion

0.2.1. An exercise on counting surjections and more

Exercise 2. Let A, B and C be three finite sets such that $C \subseteq B$. Let $a = |A|$, $b = |B|$ and $c = |C|$.

(a) Prove that the number of maps $f : A \to B$ satisfying $C \subseteq f(A)$ is

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

(b) Prove that the number of surjective maps $f : A \to B$ is

$$\operatorname{sur}(a,b) = \sum_{k=0}^{b} (-1)^k \binom{b}{k} (b-k)^a = \sum_{k=0}^{b} (-1)^{b-k} \binom{b}{k} k^a.$$

[This is a formula I mentioned but did not prove in class; of course, you cannot use it without proof.]

(c) Prove that

$$\sum_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a = 0$$

whenever $c > a$.

(d) Prove that

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

[**Hint:** For part **(a)**, notice that a map $f : A \to B$ satisfies $C \subseteq f(A)$ if and only if the image of f misses none of the c elements of C. Parts **(b)**, **(c)** and **(d)** should follow from **(a)**.]

This exercise has turned out to be harder than I thought when I posed it. In particular, deriving its part **(d)** from the previous parts is tricky, since you cannot just set $c = a$ and apply part **(a)** (indeed, in order to set $c = a$, you would need to find a b-element set B with an a-element subset C, but this is only possible when $a \leq b$). We shall below outline a solution that addresses these issues by using the "polynomial identity trick" (Lemma 0.3 below). It is also possible to prove parts **(c)** and **(d)** of Exercise 2 algebraically, completely avoiding the use of the previous parts; this is done in [Grinbe09, proof of Corollary 2].

0.2.2. Solutions to Exercise 2 (a), (b) and (c)

Let us first recall the principle of inclusion and exclusion:

Theorem 0.1. Let $n \in \mathbb{N}$. Let $A_1, A_2, \ldots, A_n$ be finite sets.

(a) We have

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\substack{I \subseteq [n]; \\ I \neq \varnothing}} (-1)^{|I|-1} \left| \bigcap_{i \in I} A_i \right|.$$

(b) Let S be a finite set. Assume that each of $A_1, A_2, \ldots, A_n$ is a subset of S. Then,

$$\left| S \setminus \bigcup_{i=1}^{n} A_i \right| = \sum_{I \subseteq [n]} (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right|.$$

Here, the "empty" intersection $\bigcap_{i \in \varnothing} A_i$ is understood to mean the set S.

Solution to Exercise 2 (sketched). Let us first solve parts **(a)**, **(b)** and **(c)** of the exercise. Then, we'll prepare a bit further for the solution to **(d)**.

We shall use the symbol $\cong$ for "in bijection with"; i.e., for two sets U and V, we write "$U \cong V$" if and only if there is a bijection from U to V. (In abstract algebra, the symbol $\cong$ stands for "is isomorphic to", which is a generalization of our use of this symbol.)

(a) We WLOG assume that $C = [c]$ (since we can relabel the c elements of C as we wish, in particular as $1, 2, \ldots, c$). Thus, $[c] = C \subseteq B$.

Let S be the set of all maps $f : A \to B$. For each $i \in C$, we let T_i be the set of all maps $f : A \to B$ that satisfy $i \notin f(A)$ (in other words, that never take the value i). Thus, we have defined c sets $T_1, T_2, \ldots, T_c$ (since $C = [c]$). These c sets are subsets of S. Their union is

$$\begin{aligned} \bigcup_{i=1}^{c} T_i &= \{ f : A \to B \mid \text{there is some } i \in [c] \text{ such that } i \notin f(A) \} \\ &= \left\{ f : A \to B \mid \underbrace{[c]}_{=C} \not\subseteq f(A) \right\} \\ &= \{ f : A \to B \mid C \not\subseteq f(A) \} . \end{aligned} \tag{12}$$

On the other hand, each subset I of $[c]$ satisfies

$$\begin{aligned} \bigcap_{i \in I} T_i &= \{ f : A \to B \mid f \text{ takes none of the } i \in I \text{ as values} \} \\ &= \{ f : A \to B \mid \text{all values of } f \text{ belong to } B \setminus I \} \\ &\cong \{ f : A \to B \setminus I \} \end{aligned}$$

(because maps $f : A \to B$ whose all values belong to $B \setminus I$ are in bijection with maps $f : A \to B \setminus I$ [1]) and thus

$$\begin{aligned} \left| \bigcap_{i \in I} T_i \right| &= |\{ f : A \to B \setminus I \}| = |B \setminus I|^{|A|} \\ &= (|B| - |I|)^{|A|} \qquad (\text{since } |B \setminus I| = |B| - |I| \text{ (because } I \subseteq [c] \subseteq B)) \\ &= (b - |I|)^{a} \end{aligned} \tag{13}$$

(since $|B| = b$ and $|A| = a$). Note that this holds even when $I = \varnothing$, as long as the "empty" intersection $\bigcap_{i \in \varnothing} T_i$ is understood to mean the set S.

[1] To transform a former map into a latter map, just switch its target from B to $B \setminus I$, without changing its values. Similarly for the backwards transformation.

Theorem 0.1 **(b)** (applied to c and T_i instead of n and A_i) yields

$$\begin{aligned}
&\left| S \setminus \bigcup_{i=1}^{c} T_i \right| \\
&= \underbrace{\sum_{I \subseteq [c]}}_{= \sum_{k=0}^{c} \sum_{I \subseteq [c]; \ |I|=k}} (-1)^{|I|} \underbrace{\left| \bigcap_{i \in I} T_i \right|}_{\substack{=(b-|I|)^a \\ \text{(by (13))}}} = \sum_{k=0}^{c} \sum_{I \subseteq [c]; \ |I|=k} \underbrace{(-1)^{|I|} (b-|I|)^a}_{\substack{=(-1)^k (b-k)^a \\ \text{(since } |I|=k)}} \\
&= \sum_{k=0}^{c} \underbrace{\sum_{I \subseteq [c]; \ |I|=k} (-1)^k (b-k)^a}_{(\text{the number of all subsets } I \subseteq [c] \text{ satisfying } |I|=k) \cdot (-1)^k (b-k)^a} \\
&= \sum_{k=0}^{c} \underbrace{(\text{the number of all subsets } I \subseteq [c] \text{ satisfying } |I|=k)}_{\substack{=(\text{the number of all } k\text{-element subsets of } [c]) \\ = \binom{c}{k} \\ \text{(by (9))}}} \cdot (-1)^k (b-k)^a \\
&= \sum_{k=0}^{c} \binom{c}{k} \cdot (-1)^k (b-k)^a = \sum_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a .
\end{aligned}$$

In view of

$$\begin{aligned}
S \setminus \bigcup_{i=1}^{c} T_i &= S \setminus \{ f : A \to B \ | \ C \not\subseteq f(A) \} \qquad \text{(by (12))} \\
&= \{ f : A \to B \ | \ \text{not } C \not\subseteq f(A) \} = \{ f : A \to B \ | \ C \subseteq f(A) \} ,
\end{aligned}$$

this rewrites as

$$|\{ f : A \to B \ | \ C \subseteq f(A) \}| = \sum_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a .$$

This solves Exercise 2 **(a)**.

(b) The surjective maps $f : A \to B$ are precisely the maps $f : A \to B$ satisfying $B \subseteq f(A)$. Hence, Exercise 2 **(a)** (applied to B and b instead of C and c) yields that the number of surjective maps $f : A \to B$ is

$$\begin{aligned} &\sum_{k=0}^{b} (-1)^k \binom{b}{k} (b-k)^a \\ &= \sum_{k=0}^{b} (-1)^{b-k} \underbrace{\binom{b}{b-k}}_{\substack{=\binom{b}{k} \\ \text{(by (3))}}} \underbrace{(b-(b-k))^a}_{=k^a} \\ &\qquad \left(\text{here, we have substituted } b-k \text{ for } k \text{ in the sum}\right) \\ &= \sum_{k=0}^{b} (-1)^{b-k} \binom{b}{k} k^a. \end{aligned}$$

Also, of course, we know that this number is $\operatorname{sur}(a,b)$ (since $|A| = a$ and $|B| = b$). Thus, Exercise 2 **(b)** is solved.

(c) Assume that $c > a$. Thus, there exist no maps $f : A \to B$ satisfying $C \subseteq f(A)$ (because $C \subseteq f(A)$ would yield $|C| \leq |f(A)| \leq |A| = a < c = |C|$, which is absurd).

Exercise 2 **(a)** yields that the number of maps $f : A \to B$ satisfying $C \subseteq f(A)$ is

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

On the other hand, the same number must be 0 (since there exist no maps $f : A \to B$ satisfying $C \subseteq f(A)$). Comparing the two results, we obtain $\sum_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a = 0$. This solves Exercise 2 **(c)**.

Part **(d)** follows from Theorem 0.4 below. □

Let us now take a break and discuss a generalization of part **(c)**. In fact, our above solution of Exercise 2 **(c)** presumes that A, B, C, a, b and c are as defined in Exercise 2 (thus, A, B and C are three finite sets such that $C \subseteq B$, and $a = |A|$, $b = |B|$ and $c = |C|$). This is fine – since the exercise itself makes these assumptions. But you might wonder what happens more generally, if a, b and c are **arbitrary** nonnegative integers satisfying $c > a$, not necessarily the sizes of three finite sets A, B and C satisfying $C \subseteq B$. Does $\sum_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a = 0$ still hold in this case?

The answer is "yes". We can even generalize a bit further and state this for all $b \in \mathbb{Q}$:

Theorem 0.2. Let $a \in \mathbb{N}$, $b \in \mathbb{Q}$ and $c \in \mathbb{N}$ be such that $c > a$. Then,

$$\sum_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a = 0. \tag{14}$$

In order to prove this theorem in full generality, we will need the "polynomial identity trick"; specifically, we shall use the following fact (Theorem 3.20 **(b)** in the classwork from 26 February 2018):

Lemma 0.3. If a polynomial P with rational coefficients has infinitely many roots, then P is the zero polynomial.

Our proof of Theorem 0.2 will be similar to how we salvaged our first two proofs of the Chu-Vandermonde convolution identity back in class:

Proof of Theorem 0.2. Let us forget that b is fixed. Thus, a and c are fixed, but b is not. We must prove that (14) holds for each $b \in \mathbb{Q}$.

We now claim that for every **integer** $b \geq c$, the equality (14) holds.

[*Proof:* Let $b \geq a$ be an integer. Let $A = [a]$, $B = [b]$ and $C = [c]$. Then, $C \subseteq B$ (since $c \leq b$). Also, clearly, $a = |A|$ and $b = |B|$ and $c = |C|$. Thus, Exercise 2 **(c)** yields $\sum\limits_{k=0}^{c} (-1)^k \binom{c}{k} (b-k)^a = 0$. In other words, the equality (14) holds. Qed.]

So we have shown that for every **integer** $b \geq c$, the equality (14) holds. In other words, every integer $b \geq c$ is a root of the polynomial $\sum\limits_{k=0}^{c} (-1)^k \binom{c}{k} (x-k)^a$ (in the indeterminate x). Thus, this polynomial has infinitely many roots (because there are infinitely many integers $b \geq c$). Thus, Lemma 0.3 shows that this polynomial is the zero polynomial. In other words,

$$\sum_{k=0}^{c} (-1)^k \binom{c}{k} (x-k)^a = 0. \tag{15}$$

Now, for any $b \in \mathbb{Q}$, we can substitute b for x in the identity (15), and we obtain precisely (14). This proves Theorem 0.2. $\square$

0.2.3. Solution to Exercise 2 (d)

Let us next prove a similarly generalized part **(d)** of Exercise 2:

Theorem 0.4. Let $a \in \mathbb{N}$ and $b \in \mathbb{Q}$. Then,

$$\sum_{k=0}^{a} (-1)^k \binom{a}{k} (b-k)^a = a!. \tag{16}$$

Our proof will rely on the following simple fact:

Proposition 0.5. Let A, B and C be three finite sets such that $C \subseteq B$ and $|C| = |A|$. Then:

(a) The number of injective maps $f : A \to B$ whose image is C is $|A|!$.

(b) The number of maps $f : A \to B$ satisfying $C \subseteq f(A)$ is $|A|!$.

(c) The number of injective maps $f : A \to B$ satisfying $f(A) \subseteq C$ is $|A|!$.

Proof of Proposition 0.5 (sketched). **(a)** The injective maps $f : A \to B$ whose image is C can be identified with the bijective maps from A to C. More precisely: If $f : A \to B$ is any map satisfying $f(A) \subseteq C$, then we can define a map $\overline{f} : A \to C$ by

$$\left(\overline{f}(a) = f(a) \qquad \text{for all } a \in A\right).$$

This latter map $\overline{f}$ is well-defined, because $f(a) \in f(A) \subseteq C$ for each $a \in A$. Moreover, f can be reconstructed from $\overline{f}$.

If $f : A \to B$ is an **injective** map satisfying $f(A) \subseteq C$, then the resulting map $\overline{f} : A \to C$ is again injective. Moreover, if $f : A \to B$ is a map whose image is C (that is, we have $f(A) = C$, not just $f(A) \subseteq C$), then the resulting map $\overline{f} : A \to C$ is surjective. Combining the preceding two statements, we conclude that if $f : A \to B$ is an injective map whose image is C, then the resulting map $\overline{f} : A \to C$ is bijective. Thus, we obtain a mapping

$$\begin{aligned} \{\text{injective maps } f : A \to B \text{ whose image is } C\} &\to \{\text{bijective maps from } A \text{ to } C\}, \\ f &\mapsto \overline{f}. \end{aligned}$$

This mapping is easily seen to be itself bijective (indeed, for each bijective map g from A to C, there is a **unique** map $f : A \to B$ whose image is C such that $\overline{f} = g$: namely, this f is just g with the target switched to B). Hence, we have found a bijection between the sets $\{\text{injective maps } f : A \to B \text{ whose image is } C\}$ and $\{\text{bijective maps from } A \text{ to } C\}$. Thus,

$$\begin{aligned} &|\{\text{injective maps } f : A \to B \text{ whose image is } C\}| \\ &= |\{\text{bijective maps from } A \text{ to } C\}|. \end{aligned} \tag{17}$$

On the other hand, there exists a bijection $\phi : C \to A$ (since $|C| = |A|$). Fix such a ϕ. Then, there is a bijection

$$\begin{aligned} \{\text{bijective maps from } A \text{ to } C\} &\to \{\text{bijective maps from } A \text{ to } A\}, \\ g &\mapsto \phi \circ g \end{aligned}$$

(its inverse is given by $g \mapsto \phi^{-1} \circ g$). Hence,

$$\begin{aligned} |\{\text{bijective maps from } A \text{ to } C\}| &= |\{\text{bijective maps from } A \text{ to } A\}| \\ &= |\{\text{permutations of } A\}| \\ &= (\text{the number of permutations of } A) \\ &= |A|!. \end{aligned}$$

Thus, (17) becomes

$$\begin{aligned} &|\{\text{injective maps } f : A \to B \text{ whose image is } C\}| \\ &= |\{\text{bijective maps from } A \text{ to } C\}| = |A|!. \end{aligned}$$

In other words, the number of injective maps $f : A \to B$ whose image is C is $|A|!$. This proves Proposition 0.5 **(a)**.

(b) The maps $f : A \to B$ satisfying $C \subseteq f(A)$ are the same as the maps $f : A \to B$ satisfying $C = f(A)$ (because the inequality $|C| = |A| \geq |f(A)|$ shows that C cannot be a **proper** subset of $f(A)$). Furthermore, all these maps are necessarily injective (because if such a map $f : A \to B$ was not injective, then it would satisfy $|f(A)| < |A| = |C|$, which would contradict the condition $C = f(A)$). Hence, the maps $f : A \to B$ satisfying $C \subseteq f(A)$ are precisely the injective maps $f : A \to B$ whose image is C. Thus, Proposition 0.5 **(b)** follows immediately from Proposition 0.5 **(a)**.

(c) If an injective map $f : A \to B$ satisfies $f(A) \subseteq C$, then it must also satisfy $f(A) = C$ (because the injectivity of f yields $|f(A)| = |A| = |C|$, so that $f(A)$ cannot be a **proper** subset of C). In other words, if an injective map $f : A \to B$ satisfies $f(A) \subseteq C$, then its image must be C. Thus, the injective maps $f : A \to B$ satisfying $f(A) \subseteq C$ are the same as the injective maps $f : A \to B$ whose image is C. Hence, Proposition 0.5 **(c)** follows immediately from Proposition 0.5 **(a)**. □

Proof of Theorem 0.4 (sketched). Let us forget that b is fixed. Thus, a is fixed, but b is not. We must prove that (16) holds for each $b \in \mathbb{Q}$.

We now claim that for every **integer** $b \geq a$, the equality (16) holds.

[*Proof:* Let $b \geq a$ be an integer. Let $A = [a]$, $B = [b]$ and $C = [a]$. Then, $C \subseteq B$ (since $a \leq b$). Also, clearly, $a = |A|$ and $b = |B|$ and $a = |C|$. Hence, $|C| = a = |A|$.

Thus, Proposition 0.5 **(b)** yields that the number of maps $f : A \to B$ satisfying $C \subseteq f(A)$ is $|A|! = a!$ (since $|A| = a$).

On the other hand, Exercise 2 **(a)** (applied to $c = a$) yields that this number is

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

Comparing these two results, we conclude that $\sum_{k=0}^{a} (-1)^k \binom{a}{k} (b-k)^a = a!$. In other words, the equality (16) holds. Qed.]

So we have shown that for every **integer** $b \geq a$, the equality (16) holds. In other words, every integer $b \geq a$ is a root of the polynomial $\sum_{k=0}^{a} (-1)^k \binom{a}{k} (x-k)^a - a!$ (in the indeterminate x). Thus, this polynomial has infinitely many roots (because there are infinitely many integers $b \geq a$). Thus, Lemma 0.3 shows that this polynomial is the zero polynomial. In other words,

$$\sum_{k=0}^{a} (-1)^k \binom{a}{k} (x-k)^a - a! = 0. \qquad (18)$$

Now, for any $b \in \mathbb{Q}$, we can substitute b for x in the identity (18), and we obtain $\sum_{k=0}^{a} (-1)^k \binom{a}{k} (b-k)^a - a! = 0$. This is clearly equivalent to (16). This proves Theorem 0.4. □

With Theorem 0.4 proven, Exercise 2 **(d)** immediately follows.

0.3. Multijections, set compositions and set partitions

0.3.1. Definitions and examples

Let us introduce set partitions and set compositions. These are important concepts in combinatorics, and we will see more of them.

Definition 0.6. Let X be a finite set.

(a) A *set composition* of X means a tuple $(S_1, S_2, \ldots, S_k)$ of disjoint nonempty subsets of X such that $X = S_1 \cup S_2 \cup \cdots \cup S_k$.

(b) A *set partition* of X means a set $\{S_1, S_2, \ldots, S_k\}$ of disjoint nonempty subsets of X (written in such a way that $S_1, S_2, \ldots, S_k$ are distinct) such that $X = S_1 \cup S_2 \cup \cdots \cup S_k$.

(c) The *parts* of a set composition $(S_1, S_2, \ldots, S_k)$ are the sets $S_1, S_2, \ldots, S_k$.

(d) The *parts* of a set partition $\{S_1, S_2, \ldots, S_k\}$ are the sets $S_1, S_2, \ldots, S_k$.

Example 0.7. For this example, let $X = \{1, 2, 3\}$.

(a) The 2-tuple $(\{1,3\}, \{2\})$ is a set composition of X, since $\{1,3\}$ and $\{2\}$ are disjoint nonempty subsets of X satisfying $X = \{1,3\} \cup \{2\}$.

(b) The 2-tuple $(\{1,3\}, \{2,3\})$ is **not** a set composition of X, since $\{1,3\}$ and $\{2,3\}$ are not disjoint.

(c) The 3-tuple $(\{1,3\}, \{\}, \{2\})$ is **not** a set composition of X, since $\{\}$ is not nonempty.

(d) The 2-tuple $(\{1\}, \{3\})$ is **not** a set composition of X, since $X \neq \{1\} \cup \{3\}$.

(e) The set $\{\{1,3\}, \{2\}\}$ is a set partition of X, since $\{1,3\}$ and $\{2\}$ are disjoint nonempty subsets of X satisfying $X = \{1,3\} \cup \{2\}$.

(f) If $(S_1, S_2, \ldots, S_k)$ is a set composition of X, then $\{S_1, S_2, \ldots, S_k\}$ is a set partition of X. The converse also holds if you assume $S_1, S_2, \ldots, S_k$ to be distinct.

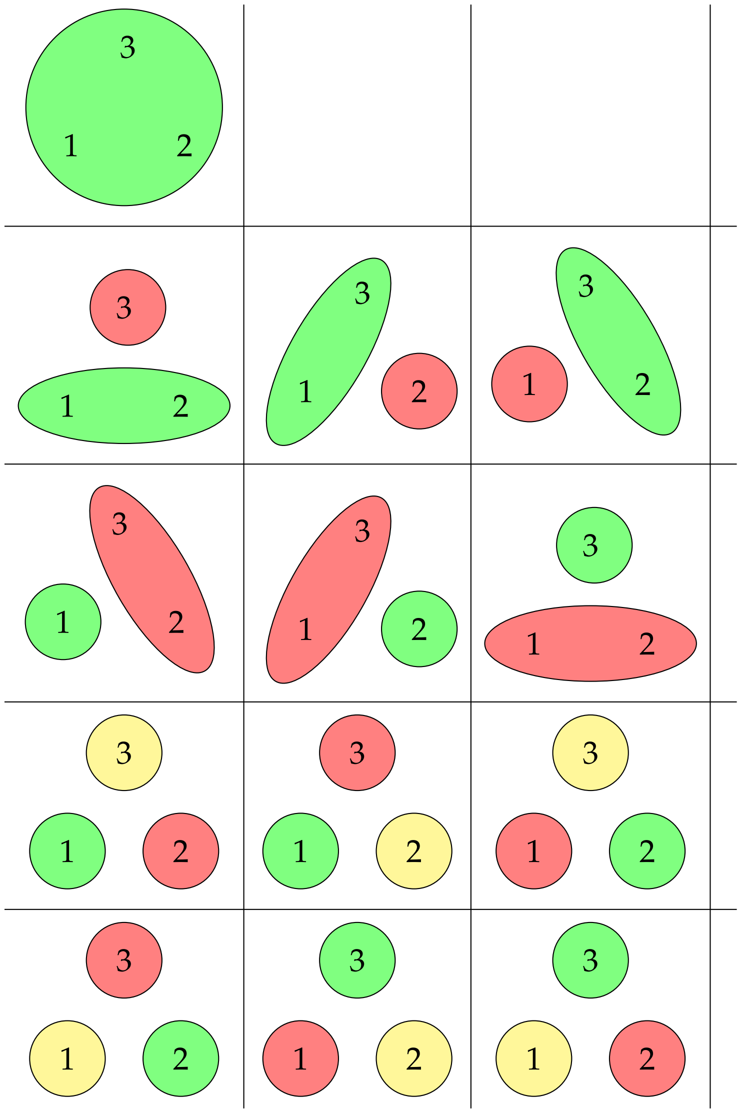
(g) The set **compositions** $(\{1,3\}, \{2\})$ and $(\{2\}, \{1,3\})$ of X are distinct, but the set **partitions** $\{\{1,3\}, \{2\}\}$ and $\{\{2\}, \{1,3\}\}$ are identical.

This illustrates the difference between set compositions and set partitions: The former come with an ordering of their parts, while the latter don't. This is why set compositions are often called *ordered set partitions*.

Example 0.8. (a) Here are all set compositions of the set $X = \{1, 2, 3\}$:

$$\begin{array}{lll}
(\{1,2,3\}), & & \\
(\{1,2\}, \{3\}), & (\{1,3\}, \{2\}), & (\{2,3\}, \{1\}), \\
(\{1\}, \{2,3\}), & (\{2\}, \{1,3\}), & (\{3\}, \{1,2\}), \\
(\{1\}, \{2\}, \{3\}), & (\{1\}, \{3\}, \{2\}), & (\{2\}, \{1\}, \{3\}), \\
(\{2\}, \{3\}, \{1\}), & (\{3\}, \{1\}, \{2\}), & (\{3\}, \{2\}, \{1\}).
\end{array}$$

And here are the same set compositions, drawn as pictures (each part of the set composition corresponds to a colored blob):

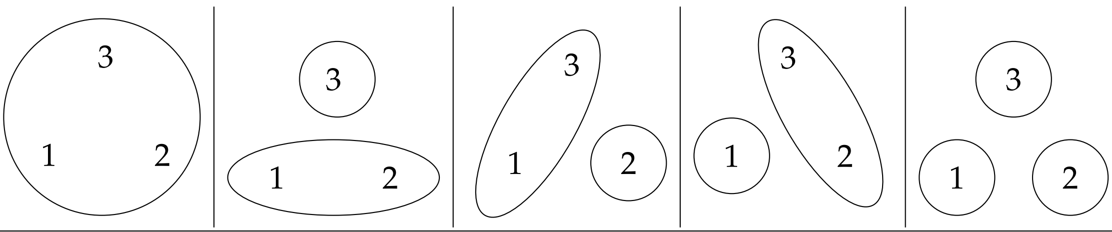


Here, the colors have been chosen as follows: The green-colored blob is the first part of the set composition; the red-colored blob is the second; the yellow-colored blob is the third.

(b) Here are all set partitions of the set $X = \{1, 2, 3\}$:

$$\{\{1,2,3\}\}, \qquad \{\{1,2\},\{3\}\}, \qquad \{\{1,3\},\{2\}\}, \qquad \{\{2,3\},\{1\}\}, \qquad \{\{1\},\{2\},\{3\}\}.$$

And here are the same set partitions, drawn as pictures (each part of the set partition corresponds to a blob):



Definition 0.9. Let X be a set. Let $k \in \mathbb{N}$.

(a) A *set composition of X into k parts* is a set composition of X having exactly k parts.

(b) A *set partition of X into k parts* is a set partition of X having exactly k parts.

For example, $(\{1,5\}, \{2\}, \{3,4,6,7\})$ is a set composition of $[7]$ into 3 parts.

0.3.2. Counting set partitions and set compositions

There are several natural things to be counted now:

- set compositions of a given set X;
- set partitions of a given set X;
- set compositions of a given set X into k parts for a given $k \in \mathbb{N}$;
- set partitions of a given set X into k parts for a given $k \in \mathbb{N}$.

Let's only briefly comment on the first two questions, and then discuss the last two.

- If X is an n-element set, then the number of set compositions of X is the n-th *ordered Bell number* $\widetilde{B}(n)$. Here is a list of the first values of $\widetilde{B}(n)$ (see A000670 at OEIS for more):

$$\widetilde{B}(0) = 1, \quad \widetilde{B}(1) = 1, \quad \widetilde{B}(2) = 3, \quad \widetilde{B}(3) = 13,$$
$$\widetilde{B}(4) = 75, \quad \widetilde{B}(5) = 541, \quad \widetilde{B}(6) = 4683, \quad \widetilde{B}(7) = 47293.$$

  No explicit formulas for $\widetilde{B}(n)$ are known. A quick way to compute $\widetilde{B}(n)$ for arbitrary $n \in \mathbb{N}$ is using the recursive equation

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

  (The proof is easy: Classify all set compositions of $[n]$ according to the size of their last part, and treat all the remaining parts as a set composition of a smaller set.)

- If X is an n-element set, then the number of set partitions of X is the n-th *Bell number* $B(n)$. Here is a list of the first values of $B(n)$ (see A000110 at OEIS for more):

$$B(0) = 1, \quad B(1) = 1, \quad B(2) = 2, \quad B(3) = 5,$$
$$B(4) = 15, \quad B(5) = 52, \quad B(6) = 203, \quad B(7) = 877.$$

  No explicit formulas for $B(n)$ are known. A quick way to compute $B(n)$ for arbitrary $n \in \mathbb{N}$ is using the recursive equation

$$B(n+1) = \sum_{i=0}^{n} \binom{n}{i} B(i) \qquad \text{for all } n \in \mathbb{N}.$$

  (The proof is easy: Classify all set partitions of $[n+1]$ according to how many elements lie in the same part as $n+1$, and treat all the remaining parts as a set partition of a smaller set.)

Now, let us study the other two questions.

Definition 0.10. Let X be a set. Let $k \in \mathbb{N}$.
(a) We let $\operatorname{SC}_k(X)$ denote the set of all set compositions of X into k parts.
(b) We let $\operatorname{SP}_k(X)$ denote the set of all set partitions of X into k parts.

Proposition 0.11. Let X be a finite set. Let $k \in \mathbb{N}$. Then, $|\operatorname{SC}_k(X)| = \operatorname{sur}(|X|, k)$.

Proof of Proposition 0.11 (sketched). For any sets A and B, we let $\operatorname{Sur}(A, B)$ be the set of all surjections from A to B. If A and B are finite sets, then

$$|\operatorname{Sur}(A, B)| = (\text{the number of surjections from } A \text{ to } B) = \operatorname{sur}(|A|, |B|)$$

(as we have shown in class). Applying this to $A = X$ and $B = [k]$, we obtain

$$|\operatorname{Sur}(X, [k])| = \operatorname{sur}\left(|X|, \underbrace{|[k]|}_{=k}\right) = \operatorname{sur}(|X|, k).$$

Here is an outline of the remainder of the proof: We want to find $|\operatorname{SC}_k(X)|$; that is, we want to count all set compositions $(S_1, S_2, \ldots, S_k)$ of X into k parts. We can construct such a set composition by choosing, for each $x \in X$, which part S_i it shall belong to[2]. This information can be encoded as a map $f : X \to [k]$ (which sends each $x \in X$ to the $i \in [k]$ satisfying $x \in S_i$); this map f has to be surjective (since the parts S_i should be nonempty, so each S_i must have at least one $x \in X$ in it), but otherwise is subject to no constraints. Hence, the number of set compositions $(S_1, S_2, \ldots, S_k)$ of X into k parts equals the number of surjective maps $X \to [k]$; in other words, it equals $|\operatorname{Sur}(X, [k])| = \operatorname{sur}(|X|, k)$. This proves Proposition 0.11.

Here is a more rigorous way to make this argument (without vague terms such as "information" and "encoded"). We are going to construct a bijection $\operatorname{SC}_k(X) \to \operatorname{Sur}(X, [k])$ (that is, a bijection between the set compositions of X into k parts and the surjections from X to $[k]$). Here is how:

- We define a map $\Phi : \operatorname{SC}_k(X) \to \operatorname{Sur}(X, [k])$ to be the map that sends any set composition $(S_1, S_2, \ldots, S_k) \in \operatorname{SC}_k(X)$ to the map

$$f : X \to [k],$$
$$x \mapsto (\text{the unique } i \in [k] \text{ such that } x \in S_i).$$

  It is easy to see that Φ is well-defined (indeed, for any set composition $(S_1, S_2, \ldots, S_k)$, the resulting map $f : X \to [k]$ is surjective, because the sets $S_1, S_2, \ldots, S_k$ are nonempty).

- We define a map $\Psi : \operatorname{Sur}(X, [k]) \to \operatorname{SC}_k(X)$ to be the map that sends any surjective map $f : X \to [k]$ to the set composition $(S_1, S_2, \ldots, S_k) \in \operatorname{SC}_k(X)$, where

$$S_i = f^{-1}(\{i\}) = \{x \in X \mid f(x) = i\} \qquad \text{for each } i \in [k].$$

  Again, it is easily shown that this map Ψ is well-defined.

[2]Each $x \in X$ must belong to **exactly one** part of $(S_1, S_2, \ldots, S_k)$ (because $(S_1, S_2, \ldots, S_k)$ is a set composition of X).

It is easy to check that the maps Φ and Ψ are mutually inverse, and thus are bijections. Hence, we have found a bijection $\operatorname{SC}_k(X) \to \operatorname{Sur}(X, [k])$. This lets us conclude that

$$|\operatorname{SC}_k(X)| = |\operatorname{Sur}(X, [k])| = \operatorname{sur}(|X|, k).$$

Thus, Proposition 0.11 is proven. □

Next, let us count set partitions of a given finite set X into a given number of parts:

Proposition 0.12. Let X be a finite set. Let $k \in \mathbb{N}$. Then, $|\operatorname{SP}_k(X)| = \dfrac{\operatorname{sur}(|X|, k)}{k!}$.

Let us give an informal proof of this proposition. A formalization of it (omitting only really straightforward details) is given in the Appendix below.

Outline of an informal proof of Proposition 0.12. Let us count the set compositions $(S_1, S_2, \ldots, S_k)$ of X into k parts in two different ways:

- On the one hand, this number is $|\operatorname{SC}_k(X)| = \operatorname{sur}(|X|, k)$ (by Proposition 0.11).
- On the other hand, we can construct a set composition $(S_1, S_2, \ldots, S_k)$ of X into k parts by first choosing a set partition of X into k parts (there are $|\operatorname{SP}_k(X)|$ ways to do this), and then choosing how to order it[3] (there are $k!$ ways to do this, since it has k distinct parts). Thus, the number of set compositions $(S_1, S_2, \ldots, S_k)$ of X into k parts equals $|\operatorname{SP}_k(X)| \cdot k!$.

Comparing these two results, we conclude that $\operatorname{sur}(|X|, k) = |\operatorname{SP}_k(X)| \cdot k!$. Thus, $\dfrac{\operatorname{sur}(|X|, k)}{k!} = |\operatorname{SP}_k(X)|$. This proves Proposition 0.12. □

The above proof outline glances over technicalities such as what "ordering" a set partition means, and why there are $k!$ ways to do this; this is the reason why I am calling it informal and give a more complete version in the Appendix below.

The method we used to prove Proposition 0.12 is another example of the "shepherd's principle" ("To count sheep, you count the legs and then you divide by 4"). Our sheep here were the set partitions in $\operatorname{SP}_k(X)$, whereas their legs were the set compositions in $\operatorname{SC}_k(X)$. A leg $(S_1, S_2, \ldots, S_k) \in \operatorname{SC}_k(X)$ belongs to the sheep $\{S_1, S_2, \ldots, S_k\} \in \operatorname{SP}_k(X)$. This principle (also known as "proof by multijection") is useful whenever the legs are easier to count than the sheep. (For example, in the above case, the legs were in bijection with the surjections $X \to [k]$, whereas the sheep were not in a visible bijection with anything.)

[3] i.e., in what order to list its parts

Definition 0.13. Let $n \in \mathbb{N}$ and $k \in \mathbb{N}$. The (n,k)*-th Stirling number of the 2nd kind* is defined to be $\dfrac{\operatorname{sur}(n,k)}{k!}$; it is commonly denoted by $\left\{ n \atop k \right\}$ or by $S(n,k)$.

Proposition 0.12 shows that $\left\{ n \atop k \right\}$ is the number of set partitions of a given n-element set into k parts. Thus, in particular, $\left\{ n \atop k \right\}$ is a nonnegative integer.

The recurrence relation $\operatorname{sur}(n,k) = k \cdot (\operatorname{sur}(n-1,k) + \operatorname{sur}(n-1,k-1))$ for arbitrary $n > 0$ and $k > 0$ (see, for example, Proposition 3.12 in classwork from 21 February 2018) now leads to the following recurrence relations for Stirling numbers of the 2nd kind:

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

Meanwhile, Exercise 2 **(b)** leads to

$$\left\{ a \atop b \right\} = \frac{1}{b!} \sum_{k=0}^{b} (-1)^{b-k} \binom{b}{k} k^a.$$

Without knowing about set partitions, would you have guessed that the right hand side is a nonnegative integer?

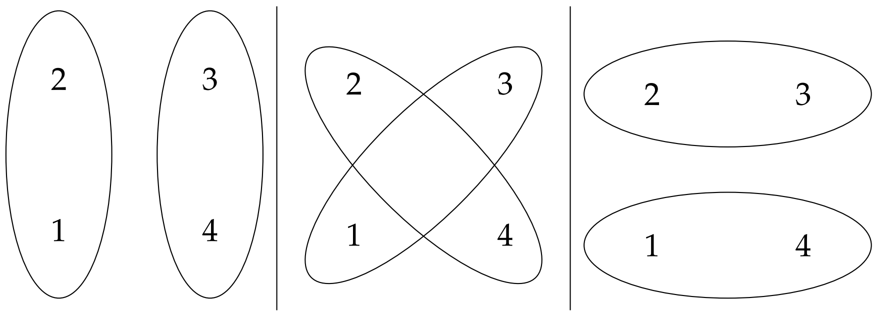
Now, I want you to count a special kind of set partitions:

Definition 0.14. A *perfect matching* of a set X means a set partition $\mathbf{P}$ of X such that each part of $\mathbf{P}$ has size 2.

Example 0.15. (a) The perfect matchings of the set $[4]$ are

$$\{\{1,2\},\{3,4\}\}, \qquad \{\{1,3\},\{2,4\}\}, \qquad \{\{1,4\},\{2,3\}\}.$$

Here are the same 3 perfect matchings, drawn as blobs:



(b) The set $[6]$ has 15 perfect matchings; three of them are

$$\{\{1,2\},\{3,4\},\{5,6\}\}, \qquad \{\{1,4\},\{2,6\},\{3,5\}\}, \qquad \{\{1,6\},\{2,5\},\{3,4\}\}.$$

Clearly, a perfect matching of a finite set X must have precisely $|X|/2$ parts (since $|X|$ is the sum of the sizes of all parts, but these sizes all equal 2). Thus, a perfect matching of a finite set X can only exist when $|X|$ is even.

Definition 0.16. Let A and B be two sets. Let $j \in \mathbb{N}$. A map $f : A \to B$ will be called *j-multijective* if for each $b \in B$, we have

$$(\text{the number of } a \in A \text{ satisfying } f(a) = b) = j.$$

In the classical picture illustrating a map, a map $f : A \to B$ is j-multijective if and only if each element of B is hit by exactly j arrows.

For example, a 1-multijective map is the same as a bijective map (make sure you understand why).

For another example, Proposition 0.19 below says that the map π in that proposition is $k!$-multijective.

0.3.3. Exercise 3: counting perfect matchings

Exercise 3. Let $n \in \mathbb{N}$. Prove the following:

(a) The number of all 2-multijective maps from $[2n]$ to $[n]$ is $\dfrac{(2n)!}{2^n}$.

(b) The number of all set compositions $\mathbf{C}$ of $[2n]$ such that each part of $\mathbf{C}$ has size 2 is $\dfrac{(2n)!}{2^n}$.

(c) The number of all perfect matchings of $[2n]$ is $\dfrac{(2n)!}{2^n n!}$.

[Hint: You don't need to imitate the level of detail that is given in the Appendix.]

Solution to Exercise 3 (sketched). **(a)** A map $f : [2n] \to [n]$ is 2-multijective if and only if for each $i \in [n]$, there are exactly two elements of $[2n]$ that are mapped to i under f. Thus, the following algorithm lets us construct any 2-multijective map f from $[2n]$ to $[n]$:

- First, decide which two elements of $[2n]$ will be mapped to 1 under f. There are $\dbinom{2n}{2}$ choices, because we are picking two distinct elements from the $2n$-element set $[2n]$ (with no regard for order).
- Then, decide which two elements of $[2n]$ will be mapped to 2 under f. There are $\dbinom{2n-2}{2}$ choices, because we are picking two distinct elements from the $2n$-element set $[2n]$ but allowing none of the 2 elements chosen before (so we are really picking two elements from a $2n-2$-element set).
- Then, decide which two elements of $[2n]$ will be mapped to 3 under f. There are $\dbinom{2n-4}{2}$ choices, because we are picking two distinct elements from the $2n$-element set $[2n]$ but allowing none of the 4 elements chosen before (so we are really picking two elements from a $2n-4$-element set).

- Proceed likewise, until finally deciding which two elements of $[2n]$ will be mapped to n.

In total, this algorithm can be performed in

$$\binom{2n}{2}\binom{2n-2}{2}\binom{2n-4}{2}\cdots\binom{2n-2(n-1)}{2}=\prod_{i=0}^{n-1}\binom{2n-2i}{2}$$

many ways. Since this algorithm constructs each 2-multijective map f from $[2n]$ to $[n]$ exactly once, we thus conclude that the number of all 2-multijective maps from $[2n]$ to $[n]$ is

$$\begin{aligned}\prod_{i=0}^{n-1}\binom{2n-2i}{2} &= \prod_{k=1}^{n}\underbrace{\binom{2k}{2}}_{=\frac{(2k-1)(2k)}{2}} \\ &\qquad \text{(here, we have substituted } k \text{ for } n-i \text{ in the product)} \\ &= \prod_{k=1}^{n}\frac{(2k-1)(2k)}{2} = \frac{1}{2^n}\underbrace{\prod_{k=1}^{n}((2k-1)(2k))}_{\substack{=(1\cdot 2)(3\cdot 4)\cdots((2n-1)\cdot(2n)) \\ =1\cdot 2\cdot 3\cdot 4\cdot\cdots\cdot(2n-1)\cdot(2n)=(2n)!}} \\ &= \frac{1}{2^n}(2n)! = \frac{(2n)!}{2^n}.\end{aligned}$$

This solves Exercise 3 **(a)**.

Now, let X be the set $[2n]$. A 2-*block composition* of X shall mean a set composition $\mathbf{C}$ of X such that each part of $\mathbf{C}$ has size 2. Thus, each 2-block composition of X has exactly n parts. (Indeed, each part of a 2-block composition has size 2, but the sum of the sizes of all its parts is $|X| = 2n$; therefore, the number of parts must be $2n/2 = n$.) Thus, each 2-block composition of X can be written in the form $(S_1, S_2, \ldots, S_n)$.

(b) The set compositions $\mathbf{C}$ of $[2n]$ such that each part of $\mathbf{C}$ has size 2 are precisely the 2-block compositions of X (because this is how we defined the latter). Thus, we must prove that the number of all 2-block compositions of X is $\frac{(2n)!}{2^n}$.

We shall derive this from Exercise 3 **(a)** in roughly the same way as we proved Proposition 0.11.

For any sets A and B, we let $\operatorname{Mulj}_2(A, B)$ be the set of all 2-multijections from A to B. Thus, Exercise 3 **(a)** says that $|\operatorname{Mulj}_2([2n], [n])| = \frac{(2n)!}{2^n}$.

We want to count all 2-block compositions $(S_1, S_2, \ldots, S_n)$ of X. We can construct such a 2-block composition by choosing, for each $x \in X$, which part S_i it shall belong to, as long as we ensure that each part S_i ends up containing exactly 2 elements $x \in X$. This information can be encoded as a map $f : X \to [n]$ (which sends each $x \in X$ to the $i \in [n]$ satisfying $x \in S_i$); this map f has to be 2-multijective (because this says precisely that each part S_i contains exactly 2 elements $x \in X$), but otherwise is subject to no constraints. Hence, the number of 2-block compositions $(S_1, S_2, \ldots, S_n)$ of X equals the number of 2-multijective maps $X \to [n]$; but this number is

$$\left| \operatorname{Mulj}_2 \left(\underbrace{X}_{=[2n]}, [n] \right) \right| = \left| \operatorname{Mulj}_2 \left([2n], [n] \right) \right| = \frac{(2n)!}{2^n}.$$

This solves Exercise 3 **(b)**.

If you want to make the above argument more rigorous, you can proceed as in the proof of Proposition 0.11, namely by constructing a bijection $\operatorname{SCTB}(X) \to \operatorname{Mulj}_2(X, [n])$, where $\operatorname{SCTB}(X)$ denotes the set of all 2-block compositions of X. This bijection is defined exactly as in the proof of Proposition 0.11 (but with k replaced by n, and other obvious changes).

(c) This is similar to how we derived Proposition 0.12 from Proposition 0.11 (with the difference that instead of Proposition 0.11, we now need to use Exercise 3 **(b)**). Let us just give the informal outline:

Let $\operatorname{PM}(X)$ be the set of all perfect matchings of X. Each perfect matching of X has exactly n parts. (Indeed, each part of a perfect matching has size 2, but the sum of the sizes of all its parts is $|X| = 2n$; therefore, the number of parts must be $2n/2 = n$.)

Recall that each 2-block composition of X can be written in the form $(S_1, S_2, \ldots, S_n)$. Let us count all 2-block compositions $(S_1, S_2, \ldots, S_n)$ of X in two different ways:

- On the one hand, this number is $\dfrac{(2n)!}{2^n}$ (by Exercise 3 **(b)**).
- On the other hand, we can construct a 2-block composition $(S_1, S_2, \ldots, S_n)$ of X by first choosing a perfect matching of X (there are $|\operatorname{PM}(X)|$ ways to do this), and then choosing how to order it (there are $n!$ ways to do this, since it has n distinct parts). Thus, the number of 2-block compositions of X equals $|\operatorname{PM}(X)| \cdot n!$.

Comparing these two results, we conclude that $\dfrac{(2n)!}{2^n} = |\operatorname{PM}(X)| \cdot n!$. Thus, $|\operatorname{PM}(X)| = \dfrac{(2n)!}{2^n} / n! = \dfrac{(2n)!}{2^n n!}$.

But $\operatorname{PM}(X)$ is the set of all perfect matchings of $X = [2n]$. Hence, the number of all perfect matchings of $[2n]$ is $|\operatorname{PM}(X)| = \dfrac{(2n)!}{2^n n!}$. This solves Exercise 3 **(c)**. □

0.4. "Image-injective maps"

If S is a set, then a map $f : S \to S$ is said to be *image-injective*[4] if and only if its restriction $f \mid_{f(S)}$ is injective. For example:

[4]This is my terminology; don't expect to see it in the literature.

- The map $[4] \to [4]$ sending $1, 2, 3, 4$ to $4, 1, 4, 1$ (respectively) is image-injective (since its image is $\{1, 4\}$, and its restriction to $\{1, 4\}$ is injective).
- The map $[6] \to [6]$ sending $1, 2, 3, 4, 5, 6$ to $2, 4, 4, 6, 6, 2$ (respectively) is image-injective (since its image is $\{2, 4, 6\}$, and its restriction to $\{2, 4, 6\}$ is injective).
- Any injective map $f : S \to S$ is image-injective. So is any constant map (i.e., any map $f : S \to S$ such that all values of f are equal).
- The map $[3] \to [3]$ sending $1, 2, 3$ to $2, 2, 1$ (respectively) is not image-injective (since its restriction to its image $\{1, 2\}$ is not injective).

As usual, if S is a set, and $f : S \to S$ is a map, then f^2 means the map $f \circ f : S \to S$. (More generally, f^k means the map $\underbrace{f \circ f \circ \cdots \circ f}_{k \text{ times}}$ whenever $k \in \mathbb{N}$.)

Exercise 4. Let $n \in \mathbb{N}$.

(a) Prove that a map $f : [n] \to [n]$ is image-injective if and only if it has the following property: For any $a \in [n]$ and $b \in [n]$ satisfying $f^2(a) = f^2(b)$, we must have $f(a) = f(b)$.

(b) Prove that the number of image-injective maps $[n] \to [n]$ is $\sum_{k=0}^{n} \binom{n}{k} k! k^{n-k}$.

(c) Prove that the number of image-injective maps $[n] \to [n]$ is divisible by n whenever n is positive.

[**Hint:** This is rather similar to [Fall2017-HW6os, Exercise 1]; feel free to imitate the solution of the latter exercise.]

The sequence of integers whose n-th term is the number of image-injective maps $[n] \to [n]$ appears in the OEIS as Sequence A006153; our definition of it is exactly what Geoffrey Critzer wrote in his comment. It doesn't seem to have many properties.

The solution to Exercise 4 outlined below imitates [Fall2017-HW6os, solution to Exercise 1], to the point that several parts of it are copied more or less verbatim.

Solution to Exercise 4 (sketched). **(a)** Let $f : [n] \to [n]$ be a map. We must prove that f is image-injective if and only if it has the following property:

$$\left(\begin{array}{c} \text{For any } a \in [n] \text{ and } b \in [n] \text{ satisfying } f^2(a) = f^2(b), \\ \text{we must have } f(a) = f(b) \end{array} \right). \tag{19}$$

$\Longrightarrow$: Assume that f is image-injective. We must prove that f has the property (19).

The map f is image-injective. In other words, its restriction $f \mid_{f([n])}$ is injective (by the definition of "image-injective").

Let $a \in [n]$ and $b \in [n]$ be such that $f^2(a) = f^2(b)$. Both $f(a)$ and $f(b)$ clearly belong to the image $f([n])$. Furthermore,

$$\left(f\mid_{f([n])}\right)(f(a)) = f(f(a)) = f^2(a) = f^2(b) = f(f(b)) = \left(f\mid_{f([n])}\right)(f(b)).$$

Since $f\mid_{f([n])}$ is injective, we thus conclude that $f(a) = f(b)$.

Now, forget that we fixed a and b. We thus have proven that for any $a \in [n]$ and $b \in [n]$ satisfying $f^2(a) = f^2(b)$, we must have $f(a) = f(b)$. In other words, f has the property (19). This proves the "$\Longrightarrow$" direction of Exercise 4 **(a)**.

$\Longleftarrow$: Assume that f has the property (19). We must prove that f is image-injective.

Let $x \in f([n])$ and $y \in f([n])$ be such that $\left(f\mid_{f([n])}\right)(x) = \left(f\mid_{f([n])}\right)(y)$. Then, $\left(f\mid_{f([n])}\right)(x) = f(x)$, so that $f(x) = \left(f\mid_{f([n])}\right)(x) = \left(f\mid_{f([n])}\right)(y) = f(y)$.

But $x \in f([n])$; thus, there exists some $a \in [n]$ such that $x = f(a)$. Consider this a.

Also, $y \in f([n])$; thus, there exists some $b \in [n]$ such that $y = f(b)$. Consider this b.

We have $f^2(a) = f\left(\underbrace{f(a)}_{=x}\right) = f(x) = f\left(\underbrace{y}_{=f(b)}\right) = f(f(b)) = f^2(b)$. Hence, (19) shows that $f(a) = f(b)$. In view of $x = f(a)$ and $y = f(b)$, this rewrites as $x = y$.

Now, forget that we fixed x and y. We thus have shown that if $x \in f([n])$ and $y \in f([n])$ are such that $\left(f\mid_{f([n])}\right)(x) = \left(f\mid_{f([n])}\right)(y)$, then $x = y$. In other words, the map $f\mid_{f([n])}$ is injective. In other words, the map f is image-injective (by the definition of "image-injective"). This proves the "$\Longleftarrow$" direction of Exercise 4 **(a)**.

Hence, Exercise 4 **(a)** is solved (since we have proven both of its directions).

(b) We first observe the following facts:

> *Observation 1:* Let $f : [n] \to [n]$ be an image-injective map. Let $S = f([n])$. Then:
>
> **(a)** The restriction $f\mid_S : S \to [n]$ is injective.
>
> **(b)** We have $(f\mid_S)(S) \subseteq S$.

[*Proof of Observation 1:* **(a)** The map f is image-injective. In other words, the restriction $f\mid_{f([n])}$ is injective (by the definition of "image-injective"). Since $S = f([n])$, this rewrites as follows: The restriction $f\mid_S$ is injective. This proves Observation 1 **(a)**.

(b) We have $(f\mid_S)(x) = f(x)$ for each $x \in S$. Thus, $(f\mid_S)(S) = f\left(\underbrace{S}_{\subseteq [n]}\right) \subseteq f([n]) = S$. This proves Observation 1 **(b)**.]

Observation 2: Let $f : [n] \to [n]$ be a map. Let S be a subset of $[n]$. Assume that the restriction $f \mid_S : S \to [n]$ is injective and satisfies $(f \mid_S)(S) \subseteq S$. Assume furthermore that $f(x) \in S$ for each $x \in [n] \setminus S$. Then:

(a) The map f is image-injective.

(b) We have $f([n]) = S$.

[*Proof of Observation 2:* **(b)** We have $f(x) \in S$ for each $x \in [n] \setminus S$. In other words, we have $f(x) \in S$ for each $x \in [n]$ that does not belong to S. But we also have $f(x) \in S$ for each $x \in [n]$ that does belong to S (because each $x \in [n]$ that does belong to S satisfies $f(x) = (f \mid_S)\left(\underbrace{x}_{\in S}\right) \in (f \mid_S)(S) \subseteq S$). Combining the previous two sentences, we conclude that $f(x) \in S$ for each $x \in [n]$ (whether or not x belongs to S). In other words, $\{f(x) \mid x \in [n]\} \subseteq S$. Hence, $f([n]) = \{f(x) \mid x \in [n]\} \subseteq S$.

But the restriction $f \mid_S : S \to [n]$ is injective. Hence, this restriction must take exactly $|S|$ distinct values. In other words, $|(f \mid_S)(S)| = |S|$. Thus, $(f \mid_S)(S)$ cannot be a **proper** subset of S (because if it was, then it would satisfy $|(f \mid_S)(S)| < |S|$, which would contradict $|(f \mid_S)(S)| = |S|$). So we know that $(f \mid_S)(S)$ is a subset of S (since $(f \mid_S)(S) \subseteq S$), but not a proper subset of S. Therefore, $(f \mid_S)(S) = S$.

But $(f \mid_S)(x) = f(x)$ for each $x \in S$. Hence, $(f \mid_S)(S) = f(S)$, so that $f(S) = (f \mid_S)(S) = S$.

Finally, $[n] \supseteq S$, so that $f([n]) \supseteq f(S) = S$. Combining this with $f([n]) \subseteq S$, we obtain $f([n]) = S$. This proves Observation 2 **(b)**.

(a) Observation 2 **(b)** yields that $f([n]) = S$. But the restriction $f \mid_S$ is injective (by assumption). In view of $f([n]) = S$, this rewrites as follows: The restriction $f \mid_{f([n])}$ is injective. In other words, the map f is image-injective (by the definition of "image-injective"). This proves Observation 2 **(a)**.]

The following algorithm constructs every image-injective map $f : [n] \to [n]$:

- First, we choose an integer $k \in \{0, 1, \ldots, n\}$. This integer k shall be the size $|f([n])|$ of the image of f. (Of course, this size has to be in $\{0, 1, \ldots, n\}$, because $f([n])$ must be a subset of $[n]$.)
- Next, we choose a k-element subset S of $[n]$. This subset S shall be the image $f([n])$ of f. There are $\dbinom{n}{k}$ choices for S (since the number of k-element subsets of $[n]$ is $\dbinom{n}{k}$).
- Now, we choose the values $f(x)$ for all $x \in S$. In other words, we choose the restriction $f \mid_S : S \to [n]$. We cannot choose it arbitrarily; in fact, this restriction $f \mid_S$ must be an injective map (because we want f to be image-injective, and

thus Observation 1 **(a)** dictates that the restriction $f\mid_S : S \to [n]$ is injective) and must satisfy $(f\mid_S)(S) \subseteq S$ (since we want f to be image-injective, and thus Observation 1 **(b)** dictates that $(f\mid_S)(S) \subseteq S$). Thus, we must choose an injective map $f\mid_S : S \to [n]$ satisfying $(f\mid_S)(S) \subseteq S$. Proposition 0.5 **(c)** (applied to $A = S$, $B = [n]$ and $C = S$) shows that the number of such maps is $|S|!$; in other words, it is $k!$ (since $|S| = k$).

- Finally, we choose the values of f on all remaining elements of $[n]$ (that is, on all elements of $[n] \setminus S$). These values must belong to S (because we want $f([n])$ to be S), but are otherwise unconstrained[5]. Thus, there are $|S|^{|[n]\setminus S|}$ choices at this step. In other words, there are k^{n-k} choices at this step (since $|S| = k$ and $|[n] \setminus S| = n - \underbrace{|S|}_{=k} = n - k$).

It is easy to check that this algorithm really constructs image-injective maps $f : [n] \to [n]$, and constructs each of them exactly once. Thus, the number of image-injective maps $f : [n] \to [n]$ is

$$\sum_{k\in\{0,1,\ldots,n\}} \binom{n}{k} k! k^{n-k}$$

(since we get to choose $k \in \{0, 1, \ldots, n\}$ in the first step of the algorithm, then we have $\binom{n}{k}$ choices in the second step, then $k!$ choices in the third step, and finally k^{n-k} choices in the fourth step). Hence, the number of image-injective maps $f : [n] \to [n]$ is

$$\sum_{k\in\{0,1,\ldots,n\}} \binom{n}{k} k! k^{n-k} = \sum_{k=0}^{n} \binom{n}{k} k! k^{n-k}.$$

This solves Exercise 4 **(b)**.

(c) Assume that n is positive.

[5]Indeed, by requiring that these values belong to S, we have ensured that our map f satisfies $f(x) \in S$ for each $x \in [n] \setminus S$. Furthermore, in the previous step, we have already ensured that the restriction $f\mid_S : S \to [n]$ is injective and satisfies $(f\mid_S)(S) \subseteq S$. Thus, Observation 2 **(a)** yields that our map f is image-injective, and Observation 2 **(b)** yields that it satisfies $f([n]) = S$ (so that the set S, which was meant to be the image of f, will indeed be the image of f).

Exercise 4 **(b)** shows that the number of image-injective maps $[n] \to [n]$ is

$$\sum_{k=0}^{n} \binom{n}{k} k! k^{n-k} = \binom{n}{0} 0! \underbrace{0^{n-0}}_{\substack{=0^n=0 \\ (\text{since } n>0)}} + \sum_{k=1}^{n} \underbrace{\binom{n}{k}}_{\substack{=\frac{n}{k}\binom{n-1}{k-1} \\ (\text{by (6)})}} \underbrace{k!}_{\substack{=k\cdot(k-1)! \\ (\text{since } k\geq 1)}} k^{n-k}$$

$$= \sum_{k=1}^{n} \underbrace{\frac{n}{k}\binom{n-1}{k-1} k}_{=n\binom{n-1}{k-1}} \cdot (k-1)! k^{n-k} = \sum_{k=1}^{n} n \underbrace{\binom{n-1}{k-1} \cdot (k-1)! k^{n-k}}_{\substack{\text{this is an integer} \\ (\text{since } n-1\in\mathbb{Z} \text{ and } k-1\in\mathbb{Z} \\ \text{and } n-k\in\mathbb{N})}}$$

$$= \sum_{k=1}^{n} n \cdot (\text{some integer}) = n \cdot (\text{some integer}).$$

Hence, this number is divisible by n. This solves Exercise 4 **(c)**. □

0.5. Counting certain tuples

Exercise 5. Let $n \in \mathbb{N}$, and let d be a positive integer.

An n-tuple $(x_1, x_2, \ldots, x_n) \in [d]^n$ will be called 1-*even* if the number 1 occurs in it an even number of times (i.e., the number of $i \in [n]$ satisfying $x_i = 1$ is even). (For example, the 3-tuples $(1,5,1)$ and $(3,2,6)$ are 1-even (yes, 0 is an even number), while the 3-tuple $(2,1,4)$ is not.)

Prove that the number of 1-even n-tuples in $[d]^n$ is $\frac{1}{2}\left(d^n + (d-2)^n\right)$.

[**Hint:** Set $e = d - 1$; then, $(d-2)^n = (e-1)^n$ and $d^n = (e+1)^n$. There might also be a bijective proof – after multiplying by 2 –, but I don't know it.]

Exercise 5 is [Masulo11, Example 1.13]. Our proof of it (taken from [Masulo11, Example 1.13]) relies on the following fact:

Lemma 0.17. Let $n \in \mathbb{N}$ and $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$. Then,

$$(x+y)^n + (-x+y)^n = 2 \sum_{\substack{k\in\{0,1,\ldots,n\}; \\ k \text{ is even}}} \binom{n}{k} x^k y^{n-k}.$$

Proof of Lemma 0.17 (sketched). The binomial formula yields

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

The binomial formula (applied to $-x$ instead of x) yields

$$(-x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \underbrace{(-x)^k}_{=(-1)^k x^k} y^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^k y^{n-k}.$$

Adding these two equalities together, we obtain

$$\begin{aligned}
&(x+y)^n + (-x+y)^n \\
&= \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^k y^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \left(1+(-1)^k\right) x^k y^{n-k} \\
&= \sum_{\substack{k\in\{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} \underbrace{\left(1+(-1)^k\right)}_{\substack{=2 \\ (\text{since } (-1)^k=1 \\ (\text{since } k \text{ is even}))}} x^k y^{n-k} + \sum_{\substack{k\in\{0,1,\ldots,n\};\\ k \text{ is odd}}} \binom{n}{k} \underbrace{\left(1+(-1)^k\right)}_{\substack{=0 \\ (\text{since } (-1)^k=-1 \\ (\text{since } k \text{ is odd}))}} x^k y^{n-k} \\
&= \sum_{\substack{k\in\{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} 2x^k y^{n-k} + \underbrace{\sum_{\substack{k\in\{0,1,\ldots,n\};\\ k \text{ is odd}}} \binom{n}{k} 0x^k y^{n-k}}_{=0} = 2 \sum_{\substack{k\in\{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} x^k y^{n-k}.
\end{aligned}$$

This proves Lemma 0.17. □

Solution to Exercise 5 (sketched). Set $e = d - 1$.

We can construct any 1-even n-tuple $(x_1, x_2, \ldots, x_n) \in [d]^n$ using the following algorithm:

- First, we choose the number k of times the entry 1 will appear in this n-tuple. This number k must be even (since we want our n-tuple to be 1-even), and must belong to $\{0, 1, \ldots, n\}$.
- Then, we choose the k positions in which this n-tuple will have the entry 1 (in other words, choose the k indices $i \in [n]$ that will satisfy $x_i = 1$). This choice can be made in $\binom{n}{k}$ many ways (since we are choosing k out of n possible indices).
- Next, choose the entries in the remaining $n - k$ positions of our n-tuple. The entries can be arbitrary, except that they must be distinct from 1 (since we have already chosen the entries that will equal 1). Thus, there are $d - 1 = e$ choices for each entry, and therefore e^{n-k} choices altogether in this step.

Thus, the total number of 1-even n-tuples is $\sum_{\substack{k\in\{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} e^{n-k}$.

But Lemma 0.17 (applied to $x = 1$ and $y = e$) yields

$$(1+e)^n + (-1+e)^n = 2 \sum_{\substack{k \in \{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} \underbrace{1^k}_{=1} e^{n-k} = 2 \sum_{\substack{k \in \{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} e^{n-k}.$$

Hence,

$$\sum_{\substack{k \in \{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} e^{n-k} = \frac{1}{2} \left(\left(1 + \underbrace{e}_{=d-1} \right)^n + \left(-1 + \underbrace{e}_{=d-1} \right)^n \right)$$

$$= \frac{1}{2} \left(\left(\underbrace{1 + d - 1}_{=d} \right)^n + \left(\underbrace{-1 + d - 1}_{=d-2} \right)^n \right) = \frac{1}{2} \left(d^n + (d-2)^n \right).$$

Hence, the total number of 1-even n-tuples is $\sum_{\substack{k \in \{0,1,\ldots,n\};\\ k \text{ is even}}} \binom{n}{k} e^{n-k} = \frac{1}{2} \left(d^n + (d-2)^n \right)$.

This solves Exercise 5. □

0.6. And more binomial identities

Exercise 6. **(a)** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Prove that every $j \in \{0, 1, \ldots, n\}$ satisfies

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

(b) Let x and y be two real numbers. Let $z = x + y$. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Prove that

$$x^{m+1} \sum_{k=0}^{n} \binom{m+k}{k} y^k z^{n-k} = \sum_{i=m+1}^{n+m+1} \binom{n+m+1}{i} x^i y^{(n+m+1)-i}$$

and

$$y^{n+1} \sum_{k=0}^{m} \binom{n+k}{k} x^k z^{m-k} = \sum_{i=0}^{m} \binom{n+m+1}{i} x^i y^{(n+m+1)-i}$$

and

$$x^{m+1} \sum_{k=0}^{n} \binom{m+k}{k} y^k z^{n-k} + y^{n+1} \sum_{k=0}^{m} \binom{n+k}{k} x^k z^{m-k} = z^{n+m+1}.$$

(c) Let $n \in \mathbb{N}$. Prove that

$$\sum_{k=0}^{n} \binom{n+k}{k} \frac{1}{2^k} = 2^n.$$

[**Hint:** Part **(a)** is a restatement of something proven in class. Derive **(b)** from **(a)**, and **(c)** from **(b)**.]

We shall outline a solution of this exercise. (For the missing details, see [Grinbe16, solution to Exercise 3.27].)

Solution to Exercise 6 (sketched). **(a)** Let $j \in \{0, 1, \ldots, n\}$. Then, (5) (applied to $n+m$, m and j instead of n, x and y) yields

$$\begin{aligned}
\binom{n+m+1}{m+j+1} &= \sum_{k=0}^{n+m} \binom{k}{m} \binom{n+m-k}{j} \\
&= \sum_{k=0}^{m-1} \underbrace{\binom{k}{m}}_{\substack{=0 \\ \text{(by (1))}}} \binom{n+m-k}{j} + \sum_{k=m}^{n+m} \binom{k}{m} \binom{n+m-k}{j} \\
&= \sum_{k=m}^{n+m} \binom{k}{m} \binom{n+m-k}{j} = \sum_{k=0}^{n} \underbrace{\binom{m+k}{m}}_{\substack{=\binom{m+k}{k} \\ \text{(by (4))}}} \underbrace{\binom{n+m-(m+k)}{j}}_{=\binom{n-k}{j}} \\
&\qquad \text{(here, we have substituted } m+k \text{ for } k \text{ in the sum)} \\
&= \sum_{k=0}^{n} \binom{m+k}{k} \binom{n-k}{j}.
\end{aligned}$$

This solves Exercise 6 **(a)**.

(b) Let $k \in \{0, 1, \ldots, n\}$. From $z = x + y$, we obtain

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

(by the binomial formula). Multiplying both sides of this equality by y^k, we find

$$y^k z^{n-k} = \sum_{j=0}^{n-k} \binom{n-k}{j} x^j \underbrace{y^{(n-k)-j} y^k}_{=y^{n-j}} = \sum_{j=0}^{n-k} \binom{n-k}{j} x^j y^{n-j}.$$

Comparing this with

$$\begin{aligned}
\sum_{j=0}^{n} \binom{n-k}{j} x^j y^{n-j} &= \sum_{j=0}^{n-k} \binom{n-k}{j} x^j y^{n-j} + \sum_{j=n-k+1}^{n-k} \underbrace{\binom{n-k}{j}}_{\substack{=0 \\ \text{(by (1))}}} x^j y^{n-j} \\
&= \sum_{j=0}^{n-k} \binom{n-k}{j} x^j y^{n-j},
\end{aligned}$$

we find

$$y^k z^{n-k} = \sum_{j=0}^{n} \binom{n-k}{j} x^j y^{n-j}. \tag{20}$$

Now, forget that we fixed k. We thus have proven the equality (20) for each $k \in \{0, 1, \ldots, n\}$.

Now,

$$\begin{aligned}
& x^{m+1} \sum_{k=0}^{n} \binom{m+k}{k} \underbrace{y^k z^{n-k}}_{\substack{= \sum_{j=0}^{n} \binom{n-k}{j} x^j y^{n-j} \\ \text{(by (20))}}} \\
&= x^{m+1} \sum_{k=0}^{n} \binom{m+k}{k} \sum_{j=0}^{n} \binom{n-k}{j} x^j y^{n-j} = x^{m+1} \sum_{j=0}^{n} \underbrace{\left(\sum_{k=0}^{n} \binom{m+k}{k} \binom{n-k}{j} \right)}_{\substack{= \binom{n+m+1}{m+j+1} \\ \text{(by Exercise 6 (a))}}} x^j y^{n-j} \\
&= x^{m+1} \sum_{j=0}^{n} \binom{n+m+1}{m+j+1} x^j y^{n-j} = \sum_{j=0}^{n} \binom{n+m+1}{m+j+1} \underbrace{x^{m+1} x^j}_{= x^{m+1+j}} y^{n-j} \\
&= \sum_{j=0}^{n} \binom{n+m+1}{m+j+1} x^{m+j+1} y^{n-j} = \sum_{i=m+1}^{n+m+1} \binom{n+m+1}{i} x^i y^{(n+m+1)-i} \qquad (21) \\
&\qquad \left(\text{here, we have substituted } i \text{ for } (m+1)+j \text{ in the sum}\right).
\end{aligned}$$

This proves the first of the three equalities we need to show for Exercise 6 **(b)**.

Now, of course, there is nothing preventing us from applying (21) to y, x, m and n instead of x, y, n and m (after all, we have $z = x + y = y + x$, so the roles of x and y can be interchanged). Thus we find

$$\begin{aligned}
& y^{n+1} \sum_{k=0}^{m} \binom{n+k}{k} x^k z^{m-k} \\
&= \sum_{i=n+1}^{m+n+1} \binom{m+n+1}{i} y^i x^{(m+n+1)-i} = \sum_{i=n+1}^{n+m+1} \underbrace{\binom{n+m+1}{i}}_{\substack{= \binom{n+m+1}{(n+m+1)-i} \\ \text{(by (3))}}} \underbrace{y^i x^{(n+m+1)-i}}_{= x^{(n+m+1)-i} y^i} \\
&= \sum_{i=n+1}^{n+m+1} \binom{n+m+1}{(n+m+1)-i} x^{(n+m+1)-i} y^i = \sum_{i=0}^{m} \binom{n+m+1}{i} x^i y^{(n+m+1)-i} \qquad (22) \\
&\qquad \left(\text{here, we have substituted } i \text{ for } (n+m+1)-i \text{ in the sum}\right).
\end{aligned}$$

This proves the second of the three equalities we need to show for Exercise 6 **(b)**.

To prove the third equality, we add the equalities (21) and (22) together. We obtain

$$\begin{aligned} & x^{m+1}\sum_{k=0}^{n}\binom{m+k}{k}y^k z^{n-k}+y^{n+1}\sum_{k=0}^{m}\binom{n+k}{k}x^k z^{m-k} \\ &= \sum_{i=m+1}^{n+m+1}\binom{n+m+1}{i}x^i y^{(n+m+1)-i}+\sum_{i=0}^{m}\binom{n+m+1}{i}x^i y^{(n+m+1)-i} \\ &= \sum_{i=0}^{m}\binom{n+m+1}{i}x^i y^{(n+m+1)-i}+\sum_{i=m+1}^{n+m+1}\binom{n+m+1}{i}x^i y^{(n+m+1)-i} \\ &= \sum_{i=0}^{n+m+1}\binom{n+m+1}{i}x^i y^{(n+m+1)-i} = (x+y)^{n+m+1} \qquad \text{(by the binomial formula)} \\ &= z^{n+m+1} \qquad (\text{since } x+y=z). \end{aligned} \tag{23}$$

This completes the solution of Exercise 6 **(b)**.

(c) Recall that $1+1=2$. Thus, applying (23) to $m=n$, $x=1$, $y=1$ and $z=2$, we find

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

This simplifies to

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

In other words, $2\sum_{k=0}^{n}\binom{n+k}{k}2^{n-k}=2^{n+n+1}$. Dividing this equality by 2^{n+1}, we obtain $\sum_{k=0}^{n}\binom{n+k}{k}\frac{1}{2^k}=2^n$. This solves Exercise 6 **(c)**. □

For two other solutions to Exercise 6 **(c)**, see [Engel98, Chapter 5, Example E18] and [Engel98, Chapter 8, Problem 4]. (I don't know if these solutions can be generalized to give proofs of (23).)

0.7. Appendix: Rigorous proof of Proposition 0.12

Let us now give a proof of Proposition 0.12 that keeps to the standards of rigor in most parts of mathematics. First, we need a lemma from the "isn't this obvious?" department:

Lemma 0.18. Let $k \in \mathbb{N}$. Let $x_1, x_2, \ldots, x_k$ be k distinct objects. Let $y_1, y_2, \ldots, y_k$ be k objects such that $\{x_1, x_2, \ldots, x_k\} = \{y_1, y_2, \ldots, y_k\}$. Then, there exists a permutation σ of $[k]$ such that every $i \in [k]$ satisfies $y_i = x_{\sigma(i)}$.

Roughly speaking, Lemma 0.18 says that if k distinct objects $x_1, x_2, \ldots, x_k$ form the same set as k objects $y_1, y_2, \ldots, y_k$ (which are not a-priori required to be distinct, but it follows easily that they are), then the objects $y_1, y_2, \ldots, y_k$ are just the objects $x_1, x_2, \ldots, x_k$ rearranged (the rearrangement is what the permutation σ is meant to take care of). Convince yourself that this is plausible before (or instead of) reading the following proof. Notice that $x_1, x_2, \ldots, x_k$ need to be distinct in Lemma 0.18; otherwise, the lemma would be easily disproven (e.g., we have $\{2,2,3\} = \{2,3,3\}$, but there is no way to get $2,3,3$ by rearranging $2,2,3$).

Proof of Lemma 0.18 (sketched). For each $i \in [k]$, there exists some $j \in [k]$ such that $y_i = x_j$ (since $y_i \in \{y_1, y_2, \ldots, y_k\} = \{x_1, x_2, \ldots, x_k\}$). Moreover, this j is unique (because if j_1 and j_2 were two such j's, then we would have $y_i = x_{j_1}$ and $y_i = x_{j_2}$, thus $x_{j_1} = y = x_{j_2}$, thus $j_1 = j_2$ because $x_1, x_2, \ldots, x_k$ are distinct). Let us denote this j by $\sigma(i)$. Thus, we have defined a $\sigma(i) \in [k]$ for each $i \in [k]$. In other words, we have defined a map $\sigma : [k] \to [k]$. Clearly, this map has the property that every $i \in [k]$ satisfies $y_i = x_{\sigma(i)}$ (because this is how $\sigma(i)$ was defined). Thus, in order to prove Lemma 0.18, it suffices to check that this map σ is a permutation of $[k]$.

Let $h \in [k]$. Then, $x_h \in \{x_1, x_2, \ldots, x_k\} = \{y_1, y_2, \ldots, y_k\}$. In other words, there exists some $i \in [k]$ such that $x_h = y_i$. Consider this i. Then, $\sigma(i)$ is the unique $j \in [k]$ such that $y_i = x_j$ (by the definition of $\sigma(i)$). But this unique j must be h (since $h \in [k]$ and $y_i = x_h$). Hence, $\sigma(i) = h$.

Now, forget that we fixed h. We thus have shown that for each $h \in [k]$, there exists some $i \in [k]$ such that $\sigma(i) = h$. In other words, the map σ is surjective. Since σ is a map from $[k]$ to $[k]$, this yields that σ is bijective (by the Pigeonhole Principle for surjections). In other words, σ is a permutation of $[k]$. As we said, this completes the proof of Lemma 0.18. □

Proposition 0.19. Let X be a set, and let $k \in \mathbb{N}$. Let $\pi : \mathrm{SC}_k(X) \to \mathrm{SP}_k(X)$ be the map that sends each set composition $(S_1, S_2, \ldots, S_k)$ to the set partition $\{S_1, S_2, \ldots, S_k\}$. (So what the map π does is forgetting the order of the parts. In the language of Example 0.8, this means forgetting the colors of the blobs.)

Then, for each set partition $\mathbf{P} \in \mathrm{SP}_k(X)$, we have

$$(\text{the number of all } \mathbf{C} \in \mathrm{SC}_k(X) \text{ satisfying } \pi(\mathbf{C}) = \mathbf{P}) = k!.$$

Example 0.20. For this example, pick $X = \{1,2,3\}$ and $k = 2$ and $\mathbf{P} = \{\{1,3\}, \{2\}\}$. Then, Proposition 0.19 says that the number of all set compositions $\mathbf{C} \in \mathrm{SC}_k(X)$ satisfying $\pi(\mathbf{C}) = \mathbf{P}$ is $2! = 2$. These two set compositions are $(\{1,3\}, \{2\})$ and $(\{2\}, \{1,3\})$.

Proof of Proposition 0.19 (sketched). Let $\mathbf{P} \in \mathrm{SP}_k(X)$. Thus, $\mathbf{P}$ is a set partition of X into k parts. We can thus write $\mathbf{P}$ as $\mathbf{P} = \{T_1, T_2, \ldots, T_k\}$, where $T_1, T_2, \ldots, T_k$ are some disjoint nonempty subsets of X satisfying $T_1 \cup T_2 \cup \cdots \cup T_k = X$. Consider these $T_1, T_2, \ldots, T_k$. These sets $T_1, T_2, \ldots, T_k$ are nonempty and disjoint, and therefore distinct (because if two of them were equal, then their disjointness would force them to be empty).

The rest is now easy: Informally speaking, the set compositions $\mathbf{C} \in \mathrm{SC}_k(X)$ satisfying $\pi(\mathbf{C}) = \mathbf{P}$ are just all possible ways to rearrange the k-tuple $(T_1, T_2, \ldots, T_k)$ (this follows from Lemma 0.18); thus, there are $k!$ of them (in fact, the sets $T_1, T_2, \ldots, T_k$

are distinct, so any way of rearranging them results in a different k-tuple). This proves Proposition 0.19.

Here is a more formal way to make this argument (read this if you care; I don't require this level of rigor in solutions): The k-tuple $(T_1, T_2, \ldots, T_k)$ is a set composition of X into k parts. Thus, for each permutation σ of $[k]$, the k-tuple $\left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right)$ also is a set composition of X into k parts, hence belongs to $\operatorname{SC}_k(X)$. Moreover, this set composition $\left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right)$ is a $\mathbf{C} \in \operatorname{SC}_k(X)$ satisfying $\pi(\mathbf{C}) = \mathbf{P}$ (because $\pi\left(\left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right)\right) = \left\{T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right\} = \{T_1, T_2, \ldots, T_k\} = \mathbf{P}$). Hence, we can define a map

$$\alpha : \{\text{permutations of } [k]\} \to \{\mathbf{C} \in \operatorname{SC}_k(X) \mid \pi(\mathbf{C}) = \mathbf{P}\},$$
$$\sigma \mapsto \left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right).$$

This map is injective[6] and surjective[7]; thus, it is bijective. We have therefore found a bijective map from $\{\text{permutations of } [k]\}$ to $\{\mathbf{C} \in \operatorname{SC}_k(X) \mid \pi(\mathbf{C}) = \mathbf{P}\}$. Hence,

$$\begin{aligned}|\{\text{permutations of } [k]\}| &= |\{\mathbf{C} \in \operatorname{SC}_k(X) \mid \pi(\mathbf{C}) = \mathbf{P}\}| \\ &= (\text{the number of all } \mathbf{C} \in \operatorname{SC}_k(X) \text{ satisfying } \pi(\mathbf{C}) = \mathbf{P}),\end{aligned}$$

so that

$$(\text{the number of all } \mathbf{C} \in \operatorname{SC}_k(X) \text{ satisfying } \pi(\mathbf{C}) = \mathbf{P}) = |\{\text{permutations of } [k]\}| = k!.$$

This proves Proposition 0.19. □

---

[6] *Proof.* Let σ and τ be two permutations of $[k]$ such that $\alpha(\sigma) = \alpha(\tau)$. We must show that $\sigma = \tau$.

Let $i \in [k]$. The definition of α yields $\alpha(\sigma) = \left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right)$ and $\alpha(\tau) = \left(T_{\tau(1)}, T_{\tau(2)}, \ldots, T_{\tau(k)}\right)$. Hence, the equality $\alpha(\sigma) = \alpha(\tau)$ rewrites as $\left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right) = \left(T_{\tau(1)}, T_{\tau(2)}, \ldots, T_{\tau(k)}\right)$. Thus, $T_{\sigma(i)} = T_{\tau(i)}$. Since the sets $T_1, T_2, \ldots, T_k$ are distinct, this entails $\sigma(i) = \tau(i)$.

Now, forget that we fixed i. We thus have shown that $\sigma(i) = \tau(i)$ for each $i \in [k]$. Therefore, $\sigma = \tau$. This completes the proof of the injectivity of α.

[7] *Proof.* Let $\mathbf{D} \in \{\mathbf{C} \in \operatorname{SC}_k(X) \mid \pi(\mathbf{C}) = \mathbf{P}\}$. We must prove that $\mathbf{D} = \alpha(\sigma)$ for some permutation σ of $[k]$.

We have $\mathbf{D} \in \{\mathbf{C} \in \operatorname{SC}_k(X) \mid \pi(\mathbf{C}) = \mathbf{P}\}$. In other words, $\mathbf{D} \in \operatorname{SC}_k(X)$ and $\pi(\mathbf{D}) = \mathbf{P}$. From $\mathbf{D} \in \operatorname{SC}_k(X)$, we conclude that $\mathbf{D}$ is a set composition of X into k parts. We can thus write $\mathbf{D}$ as $\mathbf{D} = (D_1, D_2, \ldots, D_k)$, where $D_1, D_2, \ldots, D_k$ are some disjoint nonempty subsets of X satisfying $D_1 \cup D_2 \cup \cdots \cup D_k = X$. Consider these $D_1, D_2, \ldots, D_k$.

From $\mathbf{D} = (D_1, D_2, \ldots, D_k)$, we conclude that $\pi(\mathbf{D}) = \pi((D_1, D_2, \ldots, D_k)) = \{D_1, D_2, \ldots, D_k\}$ (by the definition of π), so that $\{D_1, D_2, \ldots, D_k\} = \pi(\mathbf{D}) = \mathbf{P} = \{T_1, T_2, \ldots, T_k\}$. In other words, $\{T_1, T_2, \ldots, T_k\} = \{D_1, D_2, \ldots, D_k\}$. Since $T_1, T_2, \ldots, T_k$ are distinct, we can thus apply Lemma 0.18 to $x_i = T_i$ and $y_i = D_i$. We conclude that there exists a permutation σ of $[k]$ such that every $i \in [k]$ satisfies $D_i = T_{\sigma(i)}$. Consider this σ. We have $(D_1, D_2, \ldots, D_k) = \left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right)$ (since every $i \in [k]$ satisfies $D_i = T_{\sigma(i)}$). The definition of α yields

$$\alpha(\sigma) = \left(T_{\sigma(1)}, T_{\sigma(2)}, \ldots, T_{\sigma(k)}\right) = (D_1, D_2, \ldots, D_k) = \mathbf{D}.$$

Hence, we have found a permutation σ of $[k]$ such that $\mathbf{D} = \alpha(\sigma)$. This completes our proof of the surjectivity of α.

Proof of Proposition 0.12. Proposition 0.11 yields $|\mathrm{SC}_k(X)| = \mathrm{sur}(|X|, k)$. Thus,

$$\begin{aligned}
\mathrm{sur}(|X|, k) &= |\mathrm{SC}_k(X)| \\
&= (\text{the number of all } \mathbf{C} \in \mathrm{SC}_k(X)) \\
&= \sum_{\mathbf{P} \in \mathrm{SP}_k(X)} \underbrace{(\text{the number of all } \mathbf{C} \in \mathrm{SC}_k(X) \text{ satisfying } \pi(\mathbf{C}) = \mathbf{P})}_{\substack{=k! \\ \text{(by Proposition 0.19)}}} \\
&\qquad \left(\begin{array}{c}\text{here, we have subdivided our count according} \\ \text{to the value of } \pi(\mathbf{C})\end{array}\right) \\
&= \sum_{\mathbf{P} \in \mathrm{SP}_k(X)} k! = |\mathrm{SP}_k(X)| \cdot k!.
\end{aligned}$$

Dividing this equality by $k!$, we find $\dfrac{\mathrm{sur}(|X|, k)}{k!} = |\mathrm{SP}_k(X)|$. This proves Proposition 0.12. □

References

[Engel98] Arthur Engel, *Problem-Solving Strategies*, Springer 1998.

[Fall2017-HW6os] Darij Grinberg, *Math 4990 Fall 2017 (Darij Grinberg): homework set 6 with solutions.*
`http://www.cip.ifi.lmu.de/~grinberg/t/17f/hw6os.pdf`

[Grinbe09] Darij Grinberg, *6th QEDMO 2009, Problem 4 (the Cauchy identity).*
`http://www.cip.ifi.lmu.de/~grinberg/QEDMO6P4long.pdf`

[Grinbe16] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 10 January 2019.
`http://www.cip.ifi.lmu.de/~grinberg/primes2015/sols.pdf`
The numbering of theorems and formulas in this link might shift when the project gets updated; for a "frozen" version whose numbering is guaranteed to match that in the citations above, see `https://github.com/darijgr/detnotes/releases/tag/2019-01-10` .

[Masulo11] Dragan Mašulović, *Introduction to Discrete Mathematics*, 14 January 2011.
`http://www.algebra.uni-linz.ac.at/Students/DiskreteMathematik/ws10/DM-WS-2010-11.pdf`