Math 4990 Fall 2017 (Darij Grinberg): homework set 4 with solutions

Contents

0.1. On the Euler totient function

Let us first recall some basic facts from elementary number theory.

A *common divisor* of two integers a and b is an integer that divides both a and b. If a and b are two integers satisfying $(a,b) \neq (0,0)$, then $\gcd(a,b)$ is defined to be the greatest of all common divisors of a and b. For the sake of completeness, we also define $\gcd(0,0) = 0$ (so that $\gcd(a,b)$ is defined whenever a and b are two integers). For any two integers a and b, we call $\gcd(a,b)$ the *greatest common divisor* of a and b (even though, in the case when $(a,b) = (0,0)$, it is not literally the greatest among all common divisors of a and b). For example,

$$\begin{aligned} \gcd(2,3) &= 1, & \gcd(5,10) &= 5, & \gcd(4,6) &= 2, \\ \gcd(-1,4) &= 1, & \gcd(-2,-6) &= 2, & \gcd(0,8) &= 8, \\ \gcd(6,15) &= 3, & \gcd(-6,15) &= 3, & \gcd(0,0) &= 0. \end{aligned}$$

Two integers a and b are said to be *coprime* if their greatest common divisor is 1. We say that an integer a is *coprime* to an integer b if a and b are coprime.

(Note that $\gcd(a,0) = |a|$ for any integer a. Hence, the only integers coprime to 0 are 1 and -1.)

The *Euler totient function* $\phi : \{1,2,3,\ldots\} \to \mathbb{N}$ is defined by

$$\begin{aligned} \phi(n) &= (\text{the number of all } m \in [n] \text{ that are coprime to } n) \\ &= |\{m \in [n] \mid m \text{ is coprime to } n\}|. \end{aligned} \tag{1}$$

More about greatest common divisors and about this function can be found in [LeLeMe17, Chapter 9].[1] That said, you won't need anything but the definitions in this homework set.

[1] The probably most important fact is the following:

The greatest common divisor $\gcd(a,b)$ of two integers a and b can be characterized by the following property: It is the unique nonnegative common divisor g of a and b such that every common divisor of a and b must divide g.

Exercise 1. Let n be a positive integer.

(a) Prove that if m is an integer coprime to n, then $n - m$ is also an integer coprime to n.

(b) Prove that $\phi(n)$ is even if $n > 2$.

[**Hint:** If you haven't used the $n > 2$ requirement, then you must have missed something. $\phi(2) = 1$, which is not even!]

Before we solve this exercise, let us state a basic fact about the greatest common divisor of integers:

Lemma 0.1. Let n be a nonzero integer. Let m be an integer. Let $g \in \mathbb{Z}$. Then,

$$\gcd(gn - m, n) = \gcd(m, n).$$

Proof of Lemma 0.1. Each common divisor of m and n is also a common divisor of $gn - m$ and n [2]. In other words, we have

$$\begin{aligned} &(\text{the set of all common divisors of } m \text{ and } n) \\ &\subseteq (\text{the set of all common divisors of } gn - m \text{ and } n). \end{aligned} \tag{2}$$

But we can apply the same reasoning to $gn - m$ instead of m. We thus obtain

$$\begin{aligned} &(\text{the set of all common divisors of } gn - m \text{ and } n) \\ &\subseteq (\text{the set of all common divisors of } gn - (gn - m) \text{ and } n) \\ &= (\text{the set of all common divisors of } m \text{ and } n) \end{aligned}$$

(since $gn - (gn - m) = m$). Combining this inclusion with (2), we obtain

$$\begin{aligned} &(\text{the set of all common divisors of } m \text{ and } n) \\ &= (\text{the set of all common divisors of } gn - m \text{ and } n). \end{aligned} \tag{3}$$

But n is nonzero. Hence, $n \neq 0$; thus, $(m, n) \neq (0, 0)$. Thus, $\gcd(m, n)$ is the greatest of all common divisors of m and n (by the definition of $\gcd(m, n)$). In other words,

$$\gcd(m, n) = \max(\text{the set of all common divisors of } m \text{ and } n).$$

[2] *Proof.* Let d be a common divisor of m and n. We must show that d is also a common divisor of $gn - m$ and n.

We know that d divides both m and n (since d is a common divisor of m and n). Thus, there exists some $m' \in \mathbb{Z}$ such that $m = dm'$ (since d divides m), and there exists some $n' \in \mathbb{Z}$ such that $n = dn'$ (since d divides n). Consider these m' and n'. We have $g \underbrace{n}_{=dn'} - \underbrace{m}_{=dm'} = gdn' - dm' = d(gn' - m')$. Hence, d divides $gn - m$.

Now, the integer d divides both $gn - m$ and n. In other words, d is a common divisor of $gn - m$ and n. This completes our proof.

In view of (3), this rewrites as

$$\gcd(m,n) = \max(\text{the set of all common divisors of } gn-m \text{ and } n). \tag{4}$$

On the other hand, $(gn-m,n) \neq (0,0)$ (since $n \neq 0$). Hence, $\gcd(gn-m,n)$ is the greatest of all common divisors of $gn-m$ and n (by the definition of $\gcd(gn-m,n)$). In other words,

$$\gcd(gn-m,n) = \max(\text{the set of all common divisors of } gn-m \text{ and } n).$$

Comparing this with (4), we obtain $\gcd(gn-m,n) = \gcd(m,n)$. This proves Lemma 0.1. □

Solution to Exercise 1. **(a)** Let m be an integer coprime to n. We must prove that $n-m$ is also an integer coprime to n.

Clearly, $n-m$ is an integer. Furthermore, we know that m is coprime to n. In other words, m and n are coprime. In other words, $\gcd(m,n)=1$ (by the definition of "coprime"). But n is nonzero (since n is positive); hence, Lemma 0.1 (applied to $g=1$) yields $\gcd(1n-m,n) = \gcd(m,n) = 1$. In other words, $1n-m$ and n are coprime (by the definition of "coprime"). In other words, $1n-m$ is coprime to n. In other words, $n-m$ is coprime to n (since $1n=n$). This solves Exercise 1 **(a)**.

(b) Assume that $n>2$. Define a subset K of $[n]$ by $K = \{m \in [n] \mid m \text{ is coprime to } n\}$. Thus,

$$|K| = |\{m \in [n] \mid m \text{ is coprime to } n\}| = \phi(n) \tag{5}$$

(by (1)). Also, $n \notin K$ [3] and $n/2 \notin K$ [4]. Also,

$$n-i \in K \qquad \text{for each } i \in K \tag{6}$$

[5].

[3] *Proof.* Assume the contrary. Thus, $n \in K = \{m \in [n] \mid m \text{ is coprime to } n\}$. In other words, $n \in [n]$ and n is coprime to n. Hence, n and n are coprime (since n is coprime to n). In other words, $\gcd(n,n)=1$.

But n is a positive integer; hence, $\gcd(n,n)=n$. Therefore, $n = \gcd(n,n) = 1 \leq 2$; this contradicts $n>2$. This contradiction shows that our assumption was wrong, qed.

[4] *Proof.* Assume the contrary. Thus, $n/2 \in K = \{m \in [n] \mid m \text{ is coprime to } n\}$. In other words, $n/2 \in [n]$ and $n/2$ is coprime to n. Hence, $n/2$ is a positive integer (since $n/2 \in [n]$). Also, $n/2$ and n are coprime (since $n/2$ is coprime to n). In other words, $\gcd(n/2,n)=1$.

If a and b are two positive integers satisfying $a \mid b$, then $\gcd(a,b)=a$. Applying this to $a=n/2$ and $b=n$, we obtain $\gcd(n/2,n)=n/2$ (since $n/2 \mid n$). Therefore, $n/2 = \gcd(n/2,n) = 1$, so that $n=2$; this contradicts $n>2$. This contradiction shows that our assumption was wrong, qed.

[5] *Proof.* Let $i \in K$. Thus, $i \in K = \{m \in [n] \mid m \text{ is coprime to } n\}$. In other words, $i \in [n]$, and i is coprime to n. Hence, Exercise 1 **(a)** (applied to $m=i$) shows that $n-i$ is an integer coprime to n. If we had $i=n$, then we would have $n=i\in K$, which would contradict $n \notin K$. Hence, we cannot have $i=n$. We thus have $i \neq n$. Combined with $i \in [n]$, this yields $i \in [n] \setminus \{n\} = \{1,2,\ldots,n-1\}$, so that $n-i \in \{1,2,\ldots,n-1\} \subseteq [n]$. Therefore, $n-i \in [n]$ and $n-i$ is coprime to n. In other words, $n-i \in \{m \in [n] \mid m \text{ is coprime to } n\}$. In view of $K = \{m \in [n] \mid m \text{ is coprime to } n\}$, this rewrites as $n-i \in K$. Qed.

Now, define two subsets A and B of K by

$$A = \{m \in K \mid m < n/2\} \qquad \text{and} \qquad B = \{m \in K \mid m \geq n/2\}.$$

From $A = \{m \in K \mid m < n/2\}$, we obtain

$$K \setminus A = K \setminus \{m \in K \mid m < n/2\} = \left\{ m \in K \mid \underbrace{\text{we don't have } m < n/2}_{\iff (m \geq n/2)} \right\}$$
$$= \{m \in K \mid m \geq n/2\} = B.$$

But A is a subset of K; thus, $|K \setminus A| = |K| - |A|$. Hence, $|K| - |A| = |K \setminus A| = |B|$ (since $K \setminus A = B$), so that $|K| = |A| + |B|$. Comparing this with (5), we find

$$\phi(n) = |A| + |B|. \tag{7}$$

Also, $n - i \in B$ for each $i \in A$ [6]. Hence, we can define a map $\alpha : A \to B$ by

$$\alpha(i) = n - i \qquad \text{for each } i \in A.$$

Consider this map α.

Furthermore, $n - i \in A$ for each $i \in B$ [7]. Thus, we can define a map $\beta : B \to A$ by

$$\beta(i) = n - i \qquad \text{for each } i \in B.$$

Consider this map β.

We have $\alpha \circ \beta = \text{id}$ (since each $i \in B$ satisfies

$$(\alpha \circ \beta)(i) = \alpha(\beta(i)) = n - \underbrace{\beta(i)}_{\substack{=n-i \\ \text{(by the definition of } \beta)}} \qquad \text{(by the definition of } \alpha)$$
$$= n - (n - i) = i = \text{id}(i)$$

) and similarly $\beta \circ \alpha = \text{id}$. Hence, the maps α and β are mutually inverse. Thus, the map α is invertible, i.e., a bijection. We thus have found a bijection from A to B (namely, α); we conclude that $|A| = |B|$. But (7) becomes $\phi(n) = \underbrace{|A|}_{=|B|} + |B| = |B| + |B| = 2|B|$. Hence, $\phi(n)$ is even. This solves Exercise 1 **(b)**. □

[6] *Proof.* Let $i \in A$. We must show that $n - i \in B$.

We have $i \in A = \{m \in K \mid m < n/2\}$. In other words, $i \in K$ and $i < n/2$. Now, (6) shows that $n - i \in K$. Also, $n - \underbrace{i}_{<n/2} > n - n/2 = n/2$, so that $n - i \geq n/2$. Thus, we have shown that $n - i \in K$ and $n - i \geq n/2$. In other words, $n - i \in \{m \in K \mid m \geq n/2\}$. In view of $B = \{m \in K \mid m \geq n/2\}$, this rewrites as $n - i \in B$. Qed.

[7] *Proof.* Let $i \in B$. We must show that $n - i \in A$.

We have $i \in B = \{m \in K \mid m \geq n/2\}$. In other words, $i \in K$ and $i \geq n/2$. Now, (6) shows that $n - i \in K$. If we had $i = n/2$, then we would have $n/2 = i \in K$, which would contradict $n/2 \notin K$. Hence, we cannot have $i = n/2$. We thus have $i \neq n/2$, and therefore $i > n/2$ (since $i \geq n/2$). Hence, $n - \underbrace{i}_{>n/2} < n - n/2 = n/2$. Thus, we have shown that $n - i \in K$ and $n - i < n/2$. In other words, $n - i \in \{m \in K \mid m < n/2\}$. In view of $A = \{m \in K \mid m < n/2\}$, this rewrites as $n - i \in A$. Qed.

0.2. Variations on the binomial formula

Exercise 2. A preprint recently posted on the arXiv says (in a proof) that "

$$1-\sum_{i=2}^{d}(i-1)\binom{d}{i}\left(\frac{1}{d-1}\right)^{i}=0,$$

the final equality being verified by the computer algebra system Maple (which itself employs an algorithm of Zeilberger)". Here, d is assumed to be an integer ≥ 2.

Prove this equality by hand (but feel free to use a computer to write up your proof...). More generally, find and prove a sum-less expression for

$$\sum_{i=2}^{d}(i-1)\binom{d}{i}q^{i}$$

where q is an arbitrary rational number (and d is still an integer ≥ 2).

Before we solve Exercise 2, let us recall an identity about binomial coefficients:

Proposition 0.2. We have $k\binom{n}{k}=n\binom{n-1}{k-1}$ for any $n\in\mathbb{Q}$ and any positive integer k.

Proof of Proposition 0.2. Proposition 0.2 is Exercise 2 **(b)** on homework set #1, so we don't need to prove it again. □

Let us again recall the binomial formula:

Proposition 0.3. Let $n\in\mathbb{N}$. Let $x\in\mathbb{Q}$ and $y\in\mathbb{Q}$. Then,

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

Next, let us show a lemma:

Lemma 0.4. Let d be a positive integer. Let $q\in\mathbb{Q}$. Then,

$$\sum_{i=1}^{d}i\binom{d}{i}q^{i}=dq\left(q+1\right)^{d-1}.$$

Proof of Lemma 0.4. We have $d-1\in\mathbb{N}$ (since d is a positive integer). Thus, Proposition 0.3 (applied to $n=d-1$, $x=q$ and $y=1$) yields

$$(q+1)^{d-1}=\sum_{k=0}^{d-1}\binom{d-1}{k}q^{k}\underbrace{1^{d-1-k}}_{=1}=\sum_{k=0}^{d-1}\binom{d-1}{k}q^{k}=\sum_{i=1}^{d}\binom{d-1}{i-1}q^{i-1}\qquad(8)$$

(here, we have substituted $i-1$ for k in the sum). But

$$\sum_{i=1}^{d} \underbrace{i\binom{d}{i}}_{\substack{=d\binom{d-1}{i-1} \\ \text{(by Proposition 0.2,} \\ \text{applied to } n=d \text{ and } k=i)}} \underbrace{q^i}_{=qq^{i-1}} = \sum_{i=1}^{d} d\binom{d-1}{i-1} qq^{i-1} = dq \underbrace{\sum_{i=1}^{d} \binom{d-1}{i-1} q^{i-1}}_{\substack{=(q+1)^{d-1} \\ \text{(by (8))}}} = dq\,(q+1)^{d-1}.$$

This proves Lemma 0.4. □

Solution to Exercise 2. Let d be an integer ≥ 2. Let q be an arbitrary rational number. Proposition 0.3 (applied to $n=d$, $x=q$ and $y=1$) yields

$$\begin{aligned}(q+1)^d &= \sum_{k=0}^{d} \binom{d}{k} q^k \underbrace{1^{d-1-k}}_{=1} = \sum_{k=0}^{d} \binom{d}{k} q^k = \sum_{i=0}^{d} \binom{d}{i} q^i \\ &\qquad \text{(here, we have renamed the summation index } k \text{ as } i\text{)} \\ &= \underbrace{\binom{d}{0}}_{=1} \underbrace{q^0}_{=1} + \sum_{i=1}^{d} \binom{d}{i} q^i = 1 + \sum_{i=1}^{d} \binom{d}{i} q^i.\end{aligned}$$

Hence,

$$\sum_{i=1}^{d} \binom{d}{i} q^i = (q+1)^d - 1. \tag{9}$$

We have $d \geq 2$, so that

$$\begin{aligned}\sum_{i=1}^{d} (i-1)\binom{d}{i} q^i &= \underbrace{(1-1)}_{=0} \binom{d}{1} q^1 + \sum_{i=2}^{d} (i-1)\binom{d}{i} q^i \\ &= \underbrace{0\binom{d}{1} q^1}_{=0} + \sum_{i=2}^{d} (i-1)\binom{d}{i} q^i = \sum_{i=2}^{d} (i-1)\binom{d}{i} q^i.\end{aligned}$$

Hence,

$$\begin{aligned}\sum_{i=2}^{d}(i-1)\binom{d}{i}q^i &= \sum_{i=1}^{d}\underbrace{(i-1)\binom{d}{i}q^i}_{=i\binom{d}{i}q^i-\binom{d}{i}q^i} = \sum_{i=1}^{d}\left(i\binom{d}{i}q^i-\binom{d}{i}q^i\right)\\ &= \underbrace{\sum_{i=1}^{d}i\binom{d}{i}q^i}_{\substack{=dq(q+1)^{d-1}\\ \text{(by Lemma 0.4)}}} - \underbrace{\sum_{i=1}^{d}\binom{d}{i}q^i}_{\substack{=(q+1)^d-1\\ \text{(by (9))}}}\\ &= dq(q+1)^{d-1}-\left(\underbrace{(q+1)^d}_{=(q+1)(q+1)^{d-1}}-1\right)\\ &= dq(q+1)^{d-1}-\left((q+1)(q+1)^{d-1}-1\right)\\ &= (dq-(q+1))(q+1)^{d-1}+1.\end{aligned}$$

Applying this to $q=\dfrac{1}{d-1}$, we find

$$\sum_{i=2}^{d}(i-1)\binom{d}{i}\left(\frac{1}{d-1}\right)^i = \underbrace{\left(d\cdot\frac{1}{d-1}-\left(\frac{1}{d-1}+1\right)\right)}_{=0}\left(\frac{1}{d-1}+1\right)^{d-1}+1=1.$$

In other words, $1-\sum\limits_{i=2}^{d}(i-1)\binom{d}{i}\left(\dfrac{1}{d-1}\right)^i=0$. This solves Exercise 2. □

0.3. Screaming at feet

> **Exercise 3.** Let $n>1$ be an integer. Consider n people standing in a circle. Each of them looks down at someone else's feet (i.e., at the feet of one of the other $n-1$ persons). A bell sounds, and every person (simultaneously) looks up at the eyes of the person whose feet they have been ogling. If two people make eye contact, they scream. Show that the probability that no one screams is
>
> $$\sum_{k=0}^{n}(-1)^k\frac{n(n-1)\cdots(n-2k+1)}{(n-1)^{2k}\cdot 2^k\cdot k!}.$$
>
> Combinatorial restatement (feel free to solve this instead): A pair (i,j) of elements of $[n]$ is said to *scream* at a map $f:[n]\to[n]$ if it satisfies $f(i)=j$ and

$f(j) = i$. A map $f : [n] \to [n]$ is *silent* if no pair $(i,j) \in [n] \times [n]$ screams at f. Prove that the number of all silent maps $f : [n] \to [n]$ is

$$\sum_{k=0}^{n} (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!} (n-1)^{n-2k}.$$

Exercise 3 is [Camero16, Exercise 7.9]; a solution outline can be found in that references (make sure to check the errata). The solution we give below is essentially that solution from [Camero16, Exercise 7.9], with more details included.

We notice that the "staring game" described in Exercise 3 is known as the "Zen stare". See Sequence A134362 in the OEIS.

Our solution to Exercise 3 will rely on the Principle of Inclusion and Exclusion (see, e.g., [Galvin17, Theorem 16.1 and (12)]):

Theorem 0.5. 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.

Let us restate Theorem 0.5 in a way that doesn't require the finite sets to be indexed by $1, 2, \ldots, n$, but rather indexes them by any arbitrary finite set R:

Theorem 0.6. Let R be a finite set. For each $r \in R$, let A_r be a finite set.

(a) We have

$$\left| \bigcup_{r \in R} A_r \right| = \sum_{\substack{I \subseteq R; \\ I \neq \varnothing}} (-1)^{|I|-1} \left| \bigcap_{r \in I} A_r \right|.$$

(b) Let S be a finite set. Assume that A_r is a subset of S for each $r \in R$. Then,

$$\left| S \setminus \bigcup_{r \in R} A_r \right| = \sum_{I \subseteq R} (-1)^{|I|} \left| \bigcap_{r \in I} A_r \right|.$$

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

(Theorem 0.6 also appears in [Grinbe16b]: Namely, parts **(a)** and **(b)** of Theorem 0.6 are [Grinbe16b, Theorem 3.42] and [Grinbe16b, Theorem 3.43], respectively.)

Next, let us recall a basic formula, which was Exercise 1 **(a)** on homework set #2:

Proposition 0.7. Let $n \in \mathbb{N}$. Then,

$$(2n-1) \cdot (2n-3) \cdot \cdots \cdot 1 = \frac{(2n)!}{2^n n!}.$$

(The left hand side is understood to be the product of all odd integers from 1 to $2n-1$.)

Solution to Exercise 3 (sketched). In the following, an *edge* will mean a 2-element subset of $[n]$. For example, $\{1,3\}$ and $\{2,6\}$ are edges (if $n \geq 6$). Note that $\{1,3\} = \{3,1\}$. Note also that $\{1,1\}$ is not an edge (since it is not a 2-element set). Let R be the set of all edges.

Let S be the set of all maps $f : [n] \to [n]$ that have no fixed points[8]. If $r = \{i,j\}$ is an edge, then a map $f : [n] \to [n]$ is said to *coscream* at r if and only if it satisfies $f(i) = j$ and $f(j) = i$. (Clearly, this condition does not depend on how the edge r is written as $\{i,j\}$, because switching i with j merely interchanges the two conditions $f(i) = j$ and $f(j) = i$.) Thus, a map $f : [n] \to [n]$ coscreams at an edge $\{i,j\}$ if and only if the pair (i,j) screams at f. The following fact is now easy:

> *Observation 1:* Let $f : [n] \to [n]$ be a map. Then, f is silent if and only if f has no fixed point and coscreams at no edges.

[*Proof of Observation 1:* $\Longrightarrow$: Assume that f is silent. In other words, no pair $(i,j) \in [n] \times [n]$ screams at f. Hence, f has no fixed point (because if x was a fixed point of f, then the pair (x,x) would scream at f), and coscreams at no edges (because if f would coscream at an edge $\{i,j\}$, then the pair (i,j) would scream at f). This proves the "$\Longrightarrow$" direction (i.e., the "only if" direction) of Observation 1.

$\Longleftarrow$: Assume that f has no fixed point and coscreams at no edges. Then, no pair $(i,j) \in [n] \times [n]$ screams at f: In fact, if any pair $(i,j) \in [n] \times [n]$ would scream at f, then either i would be a fixed point of f (when $i = j$), or f would coscream at the edge $\{i,j\}$ (if $i \neq j$); but both of these possibilities are impossible by our assumption. Hence, the map f is silent. This proves the "$\Longleftarrow$" direction (i.e., the "if" direction) of Observation 1.]

Next, we shall study subsets I of R such that all edges in I are disjoint. For example, $\{\{1,5\},\{2,8\},\{3,4\}\}$ is such a subset (if $n \geq 8$), but $\{\{1,5\},\{2,5\},\{3,4\}\}$ is not (since $\{1,5\}$ and $\{2,5\}$ are not disjoint). If you have seen graph theory, you will recognize that such subsets are the *matchings* of the complete graph K_n. (This is also the reason why we called our edges "edges".)

[8] A *fixed point* of a map $f : X \to X$ (where X is any set) means an element $x \in X$ satisfying $f(x) = x$.

Observation 2: Let $k \in \mathbb{N}$. Then, the number of k-element subsets I of R such that all edges in I are disjoint is $\dbinom{n}{2k} \cdot \prod\limits_{j=1}^{k} (2j-1)$.

[*Proof of Observation 2:* We first introduce a notation: If I is a subset of R such that all edges in I are disjoint, and if x is an element of one of the edges in I, then the *I-partner* of x shall mean the other element of this edge (i.e., the element of e that is distinct from x, where e is the unique edge in I that contains x). This I-partner is uniquely determined because there is only one edge in I that contains x (since all edges in I are disjoint).

It is clear that if I is a k-element subset of R such that all edges in I are disjoint, then the union of the edges in I must have size $2k$; in other words, altogether $2k$ elements of $[n]$ belong to the edges of I. Thus, the following is a way to construct any k-element subset I of R such that all edges in I are disjoint:

- First, choose which $2k$ elements of $[n]$ should belong to the edges of I. This choice can be made in $\dbinom{n}{2k}$ ways.
- Having chosen these $2k$ elements, we let Z denote their set. We must now choose I. This I has to be a decomposition of the set Z into k disjoint edges (i.e., two-element subsets). We choose I via the following k-step process:
  - In step 1, we choose the I-partner of the smallest element of Z. There are $2k-1$ choices for it (since any element of Z other than the smallest element is fine). We then remove both the smallest element of Z and its I-partner from Z; thus, Z becomes a $2k-2$-element set.
  - In step 2, we choose the I-partner of the smallest element of Z (keeping in mind that Z is now a $2k-2$-element set). There are $2k-3$ choices for it (since any element of Z other than the smallest element is fine). We then remove both the smallest element of Z and its I-partner from Z; thus, Z becomes a $2k-4$-element set.
  - In step 3, we choose the I-partner of the smallest element of Z (keeping in mind that Z is now a $2k-4$-element set). There are $2k-5$ choices for it (since any element of Z other than the smallest element is fine). We then remove both the smallest element of Z and its I-partner from Z; thus, Z becomes a $2k-6$-element set.
  - And so on, until in step k the set Z has become empty.

  Altogether, we thus have $(2k-1)(2k-3)\cdots 1 = \prod\limits_{j=1}^{k}(2j-1)$ options in this process.

The total number of possibilities how this construction can be made is $\dbinom{n}{2k} \cdot \prod\limits_{j=1}^{k} (2j-1)$. This yields Observation 2.]

Next, we ask ourselves the following question: Given a subset I of R, how many maps $f \in S$ have the property that f coscreams at each $r \in I$ (and perhaps at some other edges as well – we are neither forbidding nor requiring this)? It turns out that the answer to this question depends on whether all edges in I are disjoint or not. We shall consider these two cases in Observations 3 and 4.

> *Observation 3:* Let I be a subset of R. Assume that not all edges in I are disjoint. Then,
>
> $$\begin{aligned}&(\text{the number of all } f \in S \text{ such that } f \text{ coscreams at each } r \in I)\\&= 0.\end{aligned}$$

[*Proof of Observation 3:* We must simply prove that there exists no $f \in S$ such that f coscreams at each $r \in I$.

Indeed, consider such an f. We shall derive a contradiction.

Not all edges in I are disjoint. Thus, there exist two edges in I having the forms $\{a, b\}$ and $\{a, c\}$ for distinct $a, b, c \in [n]$. Consider two such edges. The map f coscreams at the edge $\{a, b\}$ (since f coscreams at each $r \in I$). In other words, $f(a) = b$ and $f(b) = a$. Similarly, $f(a) = c$ and $f(c) = a$. But $b = f(a) = c$ contradicts the distinctness of a, b, c.

Now, forget that we fixed f. We thus have found a contradiction for each $f \in S$ such that f coscreams at each $r \in I$. Hence, there exist no such f. This proves Observation 3.]

> *Observation 4:* Let I be a subset of R. Assume that all edges in I are disjoint. Then,
>
> $$\begin{aligned}&(\text{the number of all } f \in S \text{ such that } f \text{ coscreams at each } r \in I)\\&= (n-1)^{n-2|I|}.\end{aligned}$$

[*Proof of Observation 4:* How do we construct a map $f \in S$ such that f coscreams at each $r \in I$? Clearly, if a map $f : [n] \to [n]$ coscreams at an edge r, then this uniquely determines the values of f on both elements of this edge r (namely, f has to send the first element to the second and vice versa). Thus, if we want to construct a map $f \in S$ such that f coscreams at each $r \in I$, then we immediately know the values of f on each element of each of the edges $r \in I$. These are $2|I|$ elements in total (since all edges in I are disjoint, and each of them has 2 elements). Hence, in order to construct a map $f \in S$ such that f coscreams at each $r \in I$, we only need to choose the values of f on the remaining $n - 2|I|$ elements of $[n]$. There are $(n-1)^{n-2|I|}$ options to do that, because for each of the remaining $n - 2|I|$ elements of $[n]$ we can choose the value of f at this element in exactly $n-1$ ways[9]. Hence, the number of all $f \in S$ such that f coscreams at each $r \in I$ is $(n-1)^{n-2|I|}$. This proves Observation 4.]

For each $r \in R$, let A_r be the set of all maps $f \in S$ that coscream at the edge r. Clearly, A_r is a subset of S for each $r \in R$.

Now, we are ready for the grand computation. From Observation 1, we conclude

[9] Why $n-1$ and not n ? Well, we want f to belong to S, so we want f to have no fixed points. Thus, f cannot send our element to itself.

that

$$
\begin{aligned}
&(\text{the number of all silent maps } f : [n] \to [n]) \\
&= \left(\text{the number of all maps } f : [n] \to [n] \text{ such that } \underbrace{f \text{ has no fixed point}}_{\substack{\iff (f \in S) \\ \text{(by the definition of } S\text{)}}} \right. \\
&\qquad \left. \text{and} \qquad \underbrace{f \text{ coscreams at no edges}}_{\iff ((f \text{ does not coscream at } r) \text{ for each } r \in R)} \right) \\
&= (\text{the number of all maps } f : [n] \to [n] \text{ such that } f \in S \\
&\qquad \text{and } \ ((f \text{ does not coscream at } r) \text{ for each } r \in R)) \\
&= \left(\text{the number of all } f \in S \text{ such that } \left(\underbrace{(f \text{ does not coscream at } r)}_{\substack{\iff (f \notin A_r) \\ \text{(by the definition of } A_r\text{)}}} \text{ for each } r \in R \right) \right) \\
&= \left(\text{the number of all } f \in S \text{ such that } \underbrace{(f \notin A_r \text{ for each } r \in R)}_{\iff \left(f \notin \bigcup_{r \in R} A_r \right)} \right) \\
&= \left(\text{the number of all } f \in S \text{ such that } f \notin \bigcup_{r \in R} A_r \right) \\
&= \left| S \setminus \bigcup_{r \in R} A_r \right| = \sum_{I \subseteq R} (-1)^{|I|} \underbrace{\left| \bigcap_{r \in I} A_r \right|}_{= \left(\text{the number of all } f \in S \text{ such that } f \in \bigcap_{r \in I} A_r \right)} \\
&\qquad (\text{by Theorem 0.6 } \textbf{(b)}) \\
&= \sum_{I \subseteq R} (-1)^{|I|} \left(\text{the number of all } f \in S \text{ such that } \underbrace{f \in \bigcap_{r \in I} A_r}_{\iff (f \in A_r \text{ for all } r \in I)} \right)
\end{aligned}
$$

$$= \sum_{I \subseteq R} (-1)^{|I|} \left(\text{the number of all } f \in S \text{ such that } \left(\underbrace{f \in A_r}_{\substack{\iff (f \text{ coscreams at } r) \\ \text{(by the definition of } A_r)}} \quad \text{for all } r \in I \right) \right)$$

$$= \sum_{I \subseteq R} (-1)^{|I|} \left(\text{the number of all } f \in S \text{ such that } \underbrace{(f \text{ coscreams at } r \text{ for all } r \in I)}_{\iff (f \text{ coscreams at each } r \in I)} \right)$$

$$= \sum_{I \subseteq R} (-1)^{|I|} (\text{the number of all } f \in S \text{ such that } f \text{ coscreams at each } r \in I)$$

$$= \sum_{\substack{I \subseteq R; \\ \text{not all edges in } I \\ \text{are disjoint}}} (-1)^{|I|} \underbrace{(\text{the number of all } f \in S \text{ such that } f \text{ coscreams at each } r \in I)}_{\substack{=0 \\ \text{(by Observation 3)}}}$$

$$+ \sum_{\substack{I \subseteq R; \\ \text{all edges in } I \\ \text{are disjoint}}} (-1)^{|I|} \underbrace{(\text{the number of all } f \in S \text{ such that } f \text{ coscreams at each } r \in I)}_{\substack{=(n-1)^{n-2|I|} \\ \text{(by Observation 4)}}}$$

$$= \underbrace{\sum_{\substack{I \subseteq R; \\ \text{not all edges in } I \\ \text{are disjoint}}} (-1)^{|I|} 0}_{=0} + \sum_{\substack{I \subseteq R; \\ \text{all edges in } I \\ \text{are disjoint}}} (-1)^{|I|} (n-1)^{n-2|I|}$$

$$= \sum_{\substack{I \subseteq R; \\ \text{all edges in } I \\ \text{are disjoint}}} (-1)^{|I|} (n-1)^{n-2|I|} = \sum_{k \in \mathbb{N}} \sum_{\substack{I \subseteq R; \\ \text{all edges in } I \\ \text{are disjoint;} \\ |I|=k}} \underbrace{(-1)^{|I|} (n-1)^{n-2|I|}}_{\substack{=(-1)^k (n-1)^{n-2k} \\ \text{(since } |I|=k)}}$$

$$= \sum_{k \in \mathbb{N}} \underbrace{\sum_{\substack{I \subseteq R; \\ \text{all edges in } I \\ \text{are disjoint;} \\ |I|=k}} (-1)^k (n-1)^{n-2k}}_{=(\text{the number of all } I \subseteq R \text{ such that all edges in } I \text{ are disjoint, and } |I|=k) \cdot (-1)^k (n-1)^{n-2k}}$$

$$
\begin{aligned}
&= \sum_{k\in\mathbb{N}} \underbrace{(\text{the number of all } I \subseteq R \text{ such that all edges in } I \text{ are disjoint, and } |I| = k)}_{\substack{=(\text{the number of } k\text{-element subsets } I \text{ of } R \text{ such that all edges in } I \text{ are disjoint}) \\ = \binom{n}{2k} \cdot \prod_{j=1}^{k} (2j-1) \\ (\text{by Observation 2})}} \\
&\qquad \cdot (-1)^k (n-1)^{n-2k} \\
&= \sum_{k\in\mathbb{N}} \underbrace{\binom{n}{2k}}_{\substack{= \frac{n(n-1)\cdots(n-2k+1)}{(2k)!} \\ \text{(by the definition of} \\ \text{binomial coefficients)}}} \cdot \underbrace{\left(\prod_{j=1}^{k} (2j-1)\right)}_{\substack{=(2k-1)\cdot(2k-3)\cdot\cdots\cdot 1 \\ = \frac{(2k)!}{2^k k!} \\ \text{(by Proposition 0.7,} \\ \text{applied to } k \text{ instead of } n)}} \cdot (-1)^k (n-1)^{n-2k} \\
&= \sum_{k\in\mathbb{N}} \underbrace{\frac{n(n-1)\cdots(n-2k+1)}{(2k)!} \cdot \frac{(2k)!}{2^k k!} \cdot (-1)^k (n-1)^{n-2k}}_{=(-1)^k \frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!} (n-1)^{n-2k}} \\
&= \sum_{k\in\mathbb{N}} (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!} (n-1)^{n-2k} \\
&= \underbrace{\sum_{\substack{k\in\mathbb{N}; \\ k\leq n}}}_{=\sum_{k=0}^{n}} (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!} (n-1)^{n-2k} \\
&\qquad + \sum_{\substack{k\in\mathbb{N}; \\ k>n}} (-1)^k \underbrace{\frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!}}_{\substack{=0 \\ (\text{since the product } n(n-1)\cdots(n-2k+1) \text{ contains} \\ \text{the factor } n-n \text{ (because } n<k\leq 2k\text{) and thus} \\ \text{equals } 0 \text{ (since } n-n=0))}} (n-1)^{n-2k} \\
&\qquad (\text{since each } k \in \mathbb{N} \text{ satisfies either } k \leq n \text{ or } k > n \text{ (but not both)}) \\
&= \sum_{k=0}^{n} (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!} (n-1)^{n-2k} + \underbrace{\sum_{\substack{k\in\mathbb{N}; \\ k>n}} (-1)^k \, 0 \, (n-1)^{n-2k}}_{=0} \\
&= \sum_{k=0}^{n} (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{2^k \cdot k!} (n-1)^{n-2k} .
\end{aligned}
$$

This proves the "Combinatorial restatement". To prove the first claim, about the probability that no one screams, we have to divide this number by $(n-1)^n$, because $(n-1)^n$ is the total number of ways that everyone can look down at someone else's feet. Thus, Exercise 3 is solved. □

0.4. An alternating identity

Exercise 4. Let i and j be positive integers. Prove that

$$\sum_{k=\max\{i,j\}}^{i+j} (-1)^k \frac{(k-1)!}{(k-i)!\,(k-j)!\,(i+j-k)!} = 0.$$

I have learnt Exercise 4 from a post by Peter Scholze on Art of Problem Solving [Scholz04] (who stated it in the case when $i \geq j$ only, but the general case can be easily reduced to this). Three solutions have been suggested on this thread; feel free to add yours!

Before we solve Exercise 4, let us recall some more properties of binomial coefficients:

Proposition 0.8. Let a and b be two integers such that $a \geq b \geq 0$. Then,

$$\binom{a}{b} = \frac{a!}{b!\,(a-b)!}.$$

Proposition 0.9. For every $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$ and $n \in \mathbb{N}$, we have

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

Proposition 0.10. We have

$$\binom{n}{k} = (-1)^k \binom{k-n-1}{k}$$

for any $n \in \mathbb{Q}$ and $k \in \mathbb{N}$.

Proposition 0.11. We have

$$\binom{m}{n} = 0$$

for every $m \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfying $m < n$.

Proposition 0.8 was proven in the solutions to homework set 1. Proposition 0.9 is the *Vandermonde convolution identity*, and is proven in multiple places[10]. Proposition 0.10 is Exercise 2 **(a)** in homework set 1 (and also appears as [Grinbe16b,

[10] For an elementary proof, see, e.g., [Grinbe16b, first proof of Theorem 3.29].

Proposition 3.16]). Proposition 0.11 is fundamental and easy to prove (see, e.g., [Grinbe16b, Proposition 3.6]).

As a consequence, we obtain the following:

Proposition 0.12. Let i and j be two positive integers. Then,

$$\sum_{k=0}^{j} (-1)^k \binom{i}{j-k} \binom{k+i-1}{k} = 0.$$

Proof of Proposition 0.12. We have $0 < j$ (since j is positive). Hence, Proposition 0.11 (applied to $m = 0$ and $n = j$) yields $\binom{0}{j} = 0$.

Proposition 0.9 (applied to $x = -i$, $y = i$ and $n = j$) yields

$$\binom{(-i)+i}{j} = \sum_{k=0}^{j} \underbrace{\binom{-i}{k}}_{\substack{=(-1)^k \binom{k-(-i)-1}{k} \\ \text{(by Proposition 0.10,} \\ \text{applied to } n=-i)}} \binom{i}{j-k} = \sum_{k=0}^{j} (-1)^k \underbrace{\binom{k-(-i)-1}{k}}_{\substack{=\binom{k+i-1}{k} \\ \text{(since } k-(-i)-1=k+i-1)}} \binom{i}{j-k}$$
$$= \sum_{k=0}^{j} (-1)^k \binom{k+i-1}{k} \binom{i}{j-k} = \sum_{k=0}^{j} (-1)^k \binom{i}{j-k} \binom{k+i-1}{k}.$$

Thus,

$$\sum_{k=0}^{j} (-1)^k \binom{i}{j-k} \binom{k+i-1}{k} = \binom{(-i)+i}{j} = \binom{0}{j} = 0.$$

This proves Proposition 0.12. □

Let us now solve Exercise 4:

Solution to Exercise 4. We can WLOG assume that $i \geq j$ (since otherwise, we can simply interchange with i and j). Assume this.

Also, $i! = i \cdot (i-1)!$ (since i is a positive integer).

For every $k \in \{i, i+1, \ldots, i+j\}$, we have

$$\frac{(k-1)!}{(k-i)!\,(k-j)!\,(i+j-k)!} = \frac{1}{i} \binom{i}{i+j-k} \binom{k-1}{k-i}. \tag{10}$$

[*Proof of (10):* Let $k \in \{i, i+1, \ldots, i+j\}$. Then, $i \leq k \leq i+j$. From $k \leq i+j$, we obtain $i+j-k \geq 0$. Also, from $i \leq k$, we obtain $k \geq i \geq j$. Hence, $i + j - \underbrace{k}_{\geq j} \leq$

$i+j-j=i$. Therefore, $i \geq i+j-k \geq 0$. Thus, Proposition 0.8 (applied to $a=i$ and $b=i+j-k$) yields

$$\begin{aligned}\binom{i}{i+j-k} &= \frac{i!}{(i+j-k)!\,(i-(i+j-k))!} = \frac{i!}{(i+j-k)!\,(k-j)!} \\ &\qquad (\text{since } i-(i+j-k)=k-j) \\ &= \frac{i\cdot(i-1)!}{(i+j-k)!\,(k-j)!} \qquad (\text{since } i! = i\cdot(i-1)!) . \end{aligned} \tag{11}$$

On the other hand, $k \geq i$, so that $k-i \geq 0$. Also, $i \geq 1$ (since i is a positive integer), so that $k-1 \geq k-i \geq 0$. Thus, Proposition 0.8 (applied to $a=k-1$ and $b=k-i$) yields

$$\binom{k-1}{k-i} = \frac{(k-1)!}{(k-i)!\,((k-1)-(k-i))!} = \frac{(k-1)!}{(k-i)!\,(i-1)!} \tag{12}$$

(since $(k-1)-(k-i)=i-1$). Multiplying the equalities (11) and (12), we find

$$\begin{aligned}\binom{i}{i+j-k}\binom{k-1}{k-i} &= \frac{i\cdot(i-1)!}{(i+j-k)!\,(k-j)!}\cdot\frac{(k-1)!}{(k-i)!\,(i-1)!} \\ &= i\cdot\frac{(k-1)!}{(k-i)!\,(k-j)!\,(i+j-k)!}.\end{aligned}$$

Dividing this equality by i, we find

$$\frac{1}{i}\binom{i}{i+j-k}\binom{k-1}{k-i} = \frac{(k-1)!}{(k-i)!\,(k-j)!\,(i+j-k)!}.$$

This proves (10).]

But $\max\{i,j\} = i$ (since $i \geq j$). Hence,

$$\begin{aligned}
&\sum_{k=\max\{i,j\}}^{i+j} (-1)^k \frac{(k-1)!}{(k-i)!\,(k-j)!\,(i+j-k)!} \\
&= \sum_{k=i}^{i+j} (-1)^k \underbrace{\frac{(k-1)!}{(k-i)!\,(k-j)!\,(i+j-k)!}}_{\substack{=\frac{1}{i}\binom{i}{i+j-k}\binom{k-1}{k-i} \\ \text{(by (10))}}} \\
&= \sum_{k=i}^{i+j} (-1)^k \frac{1}{i}\binom{i}{i+j-k}\binom{k-1}{k-i} = \frac{1}{i}\sum_{k=i}^{i+j} (-1)^k \binom{i}{i+j-k}\binom{k-1}{k-i} \\
&= \frac{1}{i}\sum_{k=0}^{j} \underbrace{(-1)^{k+i}}_{=(-1)^k(-1)^i} \underbrace{\binom{i}{i+j-(k+i)}}_{\substack{=\binom{i}{j-k} \\ \text{(since } i+j-(k+i)=j-k\text{)}}} \underbrace{\binom{k+i-1}{k+i-i}}_{\substack{=\binom{k+i-1}{k} \\ \text{(since } k+i-i=k\text{)}}} \\
&\qquad \text{(here, we have substituted } k+i \text{ for } k \text{ in the sum)} \\
&= \frac{1}{i}\sum_{k=0}^{j} (-1)^k (-1)^i \binom{i}{j-k}\binom{k+i-1}{k} \\
&= \frac{1}{i}(-1)^i \underbrace{\sum_{k=0}^{j} (-1)^k \binom{i}{j-k}\binom{k+i-1}{k}}_{\substack{=0 \\ \text{(by Proposition 0.12)}}} = \frac{1}{i}(-1)^i\, 0 = 0.
\end{aligned}$$

This solves Exercise 4. □

0.5. The binomial transform

Exercise 5. Let $N \in \mathbb{N}$. Let $(a_0, a_1, \ldots, a_N)$ be a list of rational numbers. Define a second list $(b_0, b_1, \ldots, b_N)$ of rational numbers by setting

$$b_n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} a_i \qquad \text{for each } n \in \{0, 1, \ldots, N\}.$$

Prove that

$$a_n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} b_i \qquad \text{for each } n \in \{0, 1, \ldots, N\}.$$

Before we solve this exercise, let us recall two more basic facts:

Proposition 0.13. If $n \in \mathbb{Q}$ and if a and b are two integers such that $a \geq b \geq 0$, then

$$\binom{n}{a}\binom{a}{b} = \binom{n}{b}\binom{n-b}{a-b}.$$

Proposition 0.14. We have $\binom{m}{m} = 1$ for every $m \in \mathbb{N}$.

Proposition 0.13 is Exercise 2 **(c)** in homework set 1. Proposition 0.14 is easy to check.

On the other hand, an easy consequence of the binomial formula is the following fact:

Proposition 0.15. Let $m \in \mathbb{N}$. Then,

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

Proof of Proposition 0.15. Proposition 0.3 (applied to $n = m$, $x = -1$ and $y = 1$) yields

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

Hence,

$$\begin{aligned}
\sum_{k=0}^{m} (-1)^k \binom{m}{k} &= \Bigg(\underbrace{(-1)+1}_{=0}\Bigg)^m = 0^m = \begin{cases} 1, & \text{if } m = 0; \\ 0, & \text{if } m > 0 \end{cases} \\
&= \begin{cases} 1, & \text{if } m = 0; \\ 0, & \text{if } m \neq 0 \end{cases} \qquad \left(\begin{array}{c}\text{since the condition } (m > 0) \text{ is} \\ \text{equivalent to } (m \neq 0) \text{ (because } m \in \mathbb{N}\text{)}\end{array}\right) \\
&= [m = 0].
\end{aligned}$$

This proves Proposition 0.15. □

Corollary 0.16. Let $n \in \mathbb{N}$. Let $i \in \{0, 1, \ldots, n\}$. Then,

$$\sum_{j=i}^{n} (-1)^{j+i} \binom{n}{j} \binom{j}{i} = [i = n].$$

Proof of Corollary 0.16. We have $i \in \{0, 1, \ldots, n\}$, so that $i \leq n$ and $i \in \mathbb{N}$. From $i \leq n$, we obtain $n - i \geq 0$. Thus, $n - i \in \mathbb{N}$.

Let $j \in \{i, i+1, \ldots, n\}$. Then, $i \leq j$. Hence, $j \geq i \geq 0$. Therefore, Proposition 0.13 (applied to $a = j$ and $b = i$) yields

$$\binom{n}{j}\binom{j}{i} = \binom{n}{i}\binom{n-i}{j-i}. \tag{13}$$

Also, $j + i \equiv j - i \mod 2$ (since $(j+i) - (j-i) = 2i$ is even). Thus, $(-1)^{j+i} = (-1)^{j-i}$. Multiplying this equality by (13), we obtain

$$(-1)^{j+i}\binom{n}{j}\binom{j}{i} = (-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i}. \tag{14}$$

Now, forget that we fixed j. We thus have proven (14) for each $j \in \{i, i+1, \ldots, n\}$. Hence,

$$\begin{aligned}
&\sum_{j=i}^{n} \underbrace{(-1)^{j+i}\binom{n}{j}\binom{j}{i}}_{\substack{=(-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i} \\ \text{(by (14))}}} \\
&= \sum_{j=i}^{n} (-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i} = \sum_{k=0}^{n-i} (-1)^{k}\binom{n}{i}\binom{n-i}{k} \\
&\qquad \text{(here, we have substituted } k \text{ for } j-i \text{ in the sum)} \\
&= \binom{n}{i} \underbrace{\sum_{k=0}^{n-i} (-1)^{k}\binom{n-i}{k}}_{\substack{=[n-i=0] \\ \text{(by Proposition 0.15 (applied to } m=n-i\text{))}}} \\
&= \binom{n}{i}[n-i=0].
\end{aligned} \tag{15}$$

But it is easy to see that

$$\binom{n}{i}[n-i=0] = [i=n] \tag{16}$$

[11]. Thus, (15) becomes

$$\sum_{j=i}^{n} (-1)^{j+i}\binom{n}{j}\binom{j}{i} = \binom{n}{i}[n-i=0] = [i=n].$$

This proves Corollary 0.16. □

[11] *Proof of (16):* We are in one of the following two cases:

Case 1: We have $i \neq n$.

Case 2: We have $i = n$.

Let us consider Case 1 first. In this case, we have $i \neq n$. In other words, $n \neq i$. Hence, *Solution to Exercise 5.* We have assumed that

$$b_n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} a_i \tag{17}$$

for each $n \in \{0, 1, \ldots, N\}$.

$n - i \neq 0$. Thus, $[n - i = 0] = 0$, so that $\binom{n}{i} \underbrace{[n - i = 0]}_{=0} = 0$. Comparing this with $[i = n] = 0$ (since $i \neq n$), we obtain $\binom{n}{i} [n - i = 0] = [i = n]$. Hence, (16) is proven in Case 1.

Now, let us consider Case 2. In this case, we have $i = n$. In other words, $n = i$. Thus, $n - i = 0$. Thus, $[n - i = 0] = 1$. Also, from $n = i$, we obtain $\binom{n}{i} = \binom{i}{i} = 1$ (by Proposition 0.14, applied to $m = i$). Hence, $\underbrace{\binom{n}{i}}_{=1} \underbrace{[n - i = 0]}_{=1} = 1$. Comparing this with $[i = n] = 1$ (since $i = n$), we obtain $\binom{n}{i} [n - i = 0] = [i = n]$. Hence, (16) is proven in Case 2.

We thus have proven (16) in each of the two Cases 1 and 2. Thus, (16) always holds.

Now, let $n \in \{0, 1, \ldots, N\}$. Then,

$$\begin{aligned}
\sum_{i=0}^{n} (-1)^i \binom{n}{i} b_i &= \sum_{j=0}^{n} (-1)^j \binom{n}{j} \underbrace{b_j}_{\substack{=\sum_{i=0}^{j} (-1)^i \binom{j}{i} a_i \\ \text{(by (17), applied to } j \\ \text{instead of } n)}} \qquad \left(\begin{array}{c}\text{here, we have renamed the} \\ \text{summation index } i \text{ as } j\end{array}\right) \\
&= \sum_{j=0}^{n} (-1)^j \binom{n}{j} \sum_{i=0}^{j} (-1)^i \binom{j}{i} a_i \\
&= \underbrace{\sum_{j=0}^{n}}_{=\sum_{j\in\{0,1,\ldots,n\}}} \underbrace{\sum_{i=0}^{j}}_{\substack{=\sum_{i\in\{0,1,\ldots,j\}} \\ =\sum_{i\in\{0,1,\ldots,n\};\ i\leq j} \\ \text{(since the elements of } \{0,1,\ldots,j\} \\ \text{are precisely the elements } i\in\{0,1,\ldots,n\} \\ \text{satisfying } i\leq j \text{ (because } j\leq n))}} \underbrace{(-1)^j \binom{n}{j} (-1)^i}_{=(-1)^i \binom{n}{j}} \binom{j}{i} a_i \\
&= \underbrace{\sum_{j\in\{0,1,\ldots,n\}} \sum_{\substack{i\in\{0,1,\ldots,n\};\\ i\leq j}}}_{=\sum_{i\in\{0,1,\ldots,n\}} \sum_{j\in\{0,1,\ldots,n\};\ i\leq j}} \underbrace{(-1)^j (-1)^i}_{=(-1)^{j+i}} \binom{n}{j} \binom{j}{i} a_i \\
&= \sum_{i\in\{0,1,\ldots,n\}} \underbrace{\sum_{\substack{j\in\{0,1,\ldots,n\};\\ i\leq j}}}_{=\sum_{j=i}^{n}} (-1)^{j+i} \binom{n}{j} \binom{j}{i} a_i = \sum_{i\in\{0,1,\ldots,n\}} \sum_{j=i}^{n} (-1)^{j+i} \binom{n}{j} \binom{j}{i} a_i \\
&= \sum_{i\in\{0,1,\ldots,n\}} \underbrace{\left(\sum_{j=i}^{n} (-1)^{j+i} \binom{n}{j} \binom{j}{i} \right)}_{\substack{=[i=n] \\ \text{(by Corollary 0.16)}}} a_i \\
&= \sum_{i\in\{0,1,\ldots,n\}} [i=n]\, a_i = \underbrace{[n=n]}_{\substack{=1 \\ \text{(since } n=n)}} a_n + \sum_{\substack{i\in\{0,1,\ldots,n\};\\ i\neq n}} \underbrace{[i=n]}_{\substack{=0 \\ \text{(since } i\neq n)}} a_i \\
&\qquad \left(\begin{array}{c}\text{here, we have split off the addend for } i=n \\ \text{from the sum}\end{array}\right) \\
&= a_n + \underbrace{\sum_{\substack{i\in\{0,1,\ldots,n\};\\ i\neq n}} 0 a_i}_{=0} = a_n.
\end{aligned}$$

In other words, $a_n = \sum\limits_{i=0}^{n} (-1)^i \dbinom{n}{i} b_i$. This solves Exercise 5. □

0.6. The Leibniz identity for the difference operator

Now, recall some of the notations for finite differences:

Per se, the words "map", "mapping", "function", "transformation" and "operator" are synonyms in mathematics (they all mean assignments of values from one set to the elements of another set). But it is common to use some of these words selectively for certain kinds of maps. We shall follow the following rules:

- The word "map" can mean any kind of map.
- The word "function" shall mean a map from $\mathbb{Q}$ to $\mathbb{Q}$. Thus, the set of all functions is $\mathbb{Q}^{\mathbb{Q}}$.
- The word "operator" shall mean a map from $\mathbb{Q}^{\mathbb{Q}}$ to $\mathbb{Q}^{\mathbb{Q}}$. Thus, an operator is a map sending functions to functions.

For example, the map

$$\mathbb{Q} \to \mathbb{Q}, \qquad x \mapsto x^2$$

is a function, whereas the map

$$\mathbb{Q}^{\mathbb{Q}} \to \mathbb{Q}^{\mathbb{Q}}, \qquad f \mapsto f \circ f$$

("apply a function twice") is an operator.

If f and g are functions, then $f + g$ denotes the pointwise sum of f and g (that is, the function $\mathbb{Q} \to \mathbb{Q}$, $x \mapsto f(x) + g(x)$), and fg denotes the pointwise product of f and g (that is, the function $\mathbb{Q} \to \mathbb{Q}$, $x \mapsto f(x) g(x)$). We can also write $f \cdot g$ for fg.

If f is a function and $\lambda \in \mathbb{Q}$, then λf denotes the pointwise product of λ with f (that is, the function $\mathbb{Q} \to \mathbb{Q}$, $x \mapsto \lambda f(x)$). Thus, the functions form a $\mathbb{Q}$-vector space (and better yet, a commutative $\mathbb{Q}$-algebra, because of the multiplication).

The following three operators are particularly important:

- The identity operator $\operatorname{id} : \mathbb{Q}^{\mathbb{Q}} \to \mathbb{Q}^{\mathbb{Q}}$. It sends each function f to f itself.
- The shift operator $S : \mathbb{Q}^{\mathbb{Q}} \to \mathbb{Q}^{\mathbb{Q}}$. It sends each function f to the function $S(f)$ defined by $(S(f))(x) = f(x+1)$ for all $x \in \mathbb{Q}$. Speaking in terms of function plots, the operator S shifts a function by 1 to the left.
- The difference operator $\Delta : \mathbb{Q}^{\mathbb{Q}} \to \mathbb{Q}^{\mathbb{Q}}$. It sends each function f to the function $\Delta(f)$ defined by $(\Delta(f))(x) = f(x+1) - f(x)$ for all $x \in \mathbb{Q}$. Speaking in terms of function plots, the operator Δ shifts a function by 1 to the left and subtracts the original function back from it. Note that $\Delta(f) = S(f) - f$ for each $f \in \mathbb{Q}^{\mathbb{Q}}$.

Here are some examples of what the difference operator Δ does to certain functions[12]:

$$\Delta\left(x\mapsto 1\right)=\left(x\mapsto 1-1\right)=\left(x\mapsto 0\right);$$
$$\Delta\left(x\mapsto x\right)=\left(x\mapsto \left(x+1\right)-x\right)=\left(x\mapsto 1\right);$$
$$\Delta\left(x\mapsto x^{2}\right)=\left(x\mapsto \left(x+1\right)^{2}-x^{2}\right)=\left(x\mapsto 2x+1\right);$$
$$\Delta\left(x\mapsto \binom{x}{2}\right)=\left(x\mapsto \binom{x+1}{2}-\binom{x}{2}\right)=\left(x\mapsto x\right);$$
$$\Delta\left(x\mapsto 2^{x}\right)=\left(x\mapsto 2^{x+1}-2^{x}\right)=\left(x\mapsto 2^{x}\right).$$

(The last example is not completely kosher, since the function $\mathbb{Q}\to\mathbb{Q},\ x\mapsto 2^{x}$ does not actually exist: 2^{x} isn't always rational. To make sense of it, imagine that we are talking about maps from $\mathbb{Z}$ to $\mathbb{Q}$, or from $\mathbb{R}$ to $\mathbb{R}$, instead.) Actually, for every $n\in\mathbb{N}$, we have

$$\Delta\left(x\mapsto x^{n}\right)=\left(x\mapsto \left(x+1\right)^{n}-x^{n}\right)=\left(x\mapsto \sum_{k=0}^{n-1}\binom{n}{k}x^{k}\right)$$

and

$$\Delta\left(x\mapsto \binom{x}{n}\right)=\left(x\mapsto \binom{x+1}{n}-\binom{x}{n}\right)=\left(x\mapsto \binom{x}{n-1}\right).$$

If our functions were C^{∞}-functions $\mathbb{R}\to\mathbb{R}$ instead of maps $\mathbb{Q}\to\mathbb{Q}$, then $\dfrac{d}{dx}$ would be another operator (sending each function f to its derivative).

Exercise 6. **(a)** Prove that $S\left(fg\right)=S\left(f\right)\cdot S\left(g\right)$ for any two functions f and g.
(b) Prove that $\Delta\left(fg\right)=S\left(f\right)\Delta\left(g\right)+\Delta\left(f\right)g$ for any two functions f and g.
(c) Prove that $\Delta\left(fg\right)=f\Delta\left(g\right)+\Delta\left(f\right)S\left(g\right)$ for any two functions f and g.
(d) Prove that $\Delta\circ S=S\circ\Delta$.

Solution to Exercise 6. **(a)** Let f and g be two functions. Let $x\in\mathbb{Q}$. Then,

$$\begin{aligned}\left(S\left(fg\right)\right)\left(x\right)&=\left(fg\right)\left(x+1\right)\qquad\text{(by the definition of }S\text{)}\\&=f\left(x+1\right)\cdot g\left(x+1\right)\qquad\text{(by the definition of }fg\text{)}\end{aligned}$$

and

$$\begin{aligned}\left(S\left(f\right)\cdot S\left(g\right)\right)\left(x\right)&=\underbrace{\left(S\left(f\right)\right)\left(x\right)}_{\substack{=f\left(x+1\right)\\\text{(by the definition of }S\text{)}}}\cdot\underbrace{\left(S\left(g\right)\right)\left(x\right)}_{\substack{=g\left(x+1\right)\\\text{(by the definition of }S\text{)}}}\\&\qquad\text{(by the definition of }S\left(f\right)\cdot S\left(g\right)\text{)}\\&=f\left(x+1\right)\cdot g\left(x+1\right).\end{aligned}$$

[12] We use the shorthand notation "$\left(x\mapsto x^{2}\right)$" for the function $\mathbb{Q}\to\mathbb{Q},\ x\mapsto x^{2}$, because all of our functions are from $\mathbb{Q}$ to $\mathbb{Q}^{2}$ anyway.

Comparing these two equalities yields $(S(fg))(x) = (S(f) \cdot S(g))(x)$.

Now, forget that we fixed x. We thus have shown that $(S(fg))(x) = (S(f) \cdot S(g))(x)$ for each $x \in \mathbb{Q}$. In other words, $S(fg) = S(f) \cdot S(g)$ (since both $S(fg)$ and $S(f) \cdot S(g)$ are functions from $\mathbb{Q}$ to $\mathbb{Q}$). This solves Exercise 6 **(a)**.

(b) Let f and g be two functions. Let $x \in \mathbb{Q}$. Then,

$$\begin{aligned}
(\Delta(fg))(x) &= \underbrace{(fg)(x+1)}_{\substack{=f(x+1)\cdot g(x+1) \\ \text{(by the definition of } fg\text{)}}} - \underbrace{(fg)(x)}_{\substack{=f(x)\cdot g(x) \\ \text{(by the definition of } fg\text{)}}} \qquad \text{(by the definition of } \Delta\text{)} \\
&= f(x+1) \cdot g(x+1) - f(x) \cdot g(x)
\end{aligned}$$

and

$$\begin{aligned}
&(S(f)\Delta(g) + \Delta(f)g)(x) \\
&= \underbrace{(S(f)\Delta(g))(x)}_{\substack{=(S(f))(x)\cdot(\Delta(g))(x) \\ \text{(by the definition of } S(f)\Delta(g)\text{)}}} + \underbrace{(\Delta(f)g)(x)}_{\substack{=(\Delta(f))(x)\cdot g(x) \\ \text{(by the definition of } \Delta(f)g\text{)}}} \\
&\qquad \text{(by the definition of } S(f)\Delta(g) + \Delta(f)g\text{)} \\
&= \underbrace{(S(f))(x)}_{\substack{=f(x+1) \\ \text{(by the definition of } S\text{)}}} \cdot \underbrace{(\Delta(g))(x)}_{\substack{=g(x+1)-g(x) \\ \text{(by the definition of } \Delta\text{)}}} + \underbrace{(\Delta(f))(x)}_{\substack{=f(x+1)-f(x) \\ \text{(by the definition of } \Delta\text{)}}} \cdot g(x) \\
&= f(x+1) \cdot (g(x+1) - g(x)) + (f(x+1) - f(x)) \cdot g(x) \\
&= f(x+1) \cdot g(x+1) - f(x+1) \cdot g(x) + f(x+1) \cdot g(x) - f(x) \cdot g(x) \\
&= f(x+1) \cdot g(x+1) - f(x) \cdot g(x).
\end{aligned}$$

Comparing these two equalities yields $(\Delta(fg))(x) = (S(f)\Delta(g) + \Delta(f)g)(x)$.

Now, forget that we fixed x. We thus have shown that $(\Delta(fg))(x) = (S(f)\Delta(g) + \Delta(f)g)(x)$ for each $x \in \mathbb{Q}$. In other words, $\Delta(fg) = S(f)\Delta(g) + \Delta(f)g$. This solves Exercise 6 **(b)**.

(c) Let f and g be two functions. Let $x \in \mathbb{Q}$. Then,

$$\begin{aligned}
(\Delta(fg))(x) &= \underbrace{(fg)(x+1)}_{\substack{=f(x+1)\cdot g(x+1) \\ \text{(by the definition of } fg\text{)}}} - \underbrace{(fg)(x)}_{\substack{=f(x)\cdot g(x) \\ \text{(by the definition of } fg\text{)}}} \qquad \text{(by the definition of } \Delta\text{)} \\
&= f(x+1) \cdot g(x+1) - f(x) \cdot g(x)
\end{aligned}$$

and

$$\begin{aligned}
&(f\Delta(g)+\Delta(f)S(g))(x) \\
&= \underbrace{(f\Delta(g))(x)}_{\substack{=f(x)\cdot(\Delta(g))(x) \\ \text{(by the definition of } f\Delta(g))}} + \underbrace{(\Delta(f)S(g))(x)}_{\substack{=(\Delta(f))(x)\cdot(S(g))(x) \\ \text{(by the definition of } \Delta(f)S(g))}} \\
&\qquad \text{(by the definition of } f\Delta(g)+\Delta(f)S(g)) \\
&= f(x)\cdot \underbrace{(\Delta(g))(x)}_{\substack{=g(x+1)-g(x) \\ \text{(by the definition of } \Delta)}} + \underbrace{(\Delta(f))(x)}_{\substack{=f(x+1)-f(x) \\ \text{(by the definition of } \Delta)}} \cdot \underbrace{(S(g))(x)}_{\substack{=g(x+1) \\ \text{(by the definition of } S)}} \\
&= f(x)\cdot(g(x+1)-g(x))+(f(x+1)-f(x))\cdot g(x+1) \\
&= f(x)\cdot g(x+1)-f(x)\cdot g(x)+f(x+1)\cdot g(x+1)-f(x)\cdot g(x+1) \\
&= f(x+1)\cdot g(x+1)-f(x)\cdot g(x).
\end{aligned}$$

Comparing these two equalities yields $(\Delta(fg))(x)=(f\Delta(g)+\Delta(f)S(g))(x)$.

Now, forget that we fixed x. We thus have shown that $(\Delta(fg))(x)=(f\Delta(g)+\Delta(f)S(g))(x)$ for each $x\in\mathbb{Q}$. In other words, $\Delta(fg)=f\Delta(g)+\Delta(f)S(g)$. This solves Exercise 6 **(c)**.

(d) Let f be any function. Let $x\in\mathbb{Q}$. Then,

$$\begin{aligned}
(\Delta(S(f)))(x) &= \underbrace{(S(f))(x+1)}_{\substack{=f((x+1)+1) \\ \text{(by the definition of } S)}} - \underbrace{(S(f))(x)}_{\substack{=f(x+1) \\ \text{(by the definition of } S)}} \qquad \text{(by the definition of } \Delta) \\
&= f((x+1)+1)-f(x+1).
\end{aligned}$$

Comparing this with

$$\begin{aligned}
(S(\Delta(f)))(x) &= (\Delta(f))(x+1) \qquad \text{(by the definition of } S) \\
&= f((x+1)+1)-f(x+1) \qquad \text{(by the definition of } \Delta),
\end{aligned}$$

we obtain $(\Delta(S(f)))(x)=(S(\Delta(f)))(x)$.

Now, forget that we fixed x. We thus have shown that $(\Delta(S(f)))(x)=(S(\Delta(f)))(x)$ for each $x\in\mathbb{Q}$. In other words, $\Delta(S(f))=S(\Delta(f))$. Thus, $(\Delta\circ S)(f)=\Delta(S(f))=S(\Delta(f))=(S\circ\Delta)(f)$.

Now, forget that we fixed f. We thus have proven that $(\Delta\circ S)(f)=(S\circ\Delta)(f)$ for each function f. In other words, $\Delta\circ S=S\circ\Delta$. This solves Exercise 6 **(d)**. □

0.7. Necklaces 3: Fermat's Little Theorem

And finally, let's take our tale of periodic tuples and necklaces to its (temporary) conclusion:[13]

[13] To remind: You are allowed to use the exercises from previous problem sets even if you did not solve them.

Exercise 7. This exercise is a continuation of Exercise 7 on homework set #2 and of Exercise 6 on homework set #3. We shall therefore use the notations introduced in these two exercises.

Let p be a **prime** number. Let X be a set.

(a) Let $\mathbf{x} \in X^p$ be a p-tuple. Prove that:

- if all entries of $\mathbf{x}$ are equal (that is, if $\mathbf{x}$ has the form $\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$ for some $x \in X$), then $\left|[\mathbf{x}]_\sim\right| = 1$;
- otherwise, we have $\left|[\mathbf{x}]_\sim\right| = p$.

[*Example:* If $p = 3$, then the 3-tuple $\mathbf{x} = (5, 5, 5)$ satisfies $\left|[\mathbf{x}]_\sim\right| = 1$, while the 3-tuple $\mathbf{x} = (1, 3, 1)$ satisfies $\left|[\mathbf{x}]_\sim\right| = 3$.]

A p-necklace N is said to be *aperiodic* if $|N| = p$.

(b) Assume that the set X is finite. Prove that the number of all aperiodic p-necklaces (over X) is $\dfrac{|X|^p - |X|}{p}$.

[*Example:* If $p = 3$ and $X = \{1, 2, 3\}$, then the aperiodic p-necklaces over X are

$$[(1,1,2)]_\sim, \ [(1,1,3)]_\sim, \ [(1,2,2)]_\sim, \ [(1,2,3)]_\sim,$$
$$[(1,3,2)]_\sim, \ [(1,3,3)]_\sim, \ [(2,2,3)]_\sim, \ [(2,3,3)]_\sim.$$

You can, of course, write them differently: e.g., $[(1,2,3)]_\sim$ is also known as $[(2,3,1)]_\sim$ (but $[(1,3,2)]_\sim$ is different). The p-necklaces that are not aperiodic are $[(1,1,1)]_\sim$, $[(2,2,2)]_\sim$ and $[(3,3,3)]_\sim$.]

(c) Prove *Fermat's Little Theorem*, which states that $p \mid a^p - a$ for every integer a. [**Note:** a might be negative.]

(d) Assume that the set X is finite. Prove that the number of all p-necklaces (over X) is $\dfrac{|X|^p + (p-1)|X|}{p}$.

Solution to Exercise 7. Consider the map $c : X^p \to X^p$, defined as in Exercise 7 on homework set #2 (applied to $n = p$).

During our solution to Exercise 7 **(d)** on Math 4990 homework set #2, we have proven that $c^n(\mathbf{x}) = \mathbf{x}$ for every positive integer n and every $\mathbf{x} \in X^n$. Applying this to $n = p$, we conclude that

$$c^p(\mathbf{x}) = \mathbf{x} \qquad \text{for every } \mathbf{x} \in X^p. \tag{18}$$

Notice that $p > 1$ (since p is prime), so that $p \neq 1$ and $p \geq 1$.

(a) We must prove the following two observations:

Observation 1: If all entries of $\mathbf{x}$ are equal, then $\left|[\mathbf{x}]_\sim\right| = 1$.

> *Observation 2:* If not all entries of $\mathbf{x}$ are equal, then $\left|[\mathbf{x}]_\sim\right| = p$.

Before we prove these two observations, let us state two facts that follow from previously solved exercises:

> *Observation 3:* Let m be the smallest nonzero period of $\mathbf{x}$. Then, $m \mid p$ and $\left|[\mathbf{x}]_\sim\right| = m$.

> *Observation 4:* The smallest nonzero period of $\mathbf{x}$ exists.

[*Proof of Observation 3:* Exercise 7 **(d)** on homework set #2 (applied to $n = p$) shows that m divides p. In other words, $m \mid p$. It thus remains to show that $\left|[\mathbf{x}]_\sim\right| = m$.

Exercise 6 **(c)** on homework set #3 (applied to $n = p$) shows that the m tuples $c^0(\mathbf{x}), c^1(\mathbf{x}), \ldots, c^{m-1}(\mathbf{x})$ are distinct, and that $\left|[\mathbf{x}]_\sim\right| = m$. Thus, $\left|[\mathbf{x}]_\sim\right| = m$ is proven. This completes the proof of Observation 3.]

[*Proof of Observation 4:* The nonnegative integer p satisfies $c^p(\mathbf{x}) = \mathbf{x}$ (by (18)). In other words, p is a period of $\mathbf{x}$. Of course, p is nonzero. Hence, a nonzero period of $\mathbf{x}$ exists (namely, p). Thus, the smallest nonzero period of $\mathbf{x}$ exists. This proves Observation 4.]

[*Proof of Observation 1:* Assume that all entries of $\mathbf{x}$ are equal. We must show that $\left|[\mathbf{x}]_\sim\right| = 1$.

We know that all entries of $\mathbf{x}$ are equal. In other words, $\mathbf{x}$ has the form $\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$ for some $x \in X$. Consider this x. Thus, $\mathbf{x} = \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$. Hence,

$$c(\mathbf{x}) = c\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right) = \left(\underbrace{x, x, \ldots, x}_{p-1 \text{ times}}, x\right) \qquad \text{(by the definition of } c\text{)}$$
$$= \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right) = \mathbf{x}.$$

Thus, $\underbrace{c^1}_{=c}(\mathbf{x}) = c(\mathbf{x}) = \mathbf{x}$. In other words, 1 is a period of $\mathbf{x}$ (by the definition of a "period"). Thus, 1 is a nonzero period of $\mathbf{x}$ (since 1 is nonzero), but no nonzero period of $\mathbf{x}$ can be smaller than 1 (since no nonzero positive integer can be smaller than 1). Hence, 1 is the smallest nonzero period of $\mathbf{x}$. Therefore, Observation 3 (applied to $m = 1$) shows that $1 \mid p$ and $\left|[\mathbf{x}]_\sim\right| = 1$. In particular, $\left|[\mathbf{x}]_\sim\right| = 1$. This proves Observation 1.]

[*Proof of Observation 2:* Assume that not all entries of $\mathbf{x}$ are equal. We must show that $\left|[\mathbf{x}]_{\sim}\right| = p$.

Observation 4 shows that the smallest nonzero period of $\mathbf{x}$ exists. Let m be this period. Thus, Observation 3 shows that $m \mid p$ and $\left|[\mathbf{x}]_{\sim}\right| = m$.

Now, m is a nonnegative integer (since it is a period of $\mathbf{x}$) and divides p (since $m \mid p$). Hence, m is a positive divisor of p. Since the only positive divisors of p are 1 and p (because p is prime), we thus conclude that m is either 1 or p. In other words, either $m = 1$ or $m = p$.

Now, it is easy to see that $m = 1$ cannot hold[14]. Hence, $m = p$ (since either $m = 1$ or $m = p$). Thus, $\left|[\mathbf{x}]_{\sim}\right| = m = p$. This proves Observation 2.]

Combining Observation 1 and Observation 2, we obtain the statement of Exercise 7 **(a)**.

(b) The following proof is hardly a masterpiece of mathematical writing, but at least it isn't missing any steps...

We shall call a p-necklace N *uniperiodic* if $|N| = 1$. Hence, each uniperiodic p-necklace N satisfies

$$|N| = 1. \tag{19}$$

Recall that a p-necklace N is aperiodic if and only if it satisfies $|N| = p$ (by the definition of "aperiodic"). Hence, each aperiodic p-necklace N satisfies

$$|N| = p. \tag{20}$$

Recall that the p-necklaces were defined as the $\sim$-equivalence classes. Hence, these p-necklaces partition the set X^p; in particular, they are disjoint, and their

[14] *Proof.* Assume the contrary. Thus, $m = 1$. Thus, 1 is a period of $\mathbf{x}$ (since m is a period of $\mathbf{x}$). In other words, $c^1(\mathbf{x}) = \mathbf{x}$ (by the definition of a "period").

Write the p-tuple $\mathbf{x} \in X^p$ in the form $\mathbf{x} = (x_1, x_2, \ldots, x_p)$ for some $x_1, x_2, \ldots, x_p \in X$. Hence, the entries of $\mathbf{x}$ are the p elements $x_1, x_2, \ldots, x_p$. Now,

$$\begin{aligned}(x_1, x_2, \ldots, x_p) = \mathbf{x} &= \underbrace{c^1}_{=c} \left(\underbrace{\mathbf{x}}_{=(x_1, x_2, \ldots, x_p)} \right) \qquad \left(\text{since } c^1(\mathbf{x}) = \mathbf{x}\right) \\ &= c(x_1, x_2, \ldots, x_p) = (x_2, x_3, \ldots, x_{p-1}, x_p)\end{aligned}$$

(by the definition of c). In other words, we have the p equalities

$$x_1 = x_2, \qquad x_2 = x_3, \qquad x_3 = x_4, \qquad \ldots, \qquad x_{p-1} = x_p, \qquad x_p = x_1.$$

The first $p-1$ of these p equalities are

$$x_1 = x_2, \qquad x_2 = x_3, \qquad x_3 = x_4, \qquad \ldots, \qquad x_{p-1} = x_p.$$

We can combine these $p-1$ equalities to a chain of equalities:

$$x_1 = x_2 = \cdots = x_p.$$

In other words, all p elements $x_1, x_2, \ldots, x_p$ are equal. In other words, all entries of $\mathbf{x}$ are equal (since the entries of $\mathbf{x}$ are the p elements $x_1, x_2, \ldots, x_p$). This contradicts the assumption that not all entries of $\mathbf{x}$ are equal. This contradiction concludes our proof.

union is X^p. Thus, the size of X^p is the sum of the sizes of all p-necklaces. Written out as an equation, this says the following:

$$|X^p| = \sum_{N \text{ is a } p\text{-necklace}} |N|. \tag{21}$$

Next, we show some simple observations:

> *Observation 5:* Let N be a p-necklace. Then, we have the following logical equivalence:
>
> $$(N \text{ is not aperiodic}) \iff (N \text{ is uniperiodic}). \tag{22}$$

[*Proof of Observation 5:* We know that N is a p-necklace. In other words, N is a $\sim$-equivalence class (because the p-necklaces were defined as the $\sim$-equivalence classes). In other words, $N = [\mathbf{x}]_\sim$ for some $\mathbf{x} \in X^p$. Consider this $\mathbf{x}$.

Assume that N is not aperiodic. In other words, we don't have $|N| = p$ (since N is aperiodic if and only if $|N| = p$ (by the definition of "aperiodic")). Hence, $|N| \neq p$. In view of $N = [\mathbf{x}]_\sim$, this rewrites as $|[\mathbf{x}]_\sim| \neq p$. If not all entries of $\mathbf{x}$ were equal, then we would have $|[\mathbf{x}]_\sim| = p$ (by Observation 2), which would contradict $|[\mathbf{x}]_\sim| \neq p$. Hence, it is impossible that not all entries of $\mathbf{x}$ are equal. Therefore, all entries of $\mathbf{x}$ are equal. Hence, Observation 1 shows that $|[\mathbf{x}]_\sim| = 1$. In other words, $|N| = 1$ (since $N = [\mathbf{x}]_\sim$). In other words, N is uniperiodic (since N is uniperiodic if and only if $|N| = 1$ (by the definition of "uniperiodic")).

Now, forget that we assumed that N is not aperiodic. We thus have shown that if N is not aperiodic, then N is uniperiodic. In other words, we have proven the implication

$$(N \text{ is not aperiodic}) \Longrightarrow (N \text{ is uniperiodic}).$$

On the other hand, it is easy to see that the implication

$$(N \text{ is uniperiodic}) \Longrightarrow (N \text{ is not aperiodic})$$

also holds[15]. Combining these two implications, we obtain the equivalence $(N \text{ is not aperiodic}) \iff (N \text{ is uniperiodic})$. This proves Observation 5.]

On the other hand, it is easy to count the uniperiodic p-necklaces:

> *Observation 6:* The map
>
> $$X \to \{\text{uniperiodic } p\text{-necklaces}\},$$
> $$x \mapsto \left[\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)\right]_\sim$$
>
> is well-defined and bijective.

[15] *Proof.* Assume that N is uniperiodic. We must show that N is not aperiodic.
Indeed, assume the contrary. Thus, N is aperiodic. Hence, $|N| = p$ (by (20)). But N is uniperiodic. Thus, $|N| = 1$ (by (19)). Hence, $p = |N| = 1$. But $p \neq 1$ (since p is prime). This contradicts $p = 1$. This contradiction completes our proof that N is not aperiodic. Qed.

[*Proof of Observation 6:* For each $x \in X$, we have

$$\left[\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)\right]_{\sim} \in \{\text{uniperiodic } p\text{-necklaces}\}$$

[16]. Hence, the map

$$X \to \{\text{uniperiodic } p\text{-necklaces}\},$$
$$x \mapsto \left[\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)\right]_{\sim}$$

is well-defined. Denote this map by A.

[16] *Proof.* Let $x \in X$. Define a p-tuple $\mathbf{x} \in X^p$ by $\mathbf{x} = \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$. Then, all entries of $\mathbf{x}$ are equal (indeed, they are all equal to x). Hence, $|[\mathbf{x}]_{\sim}| = 1$ (by Observation 1). In other words, $[\mathbf{x}]_{\sim}$ is uniperiodic (since $[\mathbf{x}]_{\sim}$ is uniperiodic if and only if $|[\mathbf{x}]_{\sim}| = 1$ (by the definition of "uniperiodic")). In other words, $[\mathbf{x}]_{\sim} \in \{\text{uniperiodic } p\text{-necklaces}\}$. Since $\mathbf{x} = \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$, this rewrites as follows:

$$\left[\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)\right]_{\sim} \in \{\text{uniperiodic } p\text{-necklaces}\}.$$

Qed.

The map A is injective[17] and surjective[18]. Hence, this map A is bijective. This proves Observation 6 (since A is precisely the map considered in Observation 6).]

Observation 6 shows that there is a bijection $X \to \{\text{uniperiodic } p\text{-necklaces}\}$.

[17] *Proof.* Let x and y be two elements of X such that $A(x) = A(y)$. We shall prove that $x = y$.

Define a p-tuple $\mathbf{x} \in X^p$ by $\mathbf{x} = \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$. Define a p-tuple $\mathbf{y} \in X^p$ by $\mathbf{y} = \left(\underbrace{y, y, \ldots, y}_{p \text{ times}}\right)$.

The definition of A yields $A(x) = \left[\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)\right]_\sim = [\mathbf{x}]_\sim$ (since $\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right) = \mathbf{x}$). The same argument (with x and $\mathbf{x}$ replaced by y and $\mathbf{y}$) shows that $A(y) = [\mathbf{y}]_\sim$. Thus, $[\mathbf{x}]_\sim = A(x) = A(y) = [\mathbf{y}]_\sim$.

We know that $A(x) \in \{\text{uniperiodic } p\text{-necklaces}\}$ (since A is a map $X \to \{\text{uniperiodic } p\text{-necklaces}\}$). In other words, $A(x)$ is a uniperiodic p-necklace. In other words, $[\mathbf{x}]_\sim$ is a uniperiodic p-necklace (since $A(x) = [\mathbf{x}]_\sim$).

But we know that the necklace $[\mathbf{x}]_\sim$ is uniperiodic. In other words, $|[\mathbf{x}]_\sim| = 1$ (since $[\mathbf{x}]_\sim$ is uniperiodic if and only if $|[\mathbf{x}]_\sim| = 1$ (by the definition of "uniperiodic")). In other words, $[\mathbf{x}]_\sim$ is a 1-element set.

We have $\mathbf{x} \in [\mathbf{x}]_\sim$ (by the definition of the equivalence class $[\mathbf{x}]_\sim$). Similarly, $\mathbf{y} \in [\mathbf{y}]_\sim$. In view of $[\mathbf{x}]_\sim = [\mathbf{y}]_\sim$, this rewrites as $\mathbf{y} \in [\mathbf{x}]_\sim$.

Now, $\mathbf{x}$ and $\mathbf{y}$ are two elements of the 1-element set $[\mathbf{x}]_\sim$ (since $\mathbf{x} \in [\mathbf{x}]_\sim$ and $\mathbf{y} \in [\mathbf{x}]_\sim$). Thus, $\mathbf{x}$ and $\mathbf{y}$ are equal (since any two elements of a 1-element set are equal). In other words, $\mathbf{x} = \mathbf{y}$.

But the p-tuple $\mathbf{x}$ has a first entry (since $p \geq 1$), and this first entry is x (since $\mathbf{x} = \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$). Hence, $x = (\text{the first entry of } \mathbf{x})$. The same argument (with x and $\mathbf{x}$ replaced by y and $\mathbf{y}$) shows that $y = (\text{the first entry of } \mathbf{y})$. Thus, $x = \left(\text{the first entry of } \underbrace{\mathbf{x}}_{=\mathbf{y}}\right) = (\text{the first entry of } \mathbf{y}) = y$.

Now, forget that we fixed x and y. We thus have shown that if x and y are two elements of X such that $A(x) = A(y)$, then $x = y$. In other words, the map A is injective.

[18] *Proof.* Let $N \in \{\text{uniperiodic } p\text{-necklaces}\}$. We shall prove that $N \in A(X)$.

Clearly, N is a uniperiodic p-necklace (since $N \in \{\text{uniperiodic } p\text{-necklaces}\}$). Write N in the form $N = [\mathbf{x}]_\sim$ for some $\mathbf{x} \in X^p$. (This is possible since N is a necklace.)

The p-necklace N is uniperiodic. Thus, $|N| = 1$ (by (19)). From $[\mathbf{x}]_\sim = N$, we obtain $|[\mathbf{x}]_\sim| = |N| = 1 \neq p$. If not all entries of $\mathbf{x}$ were equal, then we would have $|[\mathbf{x}]_\sim| = p$ (by Observation 2), which would contradict the fact that $|[\mathbf{x}]_\sim| \neq p$. Thus, it is impossible that not all entries of $\mathbf{x}$ are equal. Therefore, all entries of $\mathbf{x}$ are equal. In other words, the p-tuple $\mathbf{x}$ has the form $\mathbf{x} = \left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)$ for some $x \in X$. Consider this x. The definition of A yields $A(x) = \left[\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right)\right]_\sim = [\mathbf{x}]_\sim$ (since $\left(\underbrace{x, x, \ldots, x}_{p \text{ times}}\right) = \mathbf{x}$). Hence, $A(x) = [\mathbf{x}]_\sim = N$, so that $N = A\left(\underbrace{x}_{\in X}\right) \in A(X)$.

Hence,

$$|X| = |\{\text{uniperiodic } p\text{-necklaces}\}| . \tag{23}$$

Now, $|X^p| = |X|^p$. Comparing this with (21), we obtain

$$\begin{aligned}
|X|^p &= \sum_{N \text{ is a } p\text{-necklace}} |N| \\
&= \sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is aperiodic}}} |N| + \underbrace{\sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is not aperiodic}}}}_{\substack{= \sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is uniperiodic}}} \\ \text{(by the equivalence (22))}}} |N| \\
&\qquad \left(\begin{array}{c}\text{since each } p\text{-necklace is either aperiodic}\\ \text{or not aperiodic (but not both)}\end{array}\right) \\
&= \sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is aperiodic}}} \underbrace{|N|}_{\substack{=p\\ \text{(by (20))}}} + \sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is uniperiodic}}} \underbrace{|N|}_{\substack{=1\\ \text{(by (19))}}} \\
&= \underbrace{\sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is aperiodic}}} p}_{=(\text{the number of all aperiodic } p\text{-necklaces})\cdot p} + \underbrace{\sum_{\substack{N \text{ is a } p\text{-necklace;}\\ N \text{ is uniperiodic}}} 1}_{\substack{=(\text{the number of all uniperiodic } p\text{-necklaces})\cdot 1\\ =(\text{the number of all uniperiodic } p\text{-necklaces})\\ =|\{\text{uniperiodic } p\text{-necklaces}\}|=|X|\\ \text{(by (23))}}} \\
&= (\text{the number of all aperiodic } p\text{-necklaces}) \cdot p + |X| .
\end{aligned}$$

Solving this equation for (the number of all aperiodic p-necklaces), we obtain

$$(\text{the number of all aperiodic } p\text{-necklaces}) = \frac{|X|^p - |X|}{p}.$$

This solves Exercise 7 **(b)**.

(c) *First solution to Exercise 7 (c):* Let a be an integer. We must show that $p \mid a^p - a$. If $p = 2$, then this is easy to prove[19]. Hence, for the rest of this proof, we WLOG assume that we don't have $p = 2$. Hence, $p \neq 2$, so that p is odd (since p is a prime). Therefore, $(-1)^p = -1$. Now, we are in one of the following two cases:

Case 1: The integer a is nonnegative.

Case 2: The integer a is negative.

Now, forget that we fixed N. We thus have shown that $N \in A(X)$ for each $N \in \{\text{uniperiodic } p\text{-necklaces}\}$. Hence, $\{\text{uniperiodic } p\text{-necklaces}\} \subseteq A(X)$. In other words, the map A is surjective.

[19] *Proof.* Assume that $p = 2$. At least one of the two integers a and $a - 1$ is even (since their difference $a - (a-1) = 1$ is odd). Hence, their product $a(a-1)$ is even. In other words, $2 \mid a(a-1)$. In other words, $2 \mid a^2 - a$ (since $a^2 - a = a(a-1)$). In other words, $p \mid a^p - a$ (since $p = 2$). Qed.

Let us first consider Case 1. In this case, the integer a is nonnegative. Thus, there exists a finite set X with $|X| = a$. Consider such an X. Then, Exercise 7 **(b)** shows that the number of all aperiodic p-necklaces (over X) is $\dfrac{|X|^p - |X|}{p}$. Hence, $\dfrac{|X|^p - |X|}{p}$ is a nonnegative integer (since the number of all aperiodic p-necklaces is clearly a nonnegative integer). In particular, $\dfrac{|X|^p - |X|}{p} \in \mathbb{Z}$. Hence, $p \mid |X|^p - |X| = a^p - a$ (since $|X| = a$). Thus, $p \mid a^p - a$ is proven in Case 1.

Let us now consider Case 2. In this case, the integer a is negative. Hence, the integer $-a$ is positive; in particular, $-a$ is nonnegative. Thus, there exists a finite set X with $|X| = -a$. Consider such an X. Then, Exercise 7 **(b)** shows that the number of all aperiodic p-necklaces (over X) is $\dfrac{|X|^p - |X|}{p}$. Hence, $\dfrac{|X|^p - |X|}{p}$ is a nonnegative integer (since the number of all aperiodic p-necklaces is clearly a nonnegative integer). In particular, $\dfrac{|X|^p - |X|}{p} \in \mathbb{Z}$. Hence,

$$\begin{aligned} p \mid |X|^p - |X| &= (-a)^p - (-a) \qquad \text{(since } |X| = -a\text{)} \\ &= \underbrace{(-a)^p}_{=(-1)^p a^p} + a = \underbrace{(-1)^p}_{=-1} a^p + a = (-1)\, a^p + a = -\left(a^p - a\right) \\ &\mid (-1)\left(-\left(a^p - a\right)\right) = a^p - a. \end{aligned}$$

Thus, $p \mid a^p - a$ is proven in Case 2.

We have now proven $p \mid a^p - a$ in both Cases 1 and 2. Thus, $p \mid a^p - a$ always holds. This solves Exercise 7 **(c)**.

Second solution to Exercise 7 ***(c)****:* The following proof is nicer, but it uses a little bit of modular arithmetic: Namely, we shall use the fact that if b and c are two integers satisfying $b \equiv c \mod p$, and if $m \in \mathbb{N}$, then

$$b^m \equiv c^m \mod p. \tag{24}$$

(This is easy to verify by induction over m. It does not matter here that p is a prime.)

Let a be an integer. We must show that $p \mid a^p - a$.

Let c be the remainder upon dividing a by p. Then, $c \in \{0, 1, \ldots, p-1\}$ and $a \equiv c \mod p$. Hence, (24) (applied to $b = a$ and $m = p$) yields $a^p \equiv c^p \mod p$. But c is a nonnegative integer (since $c \in \{0, 1, \ldots, p-1\}$). Hence, there exists a finite set X with $|X| = c$. Consider such an X. Then, Exercise 7 **(b)** shows that the number of all aperiodic p-necklaces (over X) is $\dfrac{|X|^p - |X|}{p}$. Hence, $\dfrac{|X|^p - |X|}{p}$ is a nonnegative integer (since the number of all aperiodic p-necklaces is clearly a nonnegative integer). In particular, $\dfrac{|X|^p - |X|}{p} \in \mathbb{Z}$. Hence, $p \mid |X|^p - |X| = c^p - c$

(since $|X| = c$). In other words, $c^p \equiv c \mod p$. Thus, $a^p \equiv c^p \equiv c \equiv a \mod p$. In other words, $p \mid a^p - a$. This solves Exercise 7 **(c)** again.

(d) We have

$$\begin{aligned}
&(\text{the number of all } p\text{-necklaces}) \\
&= \underbrace{(\text{the number of all } p\text{-necklaces that are aperiodic})}_{\substack{=(\text{the number of all aperiodic } p\text{-necklaces}) \\ =\dfrac{|X|^p - |X|}{p} \\ (\text{by Exercise 7 } \textbf{(b)})}} \\
&\qquad + \underbrace{(\text{the number of all } p\text{-necklaces that are not aperiodic})}_{=|\{N \text{ is a } p\text{-necklace} \mid N \text{ is not aperiodic}\}|} \\
&\qquad \left(\begin{array}{c}\text{since each } p\text{-necklace is either aperiodic} \\ \text{or not aperiodic (but not both)}\end{array}\right) \\
&= \frac{|X|^p - |X|}{p} + \left| \underbrace{\{N \text{ is a } p\text{-necklace} \mid N \text{ is not aperiodic}\}}_{\substack{=\{N \text{ is a } p\text{-necklace} \mid N \text{ is uniperiodic}\} \\ (\text{by the equivalence (22)})}} \right| \\
&= \frac{|X|^p - |X|}{p} + \underbrace{|\{N \text{ is a } p\text{-necklace} \mid N \text{ is uniperiodic}\}|}_{\substack{=|\{\text{uniperiodic } p\text{-necklaces}\}|=|X| \\ (\text{by (23)})}} \\
&= \frac{|X|^p - |X|}{p} + |X| = \frac{|X|^p - |X| + p|X|}{p} = \frac{|X|^p + (p-1)|X|}{p}.
\end{aligned}$$

This solves Exercise 7 **(d)**. □

We have thus counted p-necklaces for a prime number p. Counting n-necklaces for n composite is harder; we will learn about this later.

References

[Camero16] Peter J. Cameron, *St Andrews Notes on Advanced Combinatorics, Part 1: The Art of Counting*, 28 March 2016.
https://cameroncounts.files.wordpress.com/2016/04/acnotes1.pdf
Errata can be found at http://www.cip.ifi.lmu.de/~grinberg/algebra/acnotes1-errata.pdf

[Galvin17] David Galvin, *Basic discrete mathematics*, 13 December 2017.
http://www.cip.ifi.lmu.de/~grinberg/t/17f/60610lectures2017-Galvin.pdf

[Grinbe16b] 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` .

[LeLeMe17] Eric Lehman, F. Thomson Leighton, Albert R. Meyer, *Mathematics for Computer Science*, version 16 October 2017.
`https://courses.csail.mit.edu/6.042/fall17/mcs.pdf`

[Scholz04] Peter Scholze et al, *Art of Problem Solving topic #16482 ("zero sum")*.
`https://artofproblemsolving.com/community/c6h16482`