1 Exercise 1

1.1 Problem

Let $A$ and $B$ be two sets, and let $f : A \to B$ be a map. A left inverse of $f$ shall mean a map $g : B \to A$ such that $g \circ f = \text{id}_A$. We say that $f$ is left-invertible if and only if a left inverse of $f$ exists. (It is usually not unique.)

Assume that the sets $A$ and $B$ are finite.

(a) If the set $A$ is nonempty, then prove that $f$ is left-invertible if and only if $f$ is injective.$^1$

(b) Assume that $f$ is injective. Prove that the number of left inverses of $f$ is $|A| |B| - |A|$.

1.2 Solution

We first prove a few claims (that hold even without requiring $A$ and $B$ to be finite):

Claim 1: If $f$ is left-invertible, then $f$ is injective.

---

$^1$This holds even when $A$ and $B$ are infinite. Feel free to prove this if you wish.
[Proof of Claim 1: Assume that \( f \) is left-invertible. In other words, \( f \) has a left inverse. In other words, there exists a left inverse \( g \) of \( f \). Consider this \( g \).

We know that \( g \) is a left inverse of \( f \). In other words, \( g \) is a map from \( B \) to \( A \) such that \( g \circ f = \text{id}_A \) (by the definition of a “left inverse”).

If \( a \) and \( b \) are two elements of \( A \) satisfying \( f(a) = f(b) \), then \( a = b \) (since

\[
a = \text{id}_A(a) = (g \circ f)(a) = g\left(f(a)\right) = g(f(b)) = (g \circ f)(b) = \text{id}_A(b) = b
\]

). In other words, the map \( f \) is injective. This proves Claim 1.]

Claim 2: Assume that \( f \) is injective.

(a) For each \( b \in f(A) \), there exists exactly one \( a \in A \) satisfying \( f(a) = b \). Let us denote this \( a \) by \( a_b \).

(b) Let \( g : B \to A \) be a map. Then, \( g \) is a left inverse of \( f \) if and only if we have

\[
(g(b) = a_b \text{ for each } b \in f(A)).
\]

[Proof of Claim 2: (a) Let \( b \in f(A) \). Thus, there exists at least one \( a \in A \) satisfying \( f(a) = b \). Moreover, there exists at most one such \( a \in A \) (since \( f \) is injective). Hence, there exists exactly one \( a \in A \) satisfying \( f(a) = b \). This proves Claim 2 (a).

(b) Claim 2 (b) is an “if and only if” statement. We shall prove it by first proving the \( \implies \) part (i.e., the “only if” part), and then proving the \( \iff \) part (i.e., the “if” part).

\( \implies \): Assume that \( g \) is a left inverse of \( f \). We must prove that \( (g(b) = a_b \text{ for each } b \in f(A)) \).

We know that \( g \) is a left inverse of \( f \). In other words, \( g \circ f = \text{id}_A \).

Now, let \( b \in f(A) \). Recall that \( a_b \) is the unique \( a \in A \) satisfying \( f(a) = b \) (by the definition of \( a_b \)). Thus, \( a_b \) is an element of \( A \) and satisfies \( f(a_b) = b \). Now, applying the map \( g \) to both sides of the equality \( b = f(a_b) \), we obtain

\[
g(b) = g(f(a_b)) = (g \circ f)(a_b) = \text{id}_A(a_b) = a_b.
\]

Now, forget that we fixed \( b \). We thus have proven that \( (g(b) = a_b \text{ for each } b \in f(A)) \). This proves the \( \implies \) part of Claim 2 (b).

\( \iff \): Assume that \( (g(b) = a_b \text{ for each } b \in f(A)) \). We must prove that \( g \) is a left inverse of \( f \).

Let \( x \in A \). Then, \( f(x) \in f(A) \). Thus, the definition of \( a_{f(x)} \) yields that \( a_{f(x)} \) is the unique \( a \in A \) satisfying \( f(a) = f(x) \). But this unique \( a \in A \) is obviously \( x \) (because \( x \) is an \( a \in A \) satisfying \( f(a) = f(x) \)). Hence, we conclude that \( a_{f(x)} = x \).

But we have assumed that \( (g(b) = a_b \text{ for each } b \in f(A)) \). Applying this to \( b = f(x) \), we obtain \( g(f(x)) = a_{f(x)} = x \). Hence, \( (g \circ f)(x) = g(f(x)) = x = \text{id}_A(x) \).

Forget that we fixed \( x \). We thus have shown that \( (g \circ f)(x) = \text{id}_A(x) \) for each \( x \in A \). In other words, \( g \circ f = \text{id}_A \). In other words, \( g \) is a left inverse of \( f \). This proves the \( \iff \) part of Claim 2 (b).

Thus, Claim 2 (b) is proven.]

Claim 3: Assume that the set \( A \) is nonempty. If \( f \) is injective, then \( f \) is left-invertible.
[Proof of Claim 3: Assume that $f$ is injective. Thus, for each $b \in f(A)$, there exists exactly one $a \in A$ satisfying $f(a) = b$ (by Claim 2 (a)). Let us denote this $a$ by $a_b$.]

Also, we have assumed that the set $A$ is nonempty. In other words, there exists some $w \in A$. Fix such a $w$.

Now, let us define a map $g : B \to A$ by setting

$$g(b) = \begin{cases} a_b & \text{if } b \in f(A) \\ w & \text{if } b \notin f(A) \end{cases} \quad \text{for each } b \in B.$$  

Thus, $g(b) = a_b$ for each $b \in f(A)$. Hence, Claim 2 (b) shows that $g$ is a left inverse of $f$. Hence, the map $f$ has a left inverse. In other words, $f$ is left-invertible. This proves Claim 3.]

(a) Assume that the set $A$ is nonempty. Then, $f$ is left-invertible if and only if $f$ is injective. (Indeed, the “if” part follows from Claim 3, whereas the “only if” part follows from Claim 1.) This solves part (a) of the exercise.

[Note that we have not used the assumption that the sets $A$ and $B$ are finite.]

(b) We have assumed that $f$ is injective. Thus, $|f(A)| = |A|$. But $f(A) \subseteq B$; hence, $|B \setminus f(A)| = |B| - |f(A)| = |B| - |A|$. 

Claim 2 (a) shows that for each $b \in f(A)$, there exists exactly one $a \in A$ satisfying $f(a) = b$. Let us denote this $a$ by $a_b$.

Claim 2 (b) shows that a map $g : B \to A$ is a left inverse of $f$ if and only if we have $(g(b) = a_b$ for each $b \in f(A)$. Thus, a left inverse of $f$ is the same as a map $g : B \to A$ with the property that

$$(g(b) = a_b \text{ for each } b \in f(A)). \quad (1)$$

Hence, in order to construct a left inverse $g$ of $f$, we can proceed as follows:

- For each $b \in f(A)$, set $g(b) = a_b$. (This is the only possible choice for $g(b)$, because our $g$ should satisfy (1).) Note that we are not making any choices at this step.

- For each $b \in B \setminus f(A)$, choose the value $g(b)$ arbitrarily (among all $|A|$ elements of $A$). Note that we have $|A|$ many choices for each $b \in B \setminus f(A)$.

Thus, there are $|A|^{|B \setminus f(A)|}$ many ways to perform this construction. Hence, the number of left inverses of $f$ is $|A|^{|B \setminus f(A)|}$. In other words, this number is $|A|^{|B| - |A|}$ (since $|B \setminus f(A)| = |B| - |A|$). This solves part (b) of the exercise.

---

2 EXERCISE 2

2.1 PROBLEM

Let $A$ and $B$ be two sets, and let $f : A \to B$ be a map. A right inverse of $f$ shall mean a map $h : B \to A$ such that $f \circ h = \text{id}_B$. We say that $f$ is right-invertible if and only if a right inverse of $f$ exists. (It is usually not unique.)

Assume that the sets $A$ and $B$ are finite.

Darij Grinberg, 00000000 3 dgrinber@umn.edu
(a) Prove that $f$ is right-invertible if and only if $f$ is surjective.\footnote{This holds even when $A$ and $B$ are infinite, if you assume the axiom of choice. But this is not the subject of our class.}

(b) Prove that the number of right inverses of $f$ is $\prod_{b \in B} |f^{-1}(b)|$. Here, $f^{-1}(b)$ denotes the set of all $a \in A$ satisfying $f(a) = b$.

2.2 Solution

If $b \in B$ is arbitrary, then $f^{-1}(b)$ shall denote the set of all $a \in A$ satisfying $f(a) = b$.

We first prove a few claims (that hold even without requiring $A$ and $B$ to be finite):

Claim 1: If $f$ is right-invertible, then $f$ is surjective.

[Proof of Claim 1: Assume that $f$ is right-invertible. In other words, $f$ has a right inverse. In other words, there exists a right inverse $h$ of $f$. Consider this $h$.

We know that $h$ is a right inverse of $f$. In other words, $h$ is a map from $B$ to $A$ such that $f \circ h = \text{id}_B$ (by the definition of a “right inverse”).

Let $b \in B$. Then, $f(h(b)) = (f \circ h)(b) = \text{id}_B(b) = b$. Thus, there exists some $a \in A$ such that $f(a) = b$ (namely, $a = h(b)$).

Now, forget that we fixed $b$. We thus have shown that if $b \in B$, then there exists some $a \in A$ such that $f(a) = b$. In other words, the map $f$ is surjective. This proves Claim 1.]

Claim 2: Let $h : B \to A$ be any map. Then, $h$ is a right inverse of $f$ if and only if we have

$$(h(b) \in f^{-1}(b) \text{ for each } b \in B).$$

[Proof of Claim 2: Claim 2 is an “if and only if” statement. We shall prove it by first proving the “$\implies$” part (i.e., the “only if” part), and then proving the “$\impliedby$” part (i.e., the “if” part).

$\implies$: Assume that $h$ is a right inverse of $f$. We must prove that $h(b) \in f^{-1}(b)$ for each $b \in B$.

We know that $h$ is a right inverse of $f$. In other words, $f \circ h = \text{id}_B$.

Now, let $b \in B$. Then, $f(h(b)) = (f \circ h)(b) = \text{id}_B(b) = b$. Thus, $h(b)$ belongs to the set of all $a \in A$ satisfying $f(a) = b$. In other words, $h(b)$ belongs to $f^{-1}(b)$ (since $f^{-1}(b)$ is the set of all $a \in A$ satisfying $f(a) = b$). In other words, $h(b) \in f^{-1}(b)$.

Now, forget that we fixed $b$. We thus have proven that $(h(b) \in f^{-1}(b) \text{ for each } b \in B)$. This proves the “$\implies$” part of Claim 2.

$\impliedby$: Assume that $(h(b) \in f^{-1}(b) \text{ for each } b \in B)$. We must prove that $h$ is a right inverse of $f$.

Let $y \in B$. Recall that $(h(b) \in f^{-1}(b) \text{ for each } b \in B)$. Applying this to $b = y$, we conclude that $h(y) \in f^{-1}(y)$. In other words, $h(y)$ is an $a \in A$ satisfying $f(a) = y$ (since $f^{-1}(y)$ is the set of all $a \in A$ satisfying $f(a) = y$). In other words, $f(h(y)) = y$. Hence, $(f \circ h)(y) = f(h(y)) = y = \text{id}_B(y)$.

Forget that we fixed $y$. We thus have shown that $(f \circ h)(y) = \text{id}_B(y)$ for each $y \in B$. In other words, $f \circ h = \text{id}_B$. In other words, $h$ is a right inverse of $f$. This proves the “$\impliedby$” part of Claim 2.

Thus, Claim 2 is proven.]
Claim 3: Assume that the set $B$ is finite. If $f$ is surjective, then $f$ is right-invertible.

[Proof of Claim 3: Assume that $f$ is surjective. We define a map $h : B \to A$ as follows: Let $b \in B$. Then, there exists some $a \in A$ satisfying $f(a) = b$ (since $f$ is surjective). Choose any such $a$, and set $h(b) = a$.

Note that we are making only finitely many choices in this definition of $h$ (since $B$ is finite); thus, this argument does not rely on the Axiom of Choice (indeed, finitely many choices can be made by induction).

So we have defined a map $h : B \to A$. This map $h$ has the property that

$$f(h(b)) = b \quad \text{for each } b \in B$$

(since $h(b)$ was defined to be some $a \in A$ satisfying $f(a) = b$). Thus, for each $b \in B$, we have $(f \circ h)(b) = f(h(b)) = b = \text{id}_B(b)$. In other words, $f \circ h = \text{id}_B$. In other words, $h$ is a right inverse of $f$ (by the definition of a “right inverse”). Hence, the map $f$ has a right inverse. In other words, $f$ is right-invertible. This proves Claim 3.]

(a) The map $f$ is right-invertible if and only if $f$ is surjective. (Indeed, the “if” part follows from Claim 3, whereas the “only if” part follows from Claim 1.) This solves part (a) of the exercise.

(b) Claim 2 shows that a map $h : B \to A$ is a right inverse of $f$ if and only if we have $(h(b) \in f^{-1}(b)$ for each $b \in B$). Thus, a right inverse of $f$ is the same as a map $h : B \to A$ with the property that

$$(h(b) \in f^{-1}(b) \quad \text{for each } b \in B).$$

Hence, in order to construct a right inverse $h$ of $f$, we can proceed as follows:

• For each $b \in B$, choose the value $h(b)$ to be one of the elements of the set $f^{-1}(b)$.

Note that we have $|f^{-1}(b)|$ many choices for each $b \in B$.

Thus, there are $\prod_{b\in B} |f^{-1}(b)|$ many ways to perform this construction. Hence, the number of right inverses of $f$ is $\prod_{b\in B} |f^{-1}(b)|$. This solves part (b) of the exercise.

---

3 EXERCISE 3

3.1 PROBLEM

(a) Prove that

$$\left(\frac{-1/2}{n}\right) = \left(\frac{-1}{4}\right)^n \binom{2n}{n} \quad \text{for each } n \in \mathbb{N}.\]

(b) Prove that

$$\sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k} = 4^n \quad \text{for each } n \in \mathbb{N}.\]

[Hint: Part (b) is highly difficult to prove combinatorially. Try using part (a) instead.]
3.2 Solution

Recall the classical formula which says that
\[
\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \text{for any } n \in \mathbb{N} \text{ and } k \in \mathbb{N} \text{ satisfying } n \geq k.
\] (2)

(a) Let \( n \in \mathbb{N} \). Then, \( 2n \geq n \) and \( 2n \in \mathbb{N} \). Hence, (2) (applied to \( 2n \) and \( n \) instead of \( n \) and \( k \)) yields
\[
\binom{2n}{n} = \frac{(2n)!}{n!(2n-n)!} = \frac{(2n)!}{n!n!} = \frac{1}{n!n!} \left( \prod_{i=1}^{2n} (2i) \right) = \frac{1}{n!n!} \left( \prod_{i=1}^{n} (2i+1) \right) \cdot \left( \prod_{i=1}^{n} (2i) \right)
\]
\[
= \frac{1}{n!n!} \left( \prod_{i=0}^{n-1} (2i+1) \right) \cdot \left( \prod_{i=1}^{n} i \right) = \frac{1}{n!n!} \left( \prod_{i=0}^{n-1} (2i+1) \right) \cdot 2^n n!
\]
\[
= \frac{2^n}{n!} \prod_{i=0}^{n-1} (2i+1).
\] (3)

Solving this equality for \( \prod_{i=0}^{n-1} (2i+1) \), we obtain
\[
\prod_{i=0}^{n-1} (2i+1) = \binom{2n}{n} \frac{2^n}{n!} = n! \binom{2n}{n}.
\] (4)

For each \( a \in \mathbb{Q} \), we have
\[
\binom{a}{n} = \frac{a(a-1) \cdots (a-n+1)}{n!} = \frac{\prod_{i=0}^{n-1} (a-i)}{n!} = \frac{1}{n!} \prod_{i=0}^{n-1} (a-i).
\]
Applying this to \( a = -1/2 \), we obtain
\[
\binom{-1/2}{n} = \frac{1}{n!} \prod_{i=0}^{n-1} (-1/2 - i) = \frac{1}{n!} \prod_{i=0}^{n-1} \frac{2i+1}{2i+2} = \frac{1}{n!} \left( \prod_{i=0}^{n-1} (2i+1) \right) \frac{2i+1}{2i+2}
\]
\[
= \frac{1}{n!} \left( \prod_{i=0}^{n-1} (2i+1) \right) \frac{n!}{2^n n!} = \frac{1}{n!} \frac{2^n}{2^n} \left( \binom{2n}{n} \right) = \frac{1}{n!} \left( \frac{2n}{2^n} \right) \left( \frac{2n}{n} \right)
\]
\[
= \left( \frac{-1}{4} \right)^n \binom{2n}{n}.
\]
This solves part (a) of the exercise.

(b) There are various proofs. Complicated combinatorial proofs can be found in:

- https://math.stackexchange.com/questions/72367
- https://math.stackexchange.com/a/360780/

(I have not read them all myself.)

It is much easier to solve the exercise algebraically, using part (a).

Let \( n \in \mathbb{N} \). First, we observe that \((-1/n) = (-1)^n\). (This can be derived from the Upper Negation identity, or easily checked directly using the definition of \((-1/n)\).)

Now, recall that the Vandermonde convolution theorem (Theorem 2.18 in class work (2018-09-24)) says that

\[
\binom{x+y}{n} = \sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}
\]

for all \( x \in \mathbb{R} \) and \( y \in \mathbb{R} \).

Applying this to \( x = -1/2 \) and \( y = -1/2 \), we obtain

\[
\binom{-1/2 + (-1/2)}{n} = \sum_{k=0}^{n} \binom{-1/2}{k} \binom{-1/2}{n-k}.
\]

Comparing this with \( \binom{-1/2 + (-1/2)}{n} = \binom{-1}{n} = (-1)^n \),

we obtain

\[
(-1)^n = \sum_{k=0}^{n} \binom{-1/2}{k} \binom{-1/2}{n-k} = \left(\frac{-1}{4}\right)^k \binom{2k}{k} \left(\frac{-1}{4}\right)^{n-k} \binom{2(n-k)}{n-k} = \left(\frac{-1}{4}\right)^{n-k} \binom{2(n-k)}{n-k}.
\]

(by part (a) of the exercise, applied to \( k \) instead of \( n \))

\[
= \sum_{k=0}^{n} \left(\frac{-1}{4}\right)^k \left(\frac{-1}{4}\right)^{n-k} \binom{2k}{k} \binom{2(n-k)}{n-k} = \left(\frac{-1}{4}\right)^n \sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k}.
\]

(by part (a) of the exercise, applied to \( n-k \) instead of \( n \))

Multiplying both sides of this equality by \((-4)^n\), we obtain

\[
(-4)^n (-1)^n = (-4)^n \left(\frac{-1}{4}\right)^n \sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k} = \sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k}.
\]
Hence,
\[ \sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k} = (-4)^n (-1)^n = 4^n. \]

This solves part (b) of the exercise.

3.3 REMARK

Here is another identity, similar to part (b) of the exercise:
\[ \sum_{k=0}^{n} (-1)^k \binom{2k}{k} \binom{2(n-k)}{n-k} = \begin{cases} 2^n \binom{n}{n/2}, & \text{if } n \text{ is even;} \\ 0, & \text{if } n \text{ is odd} \end{cases} \quad \text{for each } n \in \mathbb{N}. \]

Can you prove this one?

4 EXERCISE 4

4.1 PROBLEM

Recall once again the Fibonacci sequence \((f_0, f_1, f_2, \ldots)\), which is defined recursively by
\[ f_0 = 0, \quad f_1 = 1, \quad f_n = f_{n-1} + f_{n-2} \quad \text{for all } n \geq 2. \] (5)

It is easy to see that \(f_1, f_2, f_3, \ldots\) are positive integers (which will allow us to divide by them soon).

For any \(n \in \mathbb{N}\) and \(k \in \mathbb{Z}\), define the rational number \(\binom{n}{k}_F\) (a slight variation on the corresponding binomial coefficient) by
\[ \binom{n}{k}_F = \begin{cases} \frac{f_n f_{n-1} \cdots f_{k+1}}{f_k f_{k-1} \cdots f_1}, & \text{if } n \geq k \geq 0; \\ 0, & \text{otherwise}. \end{cases} \]

(a) Let \(n\) be a positive integer, and let \(k \in \mathbb{N}\) be such that \(n \geq k\). Prove that
\[ \binom{n}{k}_F = f_{k+1} \binom{n-1}{k}_F + f_{n-k-1} \binom{n-1}{k-1}_F, \]
where we set \(f_{-1} = 1\).

(b) Prove that \(\binom{n}{k}_F \in \mathbb{N}\) for any \(n \in \mathbb{N}\) and \(k \in \mathbb{N}\).

4.2 SOLUTION

Set \(f_{-1} = 1\). Let us first show the following claims:

Claim 1: Let \(m \in \mathbb{N}\) and \(n \in \{-1, 0, 1, 2, \ldots\}\). Then, \(f_{m+n+1} = f_m f_n + f_{m+1} f_{n+1}\).
Proof of Claim 1: In the case when \( n \in \mathbb{N} \), this follows immediately from Theorem 1.37 in the class notes (2018-09-17). Thus, we WLOG assume that \( n \notin \mathbb{N} \). Combining \( n \in \{-1, 0, 1, 2, \ldots\} \) with \( n \notin \mathbb{N} \), we obtain \( n \in \{-1, 0, 1, 2, \ldots\} \setminus \mathbb{N} = \{-1\} \). Hence, \( n = -1 \). Thus,

\[
f_m n + f_{m+1} f_{n+1} = f_m \left( \sum_{i=1}^{n} f_{i-1} \right) + f_{m+1} f_{n+1} = f_m.
\]

From \( n = -1 \), we also obtain \( f_{m+n+1} = f_{m+(-1)+1} = f_m \). Comparing the last two equalities, we obtain \( f_{m+n+1} = f_m \). This proves Claim 1.

Claim 2: We have \( \binom{n}{0} = 1 \) for each \( n \in \mathbb{N} \).

Proof of Claim 2: Let \( n \in \mathbb{N} \). The definition of \( \binom{n}{0} \) yields

\[
\binom{n}{0} = \begin{cases} 
\frac{f_n f_{n-1} \cdots f_{n-0+1}}{f_0 f_{-1} \cdots f_1}, & \text{if } n \geq 0 \geq 0; \\
0, & \text{otherwise}
\end{cases}
\]

\[
= \begin{cases} 
\text{(empty product)}, & \text{if } n \geq 0 \geq 0; \\
\text{(empty product)}, & \text{otherwise}
\end{cases}
\]

This proves Claim 2.

Claim 3: We have \( \binom{n}{n} = 1 \) for each \( n \in \mathbb{N} \).

Proof of Claim 3: Let \( n \in \mathbb{N} \). The definition of \( \binom{n}{n} \) yields

\[
\binom{n}{n} = \begin{cases} 
\frac{f_n f_{n-1} \cdots f_{n-n+1}}{f_n f_{n-1} \cdots f_1}, & \text{if } n \geq n \geq 0; \\
0, & \text{otherwise}
\end{cases}
\]

\[
= \frac{f_n f_{n-1} \cdots f_1}{f_n f_{n-1} \cdots f_1} = 1.
\]

This proves Claim 3.

Claim 4: We have \( \binom{n}{-1} = 0 \) for each \( n \in \mathbb{N} \).

Proof of Claim 4: Let \( n \in \mathbb{N} \). The definition of \( \binom{n}{-1} \) yields

\[
\binom{n}{-1} = \begin{cases} 
\frac{f_n f_{n-1} \cdots f_{n-(-1)+1}}{f_{-1} f_{-1-1} \cdots f_1}, & \text{if } n \geq -1 \geq 0; \\
0, & \text{otherwise}
\end{cases}
\]

(since we don’t have \( n \geq -1 \geq 0 \) (because we don’t have \(-1 \geq 0\))). This proves Claim 4.

Claim 5: We have \( \binom{0}{k} = [k = 0] \) for each \( k \in \mathbb{N} \) (where we are using the Iverson bracket notation).
Proof of Claim 5: Let \( k \in \mathbb{N} \). We must prove that \( \binom{0}{k}_F = [k = 0] \).

Claim 2 yields \( \binom{0}{0}_F = 1 \). Comparing this with \( [0 = 0] = 1 \), we obtain \( \binom{0}{0}_F = [0 = 0] \).

In other words, Claim 5 holds for \( k = 0 \). Hence, for the rest of this proof, we WLOG assume that \( k \neq 0 \). Thus, \( k > 0 \) (since \( k \in \mathbb{N} \)). Hence, we don’t have \( 0 \geq k \geq 0 \).

Now, the definition of \( \binom{0}{k}_F \) yields

\[
\binom{0}{k}_F = \begin{cases}
\frac{f_0 f_{0-1} \cdots f_{0-k+1}}{f_k f_{k-1} \cdots f_1}, & \text{if } 0 \geq k \geq 0; \\
0, & \text{otherwise}
\end{cases}
\]

(since we don’t have \( 0 \geq k \geq 0 \)). Comparing this with \( [k = 0] = 0 \) (since \( k \neq 0 \)), we obtain \( \binom{0}{k}_F = [k = 0] \). This proves Claim 5.

(a) Note that \( n - 1 \in \mathbb{N} \) (since \( n \) is a positive integer). We are in one of the following three cases:

Case 1: We have \( k = 0 \).

Case 2: We have \( k = n \).

Case 3: We have neither \( k = 0 \) nor \( k = n \).

Let us first consider Case 1. In this case, we have \( k = 0 \). Hence,

\[
f_{k+1} \binom{n-1}{k}_F + f_{n-k-1} \binom{n-1}{k-1}_F = \frac{f_0}{f_1} \binom{n-1}{0}_F + f_{n-0-1} \binom{n-1}{0-1}_F
\]

(by Claim 2, applied to \( n-1 \) instead of \( n \))

\[
= 1 \cdot 1 + f_{n-0-1} \cdot 0 = 1.
\]

Comparing this with

\[
\binom{n}{k}_F = \binom{n}{0}_F \quad \text{(since } k = 0 \text{)}
\]

\[
= 1 \quad \text{(by Claim 2)}
\]

we obtain \( \binom{n}{k}_F = f_{k+1} \binom{n-1}{k}_F + f_{n-k-1} \binom{n-1}{k-1}_F \). Thus, part (a) of the exercise is solved in Case 1.

Let us next consider Case 2. In this case, we have \( k = n \). The definition of \( \binom{n-1}{n}_F \) yields

\[
\binom{n-1}{n}_F = \begin{cases}
f_{n-1} f_{(n-1)-1} \cdots f_{(n-1)-n+1}, & \text{if } n-1 \geq n \geq 0; \\
0, & \text{otherwise}
\end{cases}
\]

Darij Grinberg, 00000000 dgrinber@umn.edu
(since we don’t have $n - 1 \geq n \geq 0$ (since we don’t have $n - 1 \geq n$)). Now, from $k = n$, we obtain

$$f_{k+1} \left( \frac{n - 1}{k} \right)_F + f_{n-k-1} \left( \frac{n - 1}{k-1} \right)_F = f_{n+1} \left( \frac{n - 1}{n} \right)_F + f_{n-n-1} \left( \frac{n - 1}{n-1} \right)_F$$

(by Claim 1, applied to $n-1$ instead of $n$)

$$= f_{n+1} \cdot 0 + 1 \cdot 1 = 1. \quad \text{(since $k = n$)}$$

Comparing this with

$$\binom{n}{k}_F = \binom{n}{n}_F = 1 \quad \text{(by Claim 3)},$$

we obtain

$$\binom{n}{k}_F = f_{k+1} \left( \frac{n - 1}{k} \right)_F + f_{n-k-1} \left( \frac{n - 1}{k-1} \right)_F. \quad \text{Thus, part (a) of the exercise is solved in Case 2.}$$

Let us finally consider Case 3. In this case, we have neither $k = 0$ nor $k = n$. Hence, $k \neq 0$ and $k \neq n$. Combining $k \neq 0$ with $k \geq 0$, we obtain $k > 0$. Combining $k \neq n$ with $k \leq n$ (which follows from $n \geq k$), we obtain $k < n$. Combining $k > 0$ with $k < n$, we obtain $k \in \{1, 2, \ldots, n-1\}$ (since $k$ and $n$ are integers). Thus, $1 \leq k \leq n-1$, so that $n - 1 \geq k$ and $k \geq 1$. Thus, $n - 1 \geq k \geq k - 1 \geq 0$ (since $k \geq 1$). The definition of $\binom{n - 1}{k}_F$ now yields

$$\binom{n - 1}{k}_F = \begin{cases} \frac{f_{n-1}f_{n-1} \cdots f_{n-1-k+1}}{f_k f_{k-1} \cdots f_1}, & \text{if } n - 1 \geq k \geq 0; \\ 0, & \text{otherwise} \end{cases}$$

$$= \frac{f_{n-1}f_{n-1} \cdots f_{n-1-k+1}}{f_k f_{k-1} \cdots f_1} \quad \text{(since } n - 1 \geq k \geq 0 \text{)}$$

$$= \frac{f_{n-1}f_{n-2} \cdots f_{n-k}}{f_k f_{k-1} \cdots f_1} = \frac{(f_{n-1}f_{n-2} \cdots f_{n-k+1}) \cdot f_{n-k}}{f_k f_{k-1} \cdots f_1}$$

(here, we have split off the factor $f_{n-k}$ from the product in the numerator). Also, the definition of $\binom{n - 1}{k-1}_F$ yields

$$\binom{n - 1}{k-1}_F = \begin{cases} \frac{f_{n-1}f_{n-1} \cdots f_{n-1-(k-1)+1}}{f_k f_{k-1} \cdots f_1}, & \text{if } n - 1 \geq k - 1 \geq 0; \\ 0, & \text{otherwise} \end{cases}$$

$$= \frac{f_{n-1}f_{n-1} \cdots f_{n-1-(k-1)+1}}{f_k f_{k-1} \cdots f_1} \quad \text{(since } n - 1 \geq k - 1 \geq 0 \text{)}$$

$$= \frac{f_{n-1}f_{n-2} \cdots f_{n-k+1}}{f_k f_{k-2} \cdots f_1} = f_k \cdot \frac{f_{n-1}f_{n-2} \cdots f_{n-k+1}}{f_k (f_{k-1} f_{k-2} \cdots f_1)}$$

$$= f_k \cdot \frac{f_{n-1}f_{n-2} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1} \quad \text{(since } f_k \cdot (f_{k-1} f_{k-2} \cdots f_1) = f_k f_{k-1} \cdots f_1 \text{).}$$
Hence,
\[
f_{k+1} \left( \binom{n-1}{k} \right)_F + f_{n-k-1} \left( \binom{n-1}{k-1} \right)_F = \frac{(f_{n-1}f_{n-2} \cdots f_{n-k+1}) \cdot f_{n-k}}{f_k f_{k-1} \cdots f_1} + \frac{f_{n-1}f_{n-2} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1} = f_{k+1} \left( \binom{n-1}{k} \right)_F + f_{n-k-1} \left( \binom{n-1}{k-1} \right)_F.
\]

We have \(k \in \mathbb{N}\) and \(n - k - 1 \in \{-1, 0, 1, 2, \ldots\}\) (since \(n - k - 1 \geq n - 1 = -1\)). Hence, Claim 1 (applied to \(k\) and \(n - k - 1\) instead of \(m\) and \(n\)) yields
\[
f_{k+(n-k-1)+1} = \frac{f_{k} f_{n-k} + f_{k+1} f_{(n-k-1)+1}}{f_{k} f_{k-1} \cdots f_1} = f_{n-1} f_{k} + f_{k+1} f_{n-k} = f_{k+1} f_{n-k} + f_{n-k-1} f_{k}.
\]
Thus,
\[
f_{k+1} f_{n-k} + f_{n-k-1} f_{k} = f_{k+(n-k-1)+1} = f_n
\]
(since \(k + (n - k - 1) + 1 = n\)). Hence, (6) becomes
\[
f_{k+1} \left( \binom{n-1}{k} \right)_F + f_{n-k-1} \left( \binom{n-1}{k-1} \right)_F = \frac{(f_{n-1}f_{n-2} \cdots f_{n-k+1}) \cdot f_{n-k}}{f_k f_{k-1} \cdots f_1} + \frac{f_{n-1}f_{n-2} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1} = f_n \cdot \frac{f_{n-1}f_{n-2} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1} = f_{n-1} f_{n-2} \cdots f_{n-k+1}
\]
(since \(f_n \cdot (f_{n-1}f_{n-2} \cdots f_{n-k+1}) = f_{n-1}f_{n-2} \cdots f_{n-k+1}\)). Comparing this with
\[
\binom{n}{k}_F = \begin{cases} \frac{f_n f_{n-1} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1}, & \text{if } n \geq k \geq 0; \\ 0, & \text{otherwise} \end{cases} \quad \text{(by the definition of } \binom{n}{k}_F)\]
\[
= \frac{f_n f_{n-1} \cdots f_{n-k+1}}{f_k f_{k-1} \cdots f_1} \quad \text{(since } n \geq k \geq 0),
\]
we obtain
\[
\binom{n}{k}_F = f_{k+1} \left( \binom{n-1}{k} \right)_F + f_{n-k-1} \left( \binom{n-1}{k-1} \right)_F.
\]
Thus, part (a) of the exercise is solved in Case 3.

We have now proven part (a) of the exercise in all three Cases 1, 2 and 3. Thus, part (a) of the exercise always holds.

(b) We shall prove part (b) of the exercise by induction on \(n\):
Induction base: For each \( k \in \mathbb{N} \), we have
\[
\binom{0}{k}_F = [k = 0] \quad \text{(by Claim 5)}
\]
\( \in \mathbb{N} \).

In other words, part (b) of the exercise holds for \( n = 0 \). This completes the induction base.

Induction step: Let \( m \) be a positive integer. Assume (as the induction hypothesis) that part (b) of the exercise holds for \( n = m - 1 \). We must prove that part (b) of the exercise holds for \( n = m \).

We have assumed that part (b) of the exercise holds for \( n = m - 1 \). In other words, we have
\[
\binom{m-1}{k}_F \in \mathbb{N} \quad \text{for any } k \in \mathbb{N}.
\]
\[(7)\]

Now, let \( k \in \mathbb{N} \) be arbitrary. We shall show that \( \binom{m}{k}_F \in \mathbb{N} \).

We are in one of the following three cases:

Case 1: We have \( k = 0 \).
Case 2: We have \( k > m \).
Case 3: We have neither \( k = 0 \) nor \( k > m \).

Let us first consider Case 1. In this case, we have \( k = 0 \). Hence, \( \binom{m}{0}_F = \binom{m}{0}_F = 1 \) (by Claim 2). Thus, \( \binom{m}{k}_F = 1 \in \mathbb{N} \). Hence, \( \binom{m}{k}_F \in \mathbb{N} \) is proven in Case 1.

Let us next consider Case 2. In this case, we have \( k > m \). Hence, we don’t have \( m \geq k \).

Thus, we don’t have \( m \geq k \geq 0 \). Now, the definition of \( \binom{m}{k}_F \) yields
\[
\binom{m}{k}_F = \begin{cases} 
\frac{f_m f_{m-1} \cdots f_{m-k+1}}{f_k f_{k-1} \cdots f_1}, & \text{if } m \geq k \geq 0; \\
0, & \text{otherwise}
\end{cases}
\]
\[
= 0 \quad (\text{since we don’t have } m \geq k \geq 0)
\]
\( \in \mathbb{N} \).

Hence, \( \binom{m}{k}_F \in \mathbb{N} \) is proven in Case 2.

Let us finally consider Case 3. In this case, we have neither \( k = 0 \) nor \( k > m \). Hence, we have \( k \neq 0 \) and \( k \leq m \). From \( k \neq 0 \), we obtain \( k \geq 1 \) (since \( k \in \mathbb{N} \)), so that \( k-1 \in \mathbb{N} \). Hence, \[(7)\] (applied to \( k-1 \) instead of \( k \)) yields \( \binom{m-1}{k-1}_F \in \mathbb{N} \). Also, \[(7)\] yields \( \binom{m-1}{k}_F \in \mathbb{N} \).

Furthermore, the Fibonacci sequence \( (f_0, f_1, f_2, \ldots) \) is a sequence of nonnegative integers; thus, \( f_i \in \mathbb{N} \) for each \( i \in \mathbb{N} \). Since this holds for \( i = -1 \) as well (because \( f_{-1} = 1 \in \mathbb{N} \)), we can thus conclude that
\[
f_i \in \mathbb{N} \quad \text{for each } i \in \{-1, 0, 1, 2, \ldots\}.
\]
\[(8)\]

Now, \( m - k \leq m - m - 1 = -1 \), so that \( m - k - 1 \in \{-1, 0, 1, 2, \ldots\} \). Hence, \[(8)\] (applied to \( i = m - k - 1 \)) yields \( f_{m-k-1} \in \mathbb{N} \). Also, \( k + 1 \in \mathbb{N} \subseteq \{-1, 0, 1, 2, \ldots\} \). Hence,
[8] (applied to $i = k+1$) yields $f_{k+1} \in \mathbb{N}$. Now, part (a) of the exercise (applied to $n = m$) yields
\[
\binom{m}{k}_f = f_{k+1} \binom{m-1}{k}_f + f_{m-k-1} \binom{m-1}{k-1}_f \in \mathbb{N}.
\]

Hence, $\binom{m}{k}_f \in \mathbb{N}$ is proven in Case 3.

We have now proven $\binom{m}{k}_f \in \mathbb{N}$ in each of the three Cases 1, 2 and 3. Hence, $\binom{m}{k}_f \in \mathbb{N}$ always holds.

Now, forget that we fixed $k$. We thus have shown that $\binom{m}{k}_f \in \mathbb{N}$ for any $k \in \mathbb{N}$. In other words, part (b) of the exercise holds for $n = m$. This completes the induction step. Hence, part (b) of the exercise is solved by induction.

### 4.3 Remark

The numbers $\binom{n}{k}_f$ defined in this exercise are the so-called *Fibonomial coefficients*. As the name (and this exercise) suggests, they have lots of properties in common with the binomial coefficients; there are numerous papers devoted to proving some of these properties. See, for example:


You can find more by searching for “Fibonomial coefficients” on Google Scholar.

---

### 5 Exercise 5

#### 5.1 Problem

Let $j \in \mathbb{N}$, $r \in \mathbb{R}$ and $s \in \mathbb{R}$. Prove that
\[
\sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r - sk}{j} = s^j.
\]

**[Hint]**: First, argue that it suffices to prove this only for $s \in \mathbb{N}$ and $r \in \mathbb{Z}$ satisfying $r \geq sj$. Next, consider $r$ distinct stones, $sj$ of which are arranged in $j$ piles containing $s$ stones each, while the remaining $r - sj$ stones are forming a separate heap. How many ways are there to pick $j$ of these $r$ stones such that each of the $j$ piles loses at least one stone?
5.2 Solution

5.2.1 First solution (sketched)

Here is a solution following the hint. (I have learnt it from Peter Scholze in the 2000s; it is the proof of Lemma 1 in [http://artofproblemsolving.com/community/c6h41800p287507].)

We forget that we fixed \( r \) and \( s \).

Let us first recall the principle of inclusion and exclusion:

**Theorem 5.1.** Let \( n \in \mathbb{N} \). Let \( A_1, A_2, \ldots, A_n \) be finite sets.

(a) We have

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

(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 \( \bigcap_{i \in \emptyset} A_i \) is understood to mean the set \( S \).

Next, we recall the “polynomial identity trick” in the following form:

**Lemma 5.2.** If a polynomial \( P \) with real coefficients has infinitely many roots, then \( P \) is the zero polynomial.

Let us now solve the exercise under some restrictive requirements on \( r \) and \( s \):

**Claim 1:** Let \( r \in \mathbb{N} \) and \( s \in \mathbb{N} \) be such that \( r \geq sj \). Then,

\[
\sum_{k=0}^{j} (-1)^{k} \binom{j}{k} \binom{r-sk}{j} = sj.
\]

**Proof of Claim 1:** Consider \( r \) (distinguishable) stones \( s_1, s_2, \ldots, s_r \). Assume that \( sj \) of these stones are arranged in \( j \) disjoint piles \( P_1, P_2, \ldots, P_j \), with each pile \( P_k \) containing exactly \( s \) stones. The remaining \( r-sj \) stones are not contained in any pile; let’s say they form the rest-heap.

(Formally speaking, this means that \( P_1, P_2, \ldots, P_j \) are \( j \) disjoint \( s \)-element subsets of \( \{s_1, s_2, \ldots, s_r\} \); the rest-heap is then defined to be \( \{s_1, s_2, \ldots, s_r\} \setminus (P_1 \cup P_2 \cup \cdots \cup P_j) \). Of course, such an arrangement of stones and piles is only possible because we have \( r \in \mathbb{N} \) and \( s \in \mathbb{N} \) and \( r \geq sj \).)

A \( j \)-pick will mean a way to choose \( j \) of the \( r \) stones (i.e., a \( j \)-element subset of \( \{s_1, s_2, \ldots, s_r\} \)).

If \( P \) is one of the \( j \) piles (that is, \( P \in \{P_1, P_2, \ldots, P_j\} \)), and \( J \) is a \( j \)-pick, then we say that \( J \) avoids \( P \) if and only if \( J \) contains no stone from \( P \) (that is, \( J \cap P = \emptyset \)).

A \( j \)-pick is said to be legal if it avoids none of the \( j \) piles. Let \( N \) be the number of all legal \( j \)-picks. We shall compute \( N \) in two different ways:

- If \( J \) is a legal \( j \)-pick, then \( J \) must contain at least 1 stone from \( P_1 \), at least 1 stone from \( P_2 \), and so on (because it must avoid none of the \( j \) piles). But since it only contains \( j \) stones altogether (because it is a \( j \)-pick), we thus conclude that it must
contain **exactly** 1 stone from $P_1$, **exactly** 1 stone from $P_2$, and so on (and no stones from the rest-heap). Hence, in order to choose a legal $j$-pick, it suffices to decide **which** of the $s$ stones from $P_1$ it should contain, **which** of the $s$ stones from $P_2$ it should contain, and so on. This is a total of $j$ decisions (one for each pile $P_k$), and each decision allows for $s$ choices. Therefore, the total number of choices is $s^j$. Hence, the number $N$ of all legal $j$-picks is $s^j$. In other words,

$$N = s^j. \quad (9)$$

- Our second computation of $N$ relies on Theorem 5.1(b).

Indeed, let $S$ be the set of all $j$-picks. For each $i \in [j]$, we let $A_i$ be the set of all $j$-picks that avoid the pile $P_i$. Clearly, each of $A_1, A_2, \ldots, A_j$ is a subset of $S$. Moreover, if $I$ is any subset of $[j]$, then

$$\left| \bigcap_{i \in I} A_i \right| = \binom{r - s |I|}{j}. \quad (10)$$

**Proof of (10):** Let $I$ be any subset of $[j]$. The piles $P_i$ for $i \in I$ are $|I|$ many disjoint $s$-element sets; thus, their union $\bigcup_{i \in I} P_i$ is an $s|I|$-element set. In other words,

$$\left| \bigcup_{i \in I} P_i \right| = s |I|.$$ But $\bigcup_{i \in I} P_i$ is a subset of $\{s_1, s_2, \ldots, s_r\}$; thus,

$$\left| \{s_1, s_2, \ldots, s_r\} \setminus \bigcup_{i \in I} P_i \right| = \left| \{s_1, s_2, \ldots, s_r\} \right| - \left| \bigcup_{i \in I} P_i \right| = r - s |I|.$$

For each $i \in I$, we have

$$A_i = ( \text{the set of all } j \text{-picks that avoid the pile } P_i )$$

(by the definition of $A_i$)

$$= \{ J \in S \mid J \text{ avoids the pile } P_i \} \quad (11)$$

(since the set of all $j$-picks is $S$). Thus,

$$\bigcap_{i \in I} A_i = \bigcap_{i \in I} \{ J \in S \mid J \text{ avoids the pile } P_i \}$$

$$= \{ J \in S \mid J \text{ avoids the pile } P_i \text{ for each } i \in I \}$$

$$= \{ J \in S \mid J \cap P_i = \emptyset \text{ for each } i \in I \}$$

$$= \left\{ J \in S \mid J \cap \left( \bigcup_{i \in I} P_i \right) = \emptyset \right\}$$

$$= \left\{ J \in S \mid J \subseteq \{s_1, s_2, \ldots, s_r\} \setminus \bigcup_{i \in I} P_i \right\}$$

$$= \left\{ j \text{-element subsets of } \{s_1, s_2, \ldots, s_r\} \setminus \bigcup_{i \in I} P_i \right\},$$

Indeed, if the legal $j$-pick $J$ contained more than 1 stone from any single pile $P_k$, or any stone from the rest-heap, then it would contain more than $j$ stones altogether, which would contradict the fact that it only contains $j$ stones.

---

Darij Grinberg, 00000000
dgrinber@umn.edu
(since “$J \in S$” simply means that $J$ is a $j$-pick, i.e., a $j$-element subset of $\{s_1, s_2, \ldots, s_r\}$). Hence,

$$\left| \bigcap_{i \in I} A_i \right| = \left| \left\{ j\text{-element subsets of } \{s_1, s_2, \ldots, s_r\} \setminus \bigcup_{i \in I} P_i \right\} \right|$$

$$= \left( \left| \{s_1, s_2, \ldots, s_r\} \setminus \bigcup_{i \in I} P_i \right| \right) = \binom{r - s |I|}{j}$$

(since $\left| \{s_1, s_2, \ldots, s_r\} \setminus \bigcup_{i \in I} P_i \right| = r - s |I|$). This proves (10).

Now, Theorem 5.1 (b) (applied to $n = j$) yields

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

$$= \sum_{k=0}^j \sum_{|I|=k} (-1)^{|I|} \binom{r - s |I|}{j}$$

(by (10))

$$= \sum_{k=0}^j \left( \text{the number of all } I \subseteq [j] \text{ such that } |I|=k \right) \binom{r - sk}{j}$$

(since $|I|=k$)

$$= \sum_{k=0}^j \left( -1 \right)^k \binom{j}{k} \binom{r - sk}{j}$$

(12)

But for each $i \in [j]$, we have defined $A_i$ to be the set of all $j$-picks that avoid the pile $P_i$. Thus,

$$S \setminus \bigcup_{i=1}^j A_i = \text{(the set of all } j\text{-picks that avoid none of the } j\text{ piles } P_1, P_2, \ldots, P_j)$$

$$= \text{(the set of all legal } j\text{-picks})$$

(because of how we defined “legal”). Hence,

$$\left| S \setminus \bigcup_{i=1}^j A_i \right| = \text{(the number of all legal } j\text{-picks}) = N$$
(by the definition of $N$). Comparing this with (12), we obtain

$$N = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{j}. \quad (13)$$

Comparing (9) with (13), we find

$$\sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{j} = s^j.$$  

This proves Claim 1.

Next, we shall use the polynomial identity trick to extend Claim 1 to somewhat greater generality (allowing $r$ to roam freely across $\mathbb{R}$, while $s$ still has to belong to $\mathbb{N}$):

**Claim 2:** Let $r \in \mathbb{R}$ and $s \in \mathbb{N}$. Then,

$$\sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{j} = s^j.$$  

**Proof of Claim 2:** Forget that we fixed $r$. Let $P$ be the polynomial in the indeterminate $x$ (with real coefficients) defined by

$$P = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{x-sk}{j} - s^j. \quad (14)$$

Then, for each $r \in \mathbb{N}$ satisfying $r \geq sj$, we have

$$P(r) = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{j} - s^j = s^j - s^j = 0.$$  

In other words, each $r \in \mathbb{N}$ satisfying $r \geq sj$ is a root of the polynomial $P$. Hence, the polynomial $P$ has infinitely many roots (since there are infinitely many such $r$). Thus, Lemma 5.2 shows that $P$ is the zero polynomial. In other words, $P = 0$.

Now, let $r \in \mathbb{R}$. From $P = 0$, we obtain $P(r) = 0$. Thus,

$$0 = P(r) = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{j} - s^j$$  

(by (14)). In other words,

$$\sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{j} = s^j.$$  

This proves Claim 2.

Applying the polynomial identity trick one more time, we can achieve the full generality required in the exercise:
Claim 3: Let \( r \in \mathbb{R} \) and \( s \in \mathbb{R} \). Then,

\[
\sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r - sk}{j} = s^j.
\]

[Proof of Claim 3:] Forget that we fixed \( s \). Let \( P \) be the polynomial in the indeterminate \( x \) (with real coefficients) defined by

\[
P = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r - xk}{j} - x^j. \tag{15}
\]

Then, for each \( s \in \mathbb{N} \), we have

\[
P(s) = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r - sk}{j} - s^j = s^j - s^j = 0.
\]

(by Claim 2)

In other words, each \( s \in \mathbb{N} \) is a root of the polynomial \( P \). Hence, the polynomial \( P \) has infinitely many roots. Thus, Lemma \[\ref{lemma:zero_polynomial}\] shows that \( P \) is the zero polynomial. In other words, \( P = 0 \).

Now, let \( s \in \mathbb{R} \). From \( P = 0 \), we obtain \( P(s) = 0 \). Thus,

\[
0 = P(s) = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r - sk}{j} - s^j
\]

(by \eqref{eq:polynomial}). In other words,

\[
\sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r - sk}{j} = s^j.
\]

This proves Claim 3.]

But Claim 3 is precisely the claim of the exercise. Thus, the exercise is solved.

5.2.2 Second solution (sketched)

We can also solve the exercise in a purely algebraic way, without having to rely on the polynomial identity trick and coming up with counting problems.

Let us forget that we fixed \( j, r \) and \( s \).

We shall use the Iverson bracket notation. We first recall two fundamental facts about binomial coefficients:

- Every \( n \in \mathbb{N} \) satisfies

\[
\sum_{k=0}^{n} (-1)^k \binom{n}{k} = [n = 0]. \tag{16}
\]

(This was proven in class.)
• Every \( n \in \mathbb{R} \) and \( k \in \{1, 2, 3, \ldots \} \) satisfy

\[
k \binom{n}{k} = n \binom{n-1}{k-1}.
\]  

(This is the so-called absorption identity, and can easily be checked by hand.\(^4\) Note that it also holds for all \( k \in \mathbb{R} \), but we will only need it for \( k \in \{1, 2, 3, \ldots \} \).)

We now make a definition:

**Definition 5.3.** Let \( j \in \mathbb{N} \), \( i \in \mathbb{N} \), \( r \in \mathbb{R} \) and \( s \in \mathbb{R} \). Then, we define a number \( L_{r,s,j,i} \in \mathbb{R} \) by

\[
L_{r,s,j,i} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r-sk}{i}.
\]  

Notice that this definition is more general than the sum appearing in the exercise; indeed, the latter sum is \( L_{r,s,j,j} \). It turns out that this extra generality allows a recursive approach that wouldn’t be possible if we would only be considering the \( L_{r,s,j,j} \)’s in isolation.

The main thrust of our recursive approach runs through the following lemma:

**Lemma 5.4.** Let \( j \in \mathbb{N} \), \( r \in \mathbb{R} \) and \( s \in \mathbb{R} \).

(a) We have \( L_{r,s,j,0} = [j = 0] \).

(b) Let \( i \) be a positive integer. Then,

\[
iL_{r,s,j,i} = rL_{r-1,s,j,i-1} + sjL_{r-1-s,s,j-1,i-1}.
\]

**Proof of Lemma 5.4.** (a) The definition of \( L_{r,s,j,0} \) yields

\[
L_{r,s,j,0} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \binom{r}{i} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} = [j = 0] \quad (\text{by } (16), \text{ applied to } n = j).
\]

\(^4\) **Proof of (17).** Let \( n \in \mathbb{R} \) and \( k \in \{1, 2, 3, \ldots \} \). Thus, \( k-1 \in \mathbb{N} \) (since \( k \in \{1, 2, 3, \ldots \} \)); hence, the definition of \( \binom{n-1}{k-1} \) yields

\[
\binom{n-1}{k-1} = \frac{(n-1)((n-1) - 1)\cdots((n-1) - (k-1) + 1)}{(k-1)!} = \frac{(n-1)(n-2)\cdots(n-k+1)}{(k-1)!}.
\]

Multiplying both sides of this equality by \( n \), we obtain

\[
n \binom{n-1}{k-1} = n \cdot \frac{(n-1)(n-2)\cdots(n-k+1)}{(k-1)!} = \frac{n \cdot ((n-1)(n-2)\cdots(n-k+1))}{(k-1)!} = \frac{n(n-1)\cdots(n-k+1)}{(k-1)!} \quad (18)
\]

(since \( n \cdot ((n-1)(n-2)\cdots(n-k+1)) = n(n-1)\cdots(n-k+1) \). On the other hand, \( k! = k \cdot (k-1)! \) (since \( k \in \{1, 2, 3, \ldots \} \)). Also, \( k \in \{1, 2, 3, \ldots \} \subseteq \mathbb{N} \). Hence, the definition of \( \binom{n}{k} \) yields

\[
\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \quad (\text{since } k! = k \cdot (k-1)!). \quad (18)
\]

Multiplying both sides of this equality by \( k \), we obtain

\[
k \binom{n}{k} = k \cdot \frac{n(n-1)\cdots(n-k+1)}{k \cdot (k-1)!} = \frac{n(n-1)\cdots(n-k+1)}{(k-1)!} = \binom{n-1}{k-1} \quad (\text{by } (18)). \quad \text{This proves } (17).
\]
This proves Lemma 5.4 (a).

(b) We have \( i - 1 \in \mathbb{N} \) (since \( i \) is a positive integer). Thus, the definition of \( L_{r-1,s,j,i-1} \) yields

\[
L_{r-1,s,j,i-1} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \left(\frac{(r-1) - sk}{i-1}\right)
= \sum_{k=0}^{j} (-1)^k \binom{j}{k} \left(\frac{r - sk - 1}{i-1}\right).
\]

(20)

Also, the definition of \( L_{r-1-s,s,j-1,i-1} \) yields

\[
L_{r-1-s,s,j-1,i-1} = \sum_{k=1}^{j} (-1)^{k-1} \binom{j-1}{k-1} \left(\frac{(r-1) - s(k-1) - sk}{i-1}\right)
= \sum_{k=1}^{j} (-1)^{k-1} \binom{j-1}{k-1} \left(\frac{r - sk - 1}{i-1}\right)
\]

(here, we have substituted \( k - 1 \) for \( k \) in the sum)

\[
= \sum_{k=1}^{j} (-1)^{k-1} \binom{j-1}{k-1} \left(\frac{r - sk - 1}{i-1}\right).
\]

(21)

Now,

\[
\sum_{k=0}^{j} (-1)^k k \binom{j}{k} \left(\frac{r - sk - 1}{i-1}\right)
= (-1)^0 \binom{j}{0} \left(\frac{r - s0 - 1}{i-1}\right) + \sum_{k=1}^{j} (-1)^k \binom{j}{k} \left(\frac{r - sk - 1}{i-1}\right)
= \sum_{k=1}^{j} (-1)^{k-1} \binom{j-1}{k-1} \left(\frac{r - sk - 1}{i-1}\right)
\]

(by (17), applied to \( j \) instead of \( n \))

(here, we have split off the addend for \( k = 0 \) from the sum)

\[
= \sum_{k=1}^{j} (-1)^{k-1} \binom{j}{k} \left(\frac{r - sk - 1}{i-1}\right) = -j L_{r-1-s,s,j-1,i-1}
\]

(by (21))

(22)

Also, \( i \in \{1, 2, 3, \ldots\} \) (since \( i \) is a positive integer). Multiplying both sides of the equality
Theorem 5.5. the exercise demands: This proves Lemma 5.4

Proof of Theorem 5.5. Darij Grinberg, 0000000

\[iL_{r,s,j,i} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \left( \begin{array}{c} r - sk \hfill i \\ \end{array} \right) = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \left( \begin{array}{c} r - sk \\ i \end{array} \right) = \sum_{k=0}^{j} \binom{j}{k} \left( \begin{array}{c} r - sk \\ i \end{array} \right) = \sum_{k=0}^{j} \binom{j}{k} \left( \begin{array}{c} r - sk \\ i \end{array} \right) = (r - sk) \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right)
\]

Proof. Let \(H\). Hence, (a) We have (by (20))

\[L_{r,s,j,i} = \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right) \]

\[= (r - sk) \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right)
\]

\[= \sum_{k=0}^{j} \binom{j}{k} \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right) - s \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right)
\]

\[= r \sum_{k=0}^{j} \binom{j}{k} \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right) - s \sum_{k=0}^{j} \binom{j}{k} \left( \begin{array}{c} r - sk - 1 \\ i - 1 \end{array} \right)
\]

This proves Lemma 5.4 (b).

Using the above lemma and a straightforward induction, we can now prove more than the exercise demands:

Theorem 5.5. Let \(j \in \mathbb{N}, r \in \mathbb{R}\) and \(s \in \mathbb{R}\).

(a) We have \(L_{r,s,j,i} = 0\) for each \(i \in \mathbb{N}\) satisfying \(i < j\).

(b) We have \(L_{r,s,j,j} = s^j\).

Proof of Theorem 5.5. Forget that we fixed \(j, r\) and \(s\).

(a) We can rewrite the claim of Theorem 5.5 (a) as follows:

\[L_{r,s,j,i} = 0 \quad \text{for each} \quad r \in \mathbb{R}, \quad s \in \mathbb{R}, \quad j \in \mathbb{N}, \quad i \in \mathbb{N} \quad \text{satisfying} \quad i < j. \quad (23)
\]

Let us prove (23) by induction on \(i\):

Induction base: We have \(L_{r,s,j,0} = 0\) for each \(r \in \mathbb{R}, s \in \mathbb{R}\) and \(j \in \mathbb{N}\) satisfying \(0 < j\). In other words, (23) holds for \(i = 0\). This completes the induction base.

Induction step: Let \(h \in \mathbb{N}\). Assume that (23) holds for \(i = h\). We must prove that (23) holds for \(i = h + 1\).

We have assumed that (23) holds for \(i = h\). In other words, we have

\[L_{r,s,j,h} = 0 \quad \text{for each} \quad r \in \mathbb{R}, \quad s \in \mathbb{R}, \quad j \in \mathbb{N} \quad \text{satisfying} \quad h < j. \quad (24)
\]

Now, let \(r \in \mathbb{R}, s \in \mathbb{R}\) and \(j \in \mathbb{N}\) be such that \(h + 1 < j\). We shall prove that \(L_{r,s,j,h+1} = 0\). We have \(h < h + 1 < j\). Hence, (24) (applied to \(r - 1\) instead of \(r\)) yields \(L_{r-1,s,j,h} = 0\). Also, from \(h < j\), we obtain \(j > h \geq 0\); thus, \(j\) is a positive integer. Hence, \(j - 1 \in \mathbb{N}\). Also,

\[\text{Proof. Let} \quad r \in \mathbb{R}, \quad s \in \mathbb{R} \quad \text{and} \quad j \in \mathbb{N} \quad \text{be such that} \quad 0 < j. \quad \text{From} \quad 0 < j, \quad \text{we obtain} \quad j > 0, \quad \text{so that} \quad j \neq 0.
\]

Hence, \(|j = 0| = 0\). Now, Lemma 5.4 (a) yields \(L_{r,s,j,0} = |h + 1 = 0| = 0. \text{Qed.}

Darij Grinberg, 0000000000

dgrinber@umn.edu
$h < j - 1$ (since $h + 1 < j$). Thus, (24) (applied to $r - 1 - s$ and $j - 1$ instead of $r$ and $j$) yields $L_{r-1-s,s,j-1,h} = 0$.

But $h+1$ is a positive integer (since $h \in \mathbb{N}$). Hence, Lemma 5.4 (b) (applied to $i = h+1$) yields

$$(h+1) L_{r,s,j,h+1} = r L_{r-1,s,j,(h+1)-1} + s j L_{r-1-s,s,j-1,(h+1)-1} = r L_{r-1,s,j,h} + s j L_{r-1-s,s,j-1,h} = 0$$

(since $(h+1) - 1 = h$)

$= r0 + sj0 = 0$.

We can divide both sides of this equality by $h + 1$ (since $h + 1 \geq 1 > 0$), and thus obtain $L_{r,s,j,h+1} = 0$.

Now, forget that we fixed $r$, $s$ and $j$. We thus have shown that

$$L_{r,s,j,h+1} = 0 \quad \text{for each } r \in \mathbb{R}, s \in \mathbb{R} \text{ and } j \in \mathbb{N} \text{ satisfying } h + 1 < j.$$  

In other words, (23) holds for $i = h+1$. This completes the induction step. Hence, (23) is proven by induction.

In other words, Theorem 5.5 (a) is proven.

(b) Let us prove Theorem 5.5 (b) by induction on $j$:

Induction base: We have $L_{r,s,0,0} = s^0$ for each $r \in \mathbb{R}$ and $s \in \mathbb{R}$. In other words, Theorem 5.5 (b) holds for $j = 0$. This completes the induction base.

Induction step: Let $h \in \mathbb{N}$. Assume that Theorem 5.5 (b) holds for $j = h$. We must prove that Theorem 5.5 (b) holds for $j = h + 1$.

We have assumed that Theorem 5.5 (b) holds for $j = h$. In other words, we have

$$L_{r,s,h,h} = s^h \quad \text{for each } r \in \mathbb{R} \text{ and } s \in \mathbb{R}.$$ (25)

Now, let $r \in \mathbb{R}$ and $s \in \mathbb{R}$. We shall prove that $L_{r,s,h+1,h+1} = s^{h+1}$.

We have $h < h + 1$. Hence, Theorem 5.5 (a) (applied to $r - 1$, $h + 1$ and $h$ instead of $r$, $j$ and $i$) yields $L_{r-1,h+1,h} = 0$. Also, (25) (applied to $r - 1 - s$ instead of $r$) yields $L_{r-1-s,h+1,h} = s^h$.

But $h+1$ is a positive integer (since $h \in \mathbb{N}$). Hence, Lemma 5.4 (b) (applied to $j = h+1$ and $i = h + 1$) yields

$$(h+1) L_{r,s,h+1,h+1} = r L_{r-1,s,h+1,(h+1)-1} + s (h+1) L_{r-1-s,s,(h+1)-1,(h+1)-1} = r L_{r-1,s,h+1,h} + s (h+1) L_{r-1-s,s,h,h} = r0 + s (h+1) s^h = s (h+1) s^h = (h+1) s s^h.$$

We can divide both sides of this equality by $h + 1$ (since $h + 1 \geq 1 > 0$), and thus obtain $L_{r,s,h+1,h+1} = ss^h = s^{h+1}$.

Now, forget that we fixed $r$ and $s$. We thus have shown that

$$L_{r,s,h+1,h+1} = s^{h+1} \quad \text{for each } r \in \mathbb{R} \text{ and } s \in \mathbb{R}.$$ 

In other words, Theorem 5.5 (b) holds for $j = h + 1$. This completes the induction step. Hence, Theorem 5.5 (b) is proven by induction.

\[\text{Proof. Let } r \in \mathbb{R} \text{ and } s \in \mathbb{R}. \text{ Lemma 5.4 (a) (applied to } j = 0 \text{) yields } L_{r,s,0,0} = |0 = 0| = 1 = s^0. \text{ Qed.}\]
Now, let \( j \in \mathbb{N} \), \( r \in \mathbb{R} \) and \( s \in \mathbb{R} \). Then, \( L_{r,s,j,j} = \sum_{k=0}^{j} (-1)^k \binom{j}{k} \left( r - \frac{s}{j} \right) \) (by the definition of \( L_{r,s,j,j} \)). Hence,

\[
\sum_{k=0}^{j} (-1)^k \binom{j}{k} \left( r - \frac{s}{j} \right) = L_{r,s,j,j} = s^j
\]

(by Theorem 5.5 (b)).

5.2.3 Remark

Theorem 5.5 (a) can also be proven in a similar way to the First solution; this time, however, we need to consider \( i \)-picks (i.e., ways to choose \( i \) of the \( r \) stones), and argue that no \( i \)-pick is legal when \( i < j \).

Yet another algebraic solution of the above exercise (using finite differences) can be found in:

- Ronald L. Graham, Donald E. Knuth, Oren Patashnik, _Concrete Mathematics_, 2nd edition 1994, proof of (5.43).

6 Exercise 6

6.1 Problem

Let \( n \in \mathbb{N} \). The summation sign \( \sum_{I \subseteq [n]} \) shall always stand for a sum over all subsets \( I \) of \([n]\). (This sum has \( 2^n \) addends.)

Let \( A_1, A_2, \ldots, A_n \) be \( n \) numbers or polynomials or square matrices of the same size. (Allowing matrices means that \( A_i A_j \) is not necessarily equal to \( A_j A_i \), so beware of using the binomial formula or similar identities!)

(a) Show that

\[
\sum_{I \subseteq [n]} (-1)^{|I|} \left( \sum_{i \in I} A_i \right)^m = \sum_{(i_1, i_2, \ldots, i_m) \in [n]^m, \{i_1, i_2, \ldots, i_m\} \neq [n]} A_{i_1} A_{i_2} \cdots A_{i_m} \quad \text{for all } m \in \mathbb{N}.
\]

(Example: If \( n = 2 \) and \( m = 3 \), then this is saying

\[
(A + B)^3 - A^3 - B^3 + 0^3 = AAB + ABA + ABB + BAA + BAB + BBA,
\]

where we have renamed \( A_1 \) and \( A_2 \) as \( A \) and \( B \).)

(b) Show that

\[
\sum_{I \subseteq [n]} (-1)^{|I|} \left( \sum_{i \in I} A_i \right)^m = 0 \quad \text{for all } m \in \mathbb{N} \text{ satisfying } m < n.
\]
(Example: If $n = 3$ and $m = 2$, then this is saying

\[(A + B + C)^2 - (A + B)^2 - (A + C)^2 - (B + C)^2 + A^2 + B^2 + C^2 - 0^2 = 0,\]

where we have renamed $A_1, A_2, A_3$ as $A, B, C$)

(c) Show that

\[
\sum_{I \subseteq \{1, \ldots, n\}} (-1)^{n-|I|} \left( \sum_{i \in I} A_i \right)^n = \sum_{\sigma \in S_n} A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(n)},
\]

where $S_n$ stands for the set of all $(n!)$ permutations of $\{1, \ldots, n\}$.

(Example: If $n = 3$, then this is saying

\[(A + B + C)^3 - (A + B)^3 - (A + C)^3 - (B + C)^3 + A^3 + B^3 + C^3 - 0^3
= ABC + ACB + BAC + BCA + CAB + CBA,
\]

where we have renamed $A_1, A_2, A_3$ as $A, B, C$.)

[Hint: You can use the product rule, which says the following:

**Proposition 6.1** (Product rule). Let $m$ and $n$ be two nonnegative integers. Let $P_{u,v}$, for all $u \in \{1, \ldots, m\}$ and $v \in \{1, \ldots, n\}$, be numbers or polynomials or square matrices of the same size. Then,

\[
(P_{1,1} + P_{1,2} + \cdots + P_{1,n}) (P_{2,1} + P_{2,2} + \cdots + P_{2,n}) \cdots (P_{m,1} + P_{m,2} + \cdots + P_{m,n})
= \sum_{(i_1, i_2, \ldots, i_m) \in \{1, \ldots, n\}^m} P_{i_1, i_2} P_{i_2, i_3} \cdots P_{i_m, i_1}.
\]

(This frightening formula merely says that a product of sums can be expanded, and the result will be a sum of products, with each of the latter products being obtained by multiplying together one addend from each sum. You have probably used this sometime already.)

]

6.2 REMARK

This exercise is Exercise 6.50 in:

- Darij Grinberg, *Notes on the combinatorial fundamentals of algebra (version of 4 October 2018)*

with a minor difference in its wording (namely,

\[
\sum_{(i_1, i_2, \ldots, i_m) \in \{1, \ldots, n\}^m; \ \{i_1, i_2, \ldots, i_m\} = \{1, \ldots, n\}} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

is rewritten as

\[
\sum_{f: [m] \to \{1, \ldots, n\}; \ \text{f is surjective}} A_{f(1)} A_{f(2)} \cdots A_{f(m)},
\]

by substituting $(f(1), f(2), \ldots, f(m))$ for $(i_1, i_2, \ldots, i_m)$ in the sum). The solution given below is more or less identical to the solution given in the reference just mentioned.

Identities like those in the above exercise are known as “polarization identities”.

Darij Grinberg, 00000000 25  dgrinber@umn.edu
6.3 Solution (sketched)

Before we come to actually solving the problem, let us restate Proposition 6.1 in a more convenient form:

Proposition 6.2 (Product rule). Let \( m \in \mathbb{N} \). Let \( I \) be a finite set. Let \( P_{u,v} \), for all \( u \in [m] \) and \( v \in I \), be numbers or polynomials or square matrices of the same size. Then,

\[
\left( \sum_{i \in I} P_{1,i} \right) \left( \sum_{i \in I} P_{2,i} \right) \cdots \left( \sum_{i \in I} P_{m,i} \right) = \sum_{(i_1,i_2,\ldots,i_m) \in I^m} P_{1,i_1} P_{2,i_2} \cdots P_{m,i_m}.
\]

Proof of Proposition 6.2. Let \( n = |I| \). Then, there is a bijection \( \phi : I \to [n] \). Consider such a \( \phi \).

The set \( I \) is only used for labeling the elements \( P_{u,v} \); thus, we can WLOG assume that \( I = [n] \) (since otherwise, we can just rename \( P_{u,v} \) as \( P_{u,\phi(v)} \)). Assume this; then, Proposition 6.2 follows immediately from Proposition 6.1. \( \square \)

We shall use the Iverson bracket notation.

Next, we recall the principle of “destructive interference” that we have seen in class (Theorem 2.24 in class work (2018-10-01)):

Proposition 6.3. Let \( G \) be a finite set. Then,

\[
\sum_{I \subseteq G} (-1)^{|I|} = [G = \emptyset].
\]

We shall use a slightly more general version of this principle:

Proposition 6.4. Let \( G \) be a finite set. Let \( S \) be a subset of \( G \). Then,

\[
\sum_{I \subseteq G; S \subseteq I} (-1)^{|I|} = (-1)^{|S|} [G = S].
\]

Example 6.5. Applying Proposition 6.4 to \( G = \{1, 2, 3, 4\} \) and \( S = \{1, 2\} \), we find

\[
(-1)^{|\{1,2\}|} + (-1)^{|\{1,2,3\}|} + (-1)^{|\{1,2,4\}|} + (-1)^{|\{1,2,3,4\}|} = (-1)^{|\{1,2\}|} [\{1, 2, 3, 4\} = \{1, 2\}].
\]

Indeed, both sides of this equality are 0 (the left hand side because the addends cancel; the right hand side because \( \{1, 2, 3, 4\} \neq \{1, 2\} \)).

Clearly, Proposition 6.3 is the particular case of Proposition 6.4 when \( S = \emptyset \) (since every subset \( I \) of \( G \) satisfies \( \emptyset \subseteq I \)).

There are two ways to prove Proposition 6.4. One is to derive it from Proposition 6.3 (by applying the latter proposition to \( G \setminus S \) instead of \( G \), using a bijection between \( \{I \subseteq G \setminus S\} \) and \( \{I \subseteq G \mid S \subseteq I\} \)). Another is by generalizing the 2nd proof that we gave for Proposition 6.3 in class. We shall follow the second way:

Proof of Proposition 6.4. If \( G = S \), then Proposition 6.4 holds. Hence, for the rest of this proof, we WLOG assume that we don’t have \( G = S \). Thus, \([G = S] = 0\). If we had \( G \subseteq S \), then we would have \( G = S \) (since \( S \subseteq G \)), which would contradict the fact that we don’t
have \( G = S \). Hence, we cannot have \( G \subseteq S \). In other words, we have \( G \nsubseteq S \). Hence, there exists some \( g \in G \) such that \( g \notin S \). Consider such a \( g \).

The map\(^8\)

\[
\{ I \subseteq G \mid S \subseteq I \text{ and } g \notin I \} \to \{ I \subseteq G \mid S \subseteq I \text{ and } g \in I \},
\]

\[
J \mapsto J \cup \{ g \}
\]
is well-defined (since \( g \in G \)). The map

\[
\{ I \subseteq G \mid S \subseteq I \text{ and } g \in I \} \to \{ I \subseteq G \mid S \subseteq I \text{ and } g \notin I \},
\]

\[
K \mapsto K \setminus \{ g \}
\]
is also well-defined (since \( g \notin S \), and therefore every set \( K \) satisfying \( S \subseteq K \) must also satisfy \( S \subseteq K \setminus \{ g \} \)). These two maps are mutually inverse\(^9\) and thus are bijections. Hence, in particular, the map

\[
\{ I \subseteq G \mid S \subseteq I \text{ and } g \notin I \} \to \{ I \subseteq G \mid S \subseteq I \text{ and } g \in I \},
\]

\[
J \mapsto J \cup \{ g \}
\]
is a bijection. Thus, we can substitute \( J \cup \{ g \} \) for \( I \) in the sum \( \sum_{\substack{I \subseteq G; \\
S \subseteq I; \\
g \notin I}} (-1)^{|I|} \). We thus obtain

\[
\sum_{\substack{I \subseteq G; \\
S \subseteq I; \\
g \notin I}} (-1)^{|I|} = \sum_{\substack{J \subseteq G; \\
S \subseteq J; \\
g \notin J}} (-1)^{|J\cup\{g\}|} = \sum_{\substack{J \subseteq G; \\
S \subseteq J; \\
g \notin J}} (-1)^{|J|+1} = - \sum_{\substack{J \subseteq G; \\
S \subseteq J; \\
g \notin J}} (-1)^{|J|}
\]

\[
= - \sum_{\substack{I \subseteq G; \\
S \subseteq I; \\
g \notin I}} (-1)^{|I|}
\]

(here, we have renamed the summation index \( J \) as \( I \)).

But there exists only one subset \( I \) of \( G \) satisfying \( G \subseteq I \): namely, \( G \) itself. Thus, the sum \( \sum_{\substack{I \subseteq G; \\
G \subseteq I}} (-1)^{|I|} \) has only one addend: namely, the addend for \( I = G \). Hence, this sum simplifies as follows: \( \sum_{\substack{I \subseteq G; \\
G \subseteq I}} (-1)^{|I|} = (-1)^{|G|} \). Hence,

\[
\sum_{\substack{I \subseteq G; \\
G \subseteq I}} (-1)^{|I|} = \sum_{\substack{I \subseteq G; \\
G \subseteq I}} (-1)^{|I|} = (-1)^{|G|} = (-1)^{|S|} \quad \text{(since } G = S \text{)}.
\]

Comparing this with \( (-1)^{|S|} \frac{|G = S|}{|G|} = (-1)^{|S|} \), we obtain \( \sum_{\substack{I \subseteq G; \\
S \subseteq I}} (-1)^{|I|} = (-1)^{|S|} \frac{|G = S|}{|G|} \). Hence,

Proposition\(^6,3\) is proven under the assumption that \( G = S \).

\(^8\)The notation \( \{ I \subseteq G \mid S \subseteq I \text{ and } g \notin I \} \) means “the set of all subsets \( I \) of \( G \) satisfying \( S \subseteq I \) and \( g \notin I \)”. Similarly, the notation \( \{ I \subseteq G \mid S \subseteq I \text{ and } g \in I \} \) should be understood.

\(^9\)because:

- every subset \( J \) of \( G \) satisfying \( g \notin J \) must satisfy \( (J \cup \{ g \}) \setminus \{ g \} = J \);
- every subset \( K \) of \( G \) satisfying \( g \in K \) must satisfy \( (K \setminus \{ g \}) \cup \{ g \} = K \)
But each subset $I$ of $G$ satisfying $S \subseteq I$ must satisfy either $g \in I$ or $g \notin I$. Hence, we can split the sum $\sum_{I \subseteq G, \ S \subseteq I} (-1)^{|I|}$ as follows:

$$\sum_{I \subseteq G, \ S \subseteq I} (-1)^{|I|} = \sum_{I \subseteq G: \ S \subseteq I, \ g \in I} (-1)^{|I|} + \sum_{I \subseteq G: \ S \subseteq I, \ g \notin I} (-1)^{|I|} = -\sum_{I \subseteq G: \ S \subseteq I, \ g \notin I} (-1)^{|I|} = 0.$$ 

Comparing this with

$$(-1)^{|S|} [G = S] = 0,$$

we obtain $\sum_{I \subseteq G, \ S \subseteq I} (-1)^{|I|} = (-1)^{|S|} [G = S]$. This proves Proposition 6.4.

**Corollary 6.6.** Let $G$ be a finite set. Let $S$ be a subset of $G$. Then,

$$\sum_{I \subseteq G, \ S \subseteq I} (-1)^{|G|-|I|} = [G = S].$$

**Proof of Corollary 6.6.** We have

$$\sum_{I \subseteq G, \ S \subseteq I} (-1)^{|G|-|I|} = \sum_{I \subseteq G: \ S \subseteq I, \ |G|-|I| \equiv |G|+|I| \mod 2} (-1)^{|G|+|I|}$$

(by Proposition 6.4)

$$= \sum_{I \subseteq G} (-1)^{|G|-|I|} = \sum_{I \subseteq G} (-1)^{|G|} (-1)^{|I|}$$

$$= (-1)^{|G|} \sum_{I \subseteq G} (-1)^{|I|} = (-1)^{|G|} (-1)^{|S|}$$

(by Proposition 6.4)

$$= (-1)^{|G|} (-1)^{|S|} \begin{cases} 1, & \text{if } G = S; \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} (-1)^{|G|} (-1)^{|S|} \cdot 1, & \text{if } G = S; \\ 0, & \text{otherwise} \end{cases}$$

because if $G = S$,

$$\begin{align*}
\text{then } (-1)^{|G|} (-1)^{|S|} \cdot 1 &= (-1)^{|S|} (-1)^{|S|} \cdot 1 = (-1)^{|S|^2} = (-1)^{2|S|} = 1 \\
&= [G = S].
\end{align*}$$

This proves Corollary 6.6.

Let us now proceed to the solution of the problem.
(a) Fix \( m \in \mathbb{N} \). For each subset \( I \) of \([n]\), we have

\[
\left( \sum_{i \in I} A_i \right)^m = \left( \sum_{i \in I} A_i \right) \left( \sum_{i \in I} A_i \right) \cdots \left( \sum_{i \in I} A_i \right) = \sum_{(i_1, i_2, \ldots, i_m) \in I^m} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

(by Proposition 6.2 applied to \( P_{u,v} = A_v \)). Hence,

\[
\sum_{I \subseteq [n]} (-1)^{|n| - |I|} \left( \sum_{i \in I} A_i \right)^m = \sum_{(i_1, i_2, \ldots, i_m) \in I^m} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

\[
\sum_{I \subseteq [n]} (-1)^{|n| - |I|} \sum_{(i_1, i_2, \ldots, i_m) \in I^m} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

\[
\sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} (-1)^{|n| - |I|} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

\[
\sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} \sum_{\{i_1, i_2, \ldots, i_m\} \subseteq I} (-1)^{|n| - |I|} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

\[
\sum_{\{i_1, i_2, \ldots, i_m\} \subseteq I} \sum_{(i_1, i_2, \ldots, i_m) \in [n]^m} (-1)^{|n| - |I|} A_{i_1} A_{i_2} \cdots A_{i_m}
\]

\[
\sum_{\{i_1, i_2, \ldots, i_m\} \subseteq I} [\{i_1, i_2, \ldots, i_m\} = [n]] A_{i_1} A_{i_2} \cdots A_{i_m}
\]

\[
\sum_{\{i_1, i_2, \ldots, i_m\} \subseteq I} [\{i_1, i_2, \ldots, i_m\} = [n]] A_{i_1} A_{i_2} \cdots A_{i_m}
\]

Since each \((i_1, i_2, \ldots, i_m) \in [n]^m\) satisfies either \(\{i_1, i_2, \ldots, i_m\} = [n]\) or \(\{i_1, i_2, \ldots, i_m\} \neq [n] \) (but not both)
This solves part (a) of the exercise.

(b) Let $m \in \mathbb{N}$ be such that $m < n$. Then, each $(i_1, i_2, \ldots, i_m) \in [n]^m$ satisfies $\{i_1, i_2, \ldots, i_m\} = m < n = [n]$ and therefore $\{i_1, i_2, \ldots, i_m\} \neq [n]$, so that $\{i_1, i_2, \ldots, i_m\} \neq [n]$. In other words, there exists no $(i_1, i_2, \ldots, i_m) \in [n]^m$ satisfying $\{i_1, i_2, \ldots, i_m\} = [n]$. Hence, the sum
\[
\sum_{(i_1, i_2, \ldots, i_m) \in [n]^m; \{i_1, i_2, \ldots, i_m\} = [n]} A_{i_1} A_{i_2} \cdots A_{i_m} = 0.
\]

Hence, part (a) of the exercise yields
\[
\sum_{I \subseteq [n]} (-1)^{n-|I|} \left( \sum_{i \in I} A_i \right)^m = \sum_{(i_1, i_2, \ldots, i_m) \in [n]^m; \{i_1, i_2, \ldots, i_m\} = [n]} A_{i_1} A_{i_2} \cdots A_{i_m} = 0.
\]

This solves part (b) of the exercise.

(c) Consider the set $S_n$ of all permutations of the set $[n]$. If $\sigma \in S_n$, then the $n$-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ is an $(i_1, i_2, \ldots, i_n) \in [n]^n$ satisfying $\{i_1, i_2, \ldots, i_n\} = [n]$ (because $\{\sigma(1), \sigma(2), \ldots, \sigma(n)\} = \sigma([n]) = [n]$ (since $\sigma$ is surjective).10 Thus, the map
\[
S_n \to \{(i_1, i_2, \ldots, i_n) \in [n]^n \mid \{i_1, i_2, \ldots, i_n\} = [n]\},
\]
\[
\sigma \mapsto (\sigma(1), \sigma(2), \ldots, \sigma(n))
\]
is well-defined. Moreover, this map is injective (since a permutation $\sigma \in S_n$ is uniquely determined by $(\sigma(1), \sigma(2), \ldots, \sigma(n))$) and surjective.11 Hence, this map is bijective, i.e., is a bijection.

Now, part (a) of the exercise (applied to $m = n$) yields
\[
\sum_{I \subseteq [n]} (-1)^{n-|I|} \left( \sum_{i \in I} A_i \right)^n = \sum_{(i_1, i_2, \ldots, i_n) \in [n]^n; \{i_1, i_2, \ldots, i_n\} = [n]} A_{i_1} A_{i_2} \cdots A_{i_n} = \sum_{\sigma \in S_n} A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(n)}
\]

10Note that this $n$-tuple $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ is the one-line notation of the permutation $\sigma$.

11Proof. Let $(j_1, j_2, \ldots, j_n) \in \{(i_1, i_2, \ldots, i_n) \in [n]^n \mid \{i_1, i_2, \ldots, i_n\} = [n]\}$. We must prove that there exists some $\sigma \in S_n$ such that $(j_1, j_2, \ldots, j_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$.

Indeed, we have $(j_1, j_2, \ldots, j_n) \in \{(i_1, i_2, \ldots, i_n) \in [n]^n \mid \{i_1, i_2, \ldots, i_n\} = [n]\}$. In other words, $(j_1, j_2, \ldots, j_n)$ is an $n$-tuple in $[n]^n$ and satisfies $\{j_1, j_2, \ldots, j_n\} = [n]$. Now, let $f : [n] \to [n]$ be the map that sends $1, 2, \ldots, n$ to $j_1, j_2, \ldots, j_n$, respectively. Then, the image of this map $f$ is $f([n]) = \{j_1, j_2, \ldots, j_n\} = [n]$; hence, this map $f$ is surjective. But a surjective map between two finite sets of the same size must always be bijective (by the Pigeonhole Principle for surjections). Hence, $f$ is bijective (since $f$ is a surjective map between two finite sets of the same size). Thus, $f$ is a bijection from $[n]$ to $[n]$. In other words, $f$ is a permutation of $[n]$. In other words, $f \in S_n$. Also, the definition of $f$ yields $(f(1), f(2), \ldots, f(n)) = (j_1, j_2, \ldots, j_n)$, so that $(j_1, j_2, \ldots, j_n) = (f(1), f(2), \ldots, f(n))$. Hence, there exists some $\sigma \in S_n$ such that $(j_1, j_2, \ldots, j_n) = (\sigma(1), \sigma(2), \ldots, \sigma(n))$ (namely, $\sigma = f$). Qed.
(here, we have substituted \((\sigma(1), \sigma(2), \ldots, \sigma(n))\) for \((i_1, i_2, \ldots, i_n)\) in the sum, because the map

\[
S_n \to \{(i_1, i_2, \ldots, i_n) \in [n]^n \mid \{i_1, i_2, \ldots, i_n\} = [n]\},
\]

\[
\sigma \mapsto (\sigma(1), \sigma(2), \ldots, \sigma(n))
\]

is a bijection). This solves part \((c)\) of the exercise.