Math 4990 Fall 2017 (Darij Grinberg): homework set 7 with solutions

Contents

0.1. A generalized principle of inclusion/exclusion

Exercise 1. Let $n \in \mathbb{N}$. Let S be a finite set. Let $A_1, A_2, \ldots, A_n$ be finite subsets of S. Let $k \in \mathbb{N}$. Let S_k be the set of all elements of S that belong to exactly k of the subsets $A_1, A_2, \ldots, A_n$. (In other words, let $S_k = \{s \in S \mid \text{the number of } i \in [n] \text{ satisfying } s \in A_i \text{ equals } k\}$.) Prove that

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

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

Note that the principle of inclusion and exclusion (see, e.g., [Galvin17, §16]) is the particular case of Exercise 1 for $k = 0$ (since $S_0 = S \setminus \bigcup_{i=1}^{n} A_i$).

Exercise 1 is a result of Charles Jordan (see [Comtet74, §4.8, Theorem A] and [DanRot78] for fairly complicated proofs). I further generalize it in [Grinbe16, Theorem 3.44]. Let me here give a self-contained proof.

First, we recall two facts from the solutions to homework set #5:

Proposition 0.1. We have

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

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

Corollary 0.2. Let $n \in \mathbb{N}$. Let $i \in \mathbb{N}$. Then,

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

Next, we recall the classical formula for the size of a subset using Iverson brackets:

Lemma 0.3. Let S be a finite set. Let T be a subset of S. Then,

$$|T| = \sum_{s \in S} [s \in T].$$

Lemma 0.3 allows us to reduce a formula for $|S_k|$ to a formula for $[s \in S_k]$ (for any given $s \in S$). Here is the latter formula:

Lemma 0.4. Let $n \in \mathbb{N}$. Let S be a finite set. Let $A_1, A_2, \ldots, A_n$ be finite subsets of S. Let $k \in \mathbb{N}$. Let S_k be the set of all elements of S that belong to exactly k of the subsets $A_1, A_2, \ldots, A_n$. (In other words, let $S_k = \{s \in S \mid \text{the number of } i \in [n] \text{ satisfying } s \in A_i \text{ equals } k\}$.) Let $s \in S$. Then,

$$[s \in S_k] = \sum_{I \subseteq [n]} (-1)^{|I|-k} \binom{|I|}{k} \left[s \in \bigcap_{i \in I} A_i \right].$$

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

Proof of Lemma 0.4. Define a subset C of $[n]$ by

$$C = \{i \in [n] \mid s \in A_i\}.$$

Thus, for each $i \in [n]$, we have the following equivalence:

$$(i \in C) \iff (s \in A_i). \tag{1}$$

But recall the definition of S_k. From this definition, we obtain the following equivalence:

$$\begin{aligned}
(s \in S_k) &\iff \left(\underbrace{\text{the number of } i \in [n] \text{ satisfying } s \in A_i}_{=|\{i \in [n] \mid s \in A_i\}|} \text{ equals } k \right) \\
&\iff \left(\left| \underbrace{\{i \in [n] \mid s \in A_i\}}_{=C} \right| \text{ equals } k \right) \iff (|C| \text{ equals } k) \\
&\iff (|C| = k).
\end{aligned}$$

Hence, we find the following equality between truth values:

$$[s \in S_k] = [|C| = k]. \tag{2}$$

On the other hand, let I be any subset of $[n]$. Then, we have the following equivalence:

$$\left(s \in \bigcap_{i \in I} A_i\right) \iff \left(\underbrace{s \in A_i}_{\substack{\iff (i \in C) \\ \text{(by (1))}}} \text{ for each } i \in I\right) \iff (i \in C \text{ for each } i \in I)$$
$$\iff (I \subseteq C).$$

Thus,

$$\left[s \in \bigcap_{i \in I} A_i\right] = [I \subseteq C]. \tag{3}$$

Now, forget that we fixed I. We thus have proven (3) for each subset I of $[n]$.

Thus,

$$
\begin{aligned}
&\sum_{I\subseteq[n]} (-1)^{|I|-k} \binom{|I|}{k} \underbrace{\left[s \in \bigcap_{i\in I} A_i\right]}_{\substack{=[I\subseteq C] \\ \text{(by (3))}}} \\
&= \sum_{I\subseteq[n]} (-1)^{|I|-k} \binom{|I|}{k} [I \subseteq C] \\
&= \underbrace{\sum_{\substack{I\subseteq[n]; \\ I\subseteq C}}}_{\substack{=\sum_{I\subseteq C} \\ \text{(since } C\subseteq[n])}} (-1)^{|I|-k} \binom{|I|}{k} \underbrace{[I\subseteq C]}_{\substack{=1 \\ \text{(since } I\subseteq C)}} + \sum_{\substack{I\subseteq[n]; \\ \text{not } I\subseteq C}} (-1)^{|I|-k} \binom{|I|}{k} \underbrace{[I\subseteq C]}_{\substack{=0 \\ \text{(since we don't have } I\subseteq C)}} \\
&= \sum_{I\subseteq C} (-1)^{|I|-k} \binom{|I|}{k} + \underbrace{\sum_{\substack{I\subseteq[n]; \\ \text{not } I\subseteq C}} (-1)^{|I|-k} \binom{|I|}{k} 0}_{=0} = \underbrace{\sum_{I\subseteq C}}_{=\sum_{j\in\mathbb{N}} \sum_{\substack{I\subseteq C; \\ |I|=j}}} (-1)^{|I|-k} \binom{|I|}{k} \\
&= \sum_{j\in\mathbb{N}} \sum_{\substack{I\subseteq C; \\ |I|=j}} \underbrace{(-1)^{|I|-k} \binom{|I|}{k}}_{\substack{=(-1)^{j-k}\binom{j}{k} \\ \text{(since } |I|=j)}} = \sum_{j\in\mathbb{N}} \underbrace{\sum_{\substack{I\subseteq C; \\ |I|=j}} (-1)^{j-k} \binom{j}{k}}_{=(\text{the number of all subsets } I \text{ of } C \text{ satisfying } |I|=j)(-1)^{j-k}\binom{j}{k}} \\
&= \sum_{j\in\mathbb{N}} \underbrace{(\text{the number of all subsets } I \text{ of } C \text{ satisfying } |I| = j)}_{\substack{=(\text{the number of all } j\text{-element subsets of } C)=\binom{|C|}{j} \\ \text{(by the combinatorial interpretation of the binomial coefficients)}}} (-1)^{j-k} \binom{j}{k} \\
&= \sum_{j\in\mathbb{N}} \binom{|C|}{j} (-1)^{j-k} \binom{j}{k} = \sum_{j\in\mathbb{N}} (-1)^{j-k} \binom{|C|}{j} \binom{j}{k}
\end{aligned}
$$

$$= \sum_{j=0}^{|C|} \underbrace{(-1)^{j-k}}_{=(-1)^{j+k}=(-1)^j(-1)^k} \binom{|C|}{j} \binom{j}{k} + \sum_{j=|C|+1}^{\infty} (-1)^{j-k} \underbrace{\binom{|C|}{j}}_{\substack{=0 \\ \text{(by Proposition 0.1,} \\ \text{applied to } |C| \text{ and } j \\ \text{instead of } m \text{ and } n \text{ (since } |C|<j \\ \text{(because } j\geq|C|+1>|C|)))}} \binom{j}{k}$$

$$\left(\text{here, we have split the summation at } j = |C|\right)$$

$$= \sum_{j=0}^{|C|} (-1)^j (-1)^k \binom{|C|}{j} \binom{j}{k} + \underbrace{\sum_{j=|C|+1}^{\infty} (-1)^{j-k} 0 \binom{j}{k}}_{=0}$$

$$= \sum_{j=0}^{|C|} (-1)^j (-1)^k \binom{|C|}{j} \binom{j}{k} = (-1)^k \underbrace{\sum_{j=0}^{|C|} (-1)^j \binom{|C|}{j} \binom{j}{k}}_{\substack{=(-1)^k[|C|=k] \\ \text{(by Corollary 0.2,} \\ \text{applied to } |C| \text{ and } k \text{ instead of } n \text{ and } i)}}$$

$$= \underbrace{(-1)^k (-1)^k}_{=(-1)^{2k}=1} [|C| = k] = [|C| = k] = [s \in S_k] \qquad (\text{by (2)}).$$

This proves Lemma 0.4. □

Solution to Exercise 1. For every subset I of $[n]$, the intersection $\bigcap_{i \in I} A_i$ is a subset of S [1]. Thus, for every subset I of $[n]$, we have

$$\left| \bigcap_{i \in I} A_i \right| = \sum_{s \in S} \left[s \in \bigcap_{i \in I} A_i \right]$$

[1] In fact, this is obvious when I is nonempty (because all the A_i are subsets of S), but it also holds when I is empty (because in this case, the intersection $\bigcap_{i \in I} A_i = \bigcap_{i \in \varnothing} A_i$ is defined to be S).

(by Lemma 0.3, applied to $T = \bigcap_{i \in I} A_i$). Hence,

$$\begin{aligned}
&\sum_{I \subseteq [n]} (-1)^{|I|-k} \binom{|I|}{k} \underbrace{\left| \bigcap_{i \in I} A_i \right|}_{= \sum_{s \in S} \left[s \in \bigcap_{i \in I} A_i \right]} \\
&= \sum_{I \subseteq [n]} (-1)^{|I|-k} \binom{|I|}{k} \sum_{s \in S} \left[s \in \bigcap_{i \in I} A_i \right] = \underbrace{\sum_{I \subseteq [n]} \sum_{s \in S}}_{= \sum_{s \in S} \sum_{I \subseteq [n]}} (-1)^{|I|-k} \binom{|I|}{k} \left[s \in \bigcap_{i \in I} A_i \right] \\
&= \sum_{s \in S} \underbrace{\sum_{I \subseteq [n]} (-1)^{|I|-k} \binom{|I|}{k} \left[s \in \bigcap_{i \in I} A_i \right]}_{\substack{= [s \in S_k] \\ \text{(by Lemma 0.4)}}} = \sum_{s \in S} [s \in S_k] .
\end{aligned}$$

Comparing this with $|S_k| = \sum_{s \in S} [s \in S_k]$ (which follows from Lemma 0.3, applied to $T = S_k$), we obtain

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

This solves Exercise 1. □

0.2. Summing fixed point numbers of permutations

Recall that for any $n \in \mathbb{N}$, we let S_n denote the set of all permutations of $[n]$.

If S is a finite set, and if $f : S \to S$ is a map, then we let $\operatorname{Fix} f$ denote the set of all fixed points of f. (That is, $\operatorname{Fix} f = \{ s \in S \mid f(s) = s \}$.)

Exercise 2. Let n be a positive integer. Prove that $\sum_{w \in S_n} |\operatorname{Fix} w| = n!$.

[Hint: Rewrite $|\operatorname{Fix} w|$ as $\sum_{i \in [n]} [w(i) = i]$.]

(In other words, this exercise states that the average number of fixed points of a permutation of $[n]$ is 1.)

Exercise 2 was Problem 1 at the International Mathematical Olympiad (IMO) 1987.

Our solution to Exercise 2 relies on the following facts:

Lemma 0.5. Let $m \in \mathbb{N}$. Let G be an m-element set. Then, the number of all permutations of G is $m!$.

Proof of Lemma 0.5 (sketched). There is a bijection $\alpha : G \to [m]$ (since G is an m-element set). Fix such an α. Then, the map

$$\{\text{permutations of } G\} \to \{\text{permutations of } [m]\},$$
$$\sigma \mapsto \alpha \circ \sigma \circ \alpha^{-1}$$

is also a bijection[2]. Hence,

$$\begin{aligned} |\{\text{permutations of } G\}| &= |\{\text{permutations of } [m]\}| \\ &= (\text{the number of all permutations of } [m]) = m!. \end{aligned}$$

In other words, the number of all permutations of G is $m!$. This proves Lemma 0.5. $\square$

Lemma 0.6. Let n be a positive integer. Let $i \in [n]$. Then, the number of all permutations $w \in S_n$ satisfying $w(i) = i$ is $(n-1)!$.

Proof of Lemma 0.6 (sketched). Roughly speaking, a permutation $w \in S_n$ satisfying $w(i) = i$ is "nothing but" a permutation of the $(n-1)$-element set $[n] \setminus \{i\}$ (because it has to map i to i, and therefore must map the remaining elements of $[n]$ to elements other than i). This is not rigorous, because strictly speaking a permutation of $[n]$ cannot be a permutation of $[n] \setminus \{i\}$ (after all, the former has domain $[n]$ while the latter only has domain $[n] \setminus \{i\}$). Here is a rigorous version of the above statement:

To each permutation $w \in S_n$ satisfying $w(i) = i$, we can assign a permutation $\widetilde{w}$ of $[n] \setminus \{i\}$ by letting

$$\widetilde{w}(p) = w(p) \qquad \text{for each } p \in [n] \setminus \{i\}.$$

This defines a map

$$A : \{w \in S_n \mid w(i) = i\} \to \{\text{permutations of } [n] \setminus \{i\}\},$$
$$w \mapsto \widetilde{w}. \tag{4}$$

Conversely, to each permutation u of $[n] \setminus \{i\}$, we can assign a permutation $\widehat{u} \in S_n$ satisfying $\widehat{u}(i) = i$ by setting

$$\widehat{u}(p) = \begin{cases} u(p), & \text{if } p \neq i; \\ i, & \text{if } p = i \end{cases} \qquad \text{for each } p \in [n].$$

This defines a map

$$B : \{\text{permutations of } [n] \setminus \{i\}\} \to \{w \in S_n \mid w(i) = i\},$$
$$u \mapsto \widehat{u}.$$

[2]This is straightforward to verify.

The maps A and B are well-defined and mutually inverse[3]. Thus, there is a bijection from the set $\{w \in S_n \mid w(i) = i\}$ to the set $\{\text{permutations of } [n] \setminus \{i\}\}$ (namely, A). Hence,

$$\begin{aligned} |\{w \in S_n \mid w(i) = i\}| &= |\{\text{permutations of } [n] \setminus \{i\}\}| \\ &= (\text{the number of all permutations of } [n] \setminus \{i\}) \\ &= (n-1)! \end{aligned}$$

(by Lemma 0.5 (applied to $G = [n] \setminus \{i\}$ and $m = n-1$), because $[n] \setminus \{i\}$ is an $(n-1)$-element set). In other words, the number of all permutations $w \in S_n$ satisfying $w(i) = i$ is $(n-1)!$. This proves Lemma 0.6. $\square$

Solution to Exercise 2 (sketched). If $w \in S_n$ and $i \in [n]$, then

$$[i \in \operatorname{Fix} w] = [w(i) = i] \tag{5}$$

[4].

If $w \in S_n$, then $\operatorname{Fix} w$ is a subset of $[n]$, and therefore Lemma 0.3 (applied to $S = [n]$ and $T = \operatorname{Fix} w$) yields

$$\begin{aligned} |\operatorname{Fix} w| &= \sum_{s \in [n]} [s \in \operatorname{Fix} w] = \sum_{i \in [n]} \underbrace{[i \in \operatorname{Fix} w]}_{\substack{=[w(i)=i] \\ \text{(by (5))}}} \qquad \left(\begin{array}{c}\text{here, we have renamed the} \\ \text{summation index } s \text{ as } i\end{array}\right) \\ &= \sum_{i \in [n]} [w(i) = i]. \end{aligned} \tag{6}$$

But if $i \in [n]$, then $\{w \in S_n \mid w(i) = i\}$ is a subset of S_n, and therefore Lemma 0.3 (applied to $S = S_n$ and $T = \{w \in S_n \mid w(i) = i\}$) yields

$$\begin{aligned} |\{w \in S_n \mid w(i) = i\}| &= \sum_{s \in S_n} \left[\underbrace{s \in \{w \in S_n \mid w(i) = i\}}_{\iff (s(i)=i)} \right] \\ &= \sum_{s \in S_n} [s(i) = i] = \sum_{w \in S_n} [w(i) = i] \end{aligned} \tag{7}$$

(here, we have renamed the summation index s as w).

[3]This is straightforward to check (just remember that permutations are bijective).
[4]because of the equivalence $(i \in \operatorname{Fix} w) \iff (i \text{ is a fixed point of } w) \iff (w(i) = i)$

Now,

$$\begin{aligned}
\sum_{w\in S_n} \underbrace{\left|\operatorname{Fix} w\right|}_{\substack{=\sum_{i\in[n]} [w(i)=i] \\ \text{(by (6))}}} &= \underbrace{\sum_{w\in S_n} \sum_{i\in[n]}}_{=\sum_{i\in[n]} \sum_{w\in S_n}} [w(i)=i] = \sum_{i\in[n]} \underbrace{\sum_{w\in S_n} [w(i)=i]}_{\substack{=\left|\{w\in S_n \mid w(i)=i\}\right| \\ \text{(by (7))}}} \\
&= \sum_{i\in[n]} \underbrace{\left|\{w\in S_n \mid w(i)=i\}\right|}_{\substack{=(\text{the number of all permutations } w\in S_n \text{ satisfying } w(i)=i) \\ =(n-1)! \\ \text{(by Lemma 0.6)}}} \\
&= \sum_{i\in[n]} (n-1)! = \underbrace{\left|[n]\right|}_{=n} \cdot (n-1)! = n\cdot(n-1)! = n!.
\end{aligned}$$

This solves Exercise 2. □

Remark 0.7. Exercise 2 can be generalized: If $n\in\mathbb{N}$ and $k\in\mathbb{N}$ satisfy $n\geq k$, then

$$\sum_{w\in S_n} \binom{\left|\operatorname{Fix} w\right|}{k} = (n-k)!\binom{n}{k} = \frac{n!}{k!}.$$

Do you see how the above solution can be extended to cover this generalization?

0.3. Transpositions $t_{1,i}$ generate permutations

Recall a basic notation regarding permutations:

Definition 0.8. Let $n\in\mathbb{N}$. Let i and j be two distinct elements of $[n]$. We let $t_{i,j}$ be the permutation in S_n which switches i with j while leaving all other elements of $[n]$ unchanged. Such a permutation is called a *transposition*.

Exercise 3. Let $n\in\mathbb{N}$. Prove that each permutation in S_n can be written as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$.

(Note that this composition can be empty – in which case it is understood to be id –, and it can contain any given transposition multiple times.)

To solve this exercise, we recall another definition:

Definition 0.9. Let $n\in\mathbb{N}$. Let $i\in[n-1]$. Then, s_i denotes the permutation $t_{i,i+1}\in S_n$.

We shall use the following well-known fact ([Grinbe16, Exercise 5.1 **(b)**]):

Lemma 0.10. Let $n\in\mathbb{N}$. Each permutation in S_n can be written as a composition of some of the transpositions $s_1, s_2, \ldots, s_{n-1}$.

The following is easy to check:

Lemma 0.11. Let $n \in \mathbb{N}$. Let $i \in [n-1]$ be such that $i > 1$. Then, $s_i = t_{1,i+1} \circ t_{1,i} \circ t_{1,i+1}$.

Lemma 0.11 can be proven by straightforward verification (just check how s_i and $t_{1,i+1} \circ t_{1,i} \circ t_{1,i+1}$ transform a given element of $[n]$, depending on whether this element is 1, i or $i+1$ or something else). Let us give a slightly more skillful argument. The following fact is simple and well-known ([Grinbe16, Exercise 5.17 **(a)**]):

Lemma 0.12. Let $n \in \mathbb{N}$. Let $k \in [n]$. For every $\sigma \in S_n$ and every k distinct elements $i_1, i_2, \ldots, i_k$ of $[n]$, we have

$$\sigma \circ \operatorname{cyc}_{i_1, i_2, \ldots, i_k} \circ \sigma^{-1} = \operatorname{cyc}_{\sigma(i_1), \sigma(i_2), \ldots, \sigma(i_k)}.$$

Proof of Lemma 0.11 (sketched). From $i \in [n-1]$, we obtain $i \leq n-1$. But $i > 1$, so that $i \geq 2$, and thus $2 \leq i \leq n-1$. Hence, $n \geq 3$. Thus, $2 \in [n]$.

Clearly,

$$t_{u,v} = \operatorname{cyc}_{u,v} \tag{8}$$

for any two distinct elements u and v of $[n]$. Applying this to $u = 1$ and $v = i$, we obtain $t_{1,i} = \operatorname{cyc}_{1,i}$.

Now, let $\sigma = t_{1,i+1}$. Thus, $\sigma(1) = i+1$ and $\sigma(i) = i$ (since i equals neither 1 nor $i+1$). But Lemma 0.12 (applied to $k = 2$, $i_1 = 1$ and $i_2 = i$) yields

$$\sigma \circ \operatorname{cyc}_{1,i} \circ \sigma^{-1} = \operatorname{cyc}_{\sigma(1), \sigma(i)} = \operatorname{cyc}_{i+1,i} \tag{9}$$

(since $\sigma(1) = i+1$ and $\sigma(i) = i$).

The permutation σ is a transposition (since $\sigma = t_{1,i+1}$), and hence an involution. In other words, $\sigma^{-1} = \sigma$.

But the definition of s_i yields

$$\begin{aligned} s_i &= t_{i,i+1} = \operatorname{cyc}_{i,i+1} && (\text{by (8)}) \\ &= \operatorname{cyc}_{i+1,i} = \underbrace{\sigma}_{=t_{1,i+1}} \circ \underbrace{\operatorname{cyc}_{1,i}}_{=t_{1,i}} \circ \underbrace{\sigma^{-1}}_{=\sigma = t_{1,i+1}} && (\text{by (9)}) \\ &= t_{1,i+1} \circ t_{1,i} \circ t_{1,i+1}. \end{aligned}$$

This proves Lemma 0.11. □

Solution to Exercise 3 (sketched). We first show the following fact:

> *Observation 1:* Let $i \in [n-1]$. Then, s_i can be written as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$.

[*Proof of Observation 1:* If $i > 1$, then this follows immediately from Lemma 0.11. Thus, for the rest of this proof, we WLOG assume that we don't have $i > 1$. Hence, $i = 1$. Thus, $s_i = s_1 = t_{1,2}$ (by the definition of s_1). Thus, again it is clear that s_i can be written as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$. This proves Observation 1.]

Now, let $\sigma \in S_n$ be a permutation. We want to write σ as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$.

First write σ as a composition of some of the transpositions $s_1, s_2, \ldots, s_{n-1}$. (This is possible according to Lemma 0.10.) Next, write each of these transpositions $s_1, s_2, \ldots, s_{n-1}$ as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$. (This is possible according to Observation 1.) The resulting expression is now a representation of σ as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$.

Now, forget that we fixed σ. We thus have shown that each $\sigma \in S_n$ has a representation as a composition of some of the transpositions $t_{1,2}, t_{1,3}, \ldots, t_{1,n}$. This solves Exercise 3. $\square$

0.4. V-permutations as products of cycles

Recall the following notation:

Definition 0.13. Let X be a set. Let k be a positive integer. Let $i_1, i_2, \ldots, i_k$ be k distinct elements of X. We define $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ to be the permutation of X that sends $i_1, i_2, \ldots, i_k$ to $i_2, i_3, \ldots, i_k, i_1$, respectively, while leaving all other elements of X fixed. In other words, we define $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ to be the permutation of X given by

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = \begin{cases} i_{j+1}, & \text{if } p = i_j \text{ for some } j \in \{1, 2, \ldots, k\}; \\ p, & \text{otherwise} \end{cases} \qquad \text{for every } p \in X,$$

where i_{k+1} means i_1.

Exercise 4. Let $n \in \mathbb{N}$. For each $r \in [n]$, let c_r denote the permutation $\operatorname{cyc}_{r,r-1,\ldots,2,1} \in S_n$. (Thus, $c_1 = \operatorname{cyc}_1 = \mathrm{id}$ and $c_2 = \operatorname{cyc}_{2,1} = s_1$.)

Let $G = \{g_1 < g_2 < \cdots < g_p\}$ be a subset of $[n]$. (The notation "$G = \{g_1 < g_2 < \cdots < g_p\}$" is simultaneously saying that $G = \{g_1, g_2, \ldots, g_p\}$ and that $g_1 < g_2 < \cdots < g_p$.)

Let $\sigma \in S_n$ be the permutation $c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_p}$.

Prove the following:

(a) We have $\sigma(1) > \sigma(2) > \cdots > \sigma(p)$.

(b) We have $\sigma([p]) = G$.

(c) We have $\sigma(p+1) < \sigma(p+2) < \cdots < \sigma(n)$.

(Note that a chain of inequalities that involves less than two numbers is considered to be vacuously true. For example, Exercise 4 **(c)** is vacuously true when $p = n - 1$ and also when $p = n$.)

Solution to Exercise 4 (sketched). For each $r \in [n]$ and $i \in [n]$, we have

$$c_r(i) = \begin{cases} r, & \text{if } i = 1; \\ i-1, & \text{if } 1 < i \leq r; \\ i, & \text{if } i > r \end{cases} \tag{10}$$

(by the definition of c_r). Thus, each $r \in [n]$ satisfies

$$c_r(2) < c_r(3) < \cdots < c_r(n) \tag{11}$$

(because the one-line notation of the permutation c_r is $(r, 1, 2, \ldots, r-1, r+1, r+2, \ldots, n)$, which shows immediately that c_r is strictly increasing on the set $\{2, 3, \ldots, n\}$). Moreover, each $r \in [n]$ and $i \in [n]$ satisfy

$$c_r(i) \geq i - 1. \tag{12}$$

(This is easy to check using (10).)

We have $g_1 < g_2 < \cdots < g_p$. Define a further integer g_0 by $g_0 = 0$. Then, the chain of inequalities $g_1 < g_2 < \cdots < g_p$ can be extended to $g_0 < g_1 < \cdots < g_p$ (since each of $g_1, g_2, \ldots, g_p$ is $> 0 = g_0$).

For each $q \in \{0, 1, \ldots, p\}$, we let σ_q denote the permutation $c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_q} \in S_n$. Thus,

$$\sigma_0 = c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_0} = (\text{empty composition of permutations}) = \text{id}$$

and

$$\sigma_p = c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_p} = \sigma.$$

Notice that

$$\sigma_q = \sigma_{q-1} \circ c_{g_q} \qquad \text{for each } q \in [p] \tag{13}$$

[5].

Now, we claim the following:

Observation 1: For each $q \in \{0, 1, \ldots, p\}$, the following holds:

(a) We have $\sigma_q(i) = g_{q+1-i}$ for each $i \in [q]$.

(b) We have $\sigma_q(j) = j$ for each $j \in [n]$ satisfying $j > g_q$.

(c) We have $\sigma_q(q+1) < \sigma_q(q+2) < \cdots < \sigma_q(n)$.

(d) We have $g_q \geq q$.

[5] *Proof of (13):* Let $q \in [p]$. Then, the definition of σ_{q-1} yields $\sigma_{q-1} = c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_{q-1}}$. But the definition of σ_q yields

$$\sigma_q = c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_q} = \underbrace{\left(c_{g_1} \circ c_{g_2} \circ \cdots \circ c_{g_{q-1}}\right)}_{=\sigma_{q-1}} \circ c_{g_q} = \sigma_{q-1} \circ c_{g_q}.$$

This proves (13).

[*Proof of Observation 1:* We shall prove Observation 1 by induction:[6]

Induction base: Let us prove Observation 1 for $q = 0$. To do so, we must prove the following four statements:

(a$_0$) We have $\sigma_0(i) = g_{0+1-i}$ for each $i \in [0]$.

(b$_0$) We have $\sigma_0(j) = j$ for each $j \in [n]$ satisfying $j > g_0$.

(c$_0$) We have $\sigma_0(0+1) < \sigma_0(0+2) < \cdots < \sigma_0(n)$.

(d$_0$) We have $g_0 \geq 0$.

But all of these four statements are obvious. Indeed, **(a$_0$)** is vacuously true (since there exist no $i \in [0]$); furthermore, **(b$_0$)** and **(c$_0$)** are obvious (since $\sigma_0 = \mathrm{id}$); finally, **(d$_0$)** follows from $g_0 = 0$. Thus, Observation 1 has been proven for $q = 0$. This completes the induction base.

Induction step: Let $h \in [p]$. Assume that Observation 1 holds for $q = h - 1$. We must now prove that Observation 1 holds for $q = h$.

We have assumed that Observation 1 holds for $q = h - 1$. In other words, the following four statements hold:

(a$_1$) We have $\sigma_{h-1}(i) = g_{(h-1)+1-i}$ for each $i \in [h-1]$.

(b$_1$) We have $\sigma_{h-1}(j) = j$ for each $j \in [n]$ satisfying $j > g_{h-1}$.

(c$_1$) We have $\sigma_{h-1}(h) < \sigma_{h-1}(h+1) < \cdots < \sigma_{h-1}(n)$.

(d$_1$) We have $g_{h-1} \geq h - 1$.

We must prove that Observation 1 holds for $q = h$. In other words, we must prove the following four statements:

(a$_2$) We have $\sigma_h(i) = g_{h+1-i}$ for each $i \in [h]$.

(b$_2$) We have $\sigma_h(j) = j$ for each $j \in [n]$ satisfying $j > g_h$.

(c$_2$) We have $\sigma_h(h+1) < \sigma_h(h+2) < \cdots < \sigma_h(n)$.

(d$_2$) We have $g_h \geq h$.

[6] It is rather important to prove the four parts of Observation 1 **together**, rather than trying to prove them separately. This way, they can "lend each other a hand" in the induction step (as we will see below).

Recall that $g_0 < g_1 < \cdots < g_p$. Thus, $g_{h-1} < g_h$, so that $g_h > g_{h-1}$. Thus, $g_h \geq g_{h-1} + 1$ (since g_h and g_{h-1} are integers). But **(d₁)** yields $g_{h-1} \geq h - 1$, so that $g_{h-1} + 1 \geq h$. Hence, $g_h \geq g_{h-1} + 1 \geq h$. This proves statement **(d₂)**.

Let $r = g_h$. Thus, $r \in [n]$ (since $h \in [p]$ and thus $g_h \in [n]$). Applying (13) to $q = h$, we obtain

$$\sigma_h = \sigma_{h-1} \circ c_{g_h} = \sigma_{h-1} \circ c_r \qquad (\text{since } g_h = r).$$

Statement **(c₂)** is easy to derive from statement **(c₁)** with the help of (11)[7].

Statement **(b₂)** easily follows from statement **(b₁)** with the help of (10)[8].

Applying (10) to $i = 1$, we obtain $c_r(1) = r = g_h > g_{h-1}$. Hence, statement **(b₁)** (applied to $j = c_r(1)$) yields $\sigma_{h-1}(c_r(1)) = c_r(1) = g_h$. But from $\sigma_h = \sigma_{h-1} \circ c_r$, we obtain

$$\sigma_h(1) = (\sigma_{h-1} \circ c_r)(1) = \sigma_{h-1}(c_r(1)) = g_h = g_{h+1-1}. \tag{14}$$

Finally, statement **(a₂)** can be derived from statement **(a₁)** using (14)[9].

---

[7] *Proof.* We want to prove statement **(c₂)**. In other words, we want to prove that $\sigma_h(h+1) < \sigma_h(h+2) < \cdots < \sigma_h(n)$. In other words, we want to prove that $\sigma_h(k) < \sigma_h(k+1)$ for each $k \in \{h+1, h+2, \ldots, n-1\}$. So let us fix $k \in \{h+1, h+2, \ldots, n-1\}$. We must prove $\sigma_h(k) < \sigma_h(k+1)$.

We have $k \in \{h+1, h+2, \ldots, n-1\}$. Thus, $k \geq h+1 \geq 2$ (since $h \geq 1$). But (11) yields $c_r(2) < c_r(3) < \cdots < c_r(n)$. Thus, $c_r(k) < c_r(k+1)$ (since $k \geq 2$).

Also, (12) yields $c_r(k) \geq k - 1 \geq h$ (since $k \geq h+1$). Thus, $c_r(k) \in \{h, h+1, \ldots, n\}$.

Also, (12) yields $c_r(k+1) \geq (k+1) - 1 = k \geq k - 1 \geq h$. Thus, $c_r(k+1) \in \{h, h+1, \ldots, n\}$.

Statement **(c₁)** says that the map σ_{h-1} is strictly increasing on the set $\{h, h+1, \ldots, n\}$. In other words, if u and v are two elements of $\{h, h+1, \ldots, n\}$ satisfying $u < v$, then $\sigma_{h-1}(u) < \sigma_{h-1}(v)$. Applying this to $u = c_r(k)$ and $v = c_r(k+1)$, we obtain $\sigma_{h-1}(c_r(k)) < \sigma_{h-1}(c_r(k+1))$ (since $c_r(k) < c_r(k+1)$, and since both $c_r(k)$ and $c_r(k+1)$ are elements of $\{h, h+1, \ldots, n\}$).

But $\sigma_h = \sigma_{h-1} \circ c_r$, and thus

$$\sigma_h(k) = (\sigma_{h-1} \circ c_r)(k) = \sigma_{h-1}(c_r(k)) < \sigma_{h-1}(c_r(k+1)) = \underbrace{(\sigma_{h-1} \circ c_r)}_{=\sigma_h}(k+1) = \sigma_h(k+1).$$

This completes our proof of statement **(c₂)**.

[8] *Proof.* Let $j \in [n]$ be such that $j > g_h$. We want to show that $\sigma_h(j) = j$.

We have $j > g_h = r$. Thus, (10) (applied to $i = j$) simplifies to $c_r(j) = j$. But $j > g_h > g_{h-1}$; therefore, statement **(b₁)** yields $\sigma_{h-1}(j) = j$.

Now, recall that $\sigma_h = \sigma_{h-1} \circ c_r$. Hence,

$$\sigma_h(j) = (\sigma_{h-1} \circ c_r)(j) = \sigma_{h-1}\left(\underbrace{c_r(j)}_{=j}\right) = \sigma_{h-1}(j) = j.$$

This proves statement **(b₂)**.

[9] *Proof.* Let us prove statement **(a₂)**. In other words, let us prove that $\sigma_h(i) = g_{h+1-i}$ for each $i \in [h]$.

Indeed, let $i \in [h]$. We must prove that $\sigma_h(i) = g_{h+1-i}$.

If $i = 1$, then this follows from (14). Hence, for the rest of this proof, we WLOG assume that $i \neq 1$. Thus, $i > 1$. Combined with $i \in [h]$, this yields $i \in \{2, 3, \ldots, h\}$, so that $i - 1 \in [h-1]$. Therefore, statement **(a₁)** (applied to $i - 1$ instead of i) yields $\sigma_{h-1}(i-1) = g_{(h-1)+1-(i-1)} = g_{h+1-i}$ (since $(h-1) + 1 - (i-1) = h + 1 - i$).

But $i \in \{2, 3, \ldots, h\}$, so that $1 < i \leq h \leq r$ (because $r = g_h \geq h$). The equality (10) simplifies to

We have now proven all four statements **(a$_2$)**, **(b$_2$)**, **(c$_2$)** and **(d$_2$)**. Thus, Observation 1 holds for $q = h$. This completes the induction step; thus, Observation 1 is proven.]

Now, we can apply Observation 1 to $q = p$. As a result, we obtain the following four statements:

(a$_3$) We have $\sigma_p(i) = g_{p+1-i}$ for each $i \in [p]$.

(b$_3$) We have $\sigma_p(j) = j$ for each $j \in [n]$ satisfying $j > g_p$.

(c$_3$) We have $\sigma_p(p+1) < \sigma_p(p+2) < \cdots < \sigma_p(n)$.

(d$_3$) We have $g_p \geq p$.

Statement **(c$_3$)** says that $\sigma_p(p+1) < \sigma_p(p+2) < \cdots < \sigma_p(n)$. In view of $\sigma_p = \sigma$, this rewrites as $\sigma(p+1) < \sigma(p+2) < \cdots < \sigma(n)$. This solves Exercise 4 **(c)**.

Statement **(a$_3$)** says that $\sigma_p(i) = g_{p+1-i}$ for each $i \in [p]$. In view of $\sigma_p = \sigma$, this rewrites as

$$\sigma(i) = g_{p+1-i} \qquad \text{for each } i \in [p]. \tag{15}$$

In other words,

$$(\sigma(1), \sigma(2), \ldots, \sigma(p)) = (g_p, g_{p-1}, \ldots, g_1). \tag{16}$$

Hence,

$$\{\sigma(1), \sigma(2), \ldots, \sigma(p)\} = \{g_p, g_{p-1}, \ldots, g_1\} = \{g_1, g_2, \ldots, g_p\} = G$$

(since $G = \{g_1 < g_2 < \cdots < g_p\} = \{g_1, g_2, \ldots, g_p\}$). Hence

$$\sigma\left(\underbrace{[p]}_{=\{1,2,\ldots,p\}}\right) = \sigma(\{1,2,\ldots,p\}) = \{\sigma(1), \sigma(2), \ldots, \sigma(p)\} = G.$$

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

Finally, recall that $g_1 < g_2 < \cdots < g_p$. In other words, $g_p > g_{p-1} > \cdots > g_1$. In view of (15), this rewrites as follows: $\sigma(1) > \sigma(2) > \cdots > \sigma(p)$. This solves Exercise 4 **(a)**. $\square$

$c_r(i) = i - 1$ (since $1 < i \leq r$). Now, recall that $\sigma_h = \sigma_{h-1} \circ c_r$. Thus,

$$\sigma_h(i) = (\sigma_{h-1} \circ c_r)(i) = \sigma_{h-1}\left(\underbrace{c_r(i)}_{=i-1}\right) = \sigma_{h-1}(i-1) = g_{h+1-i}.$$

Thus, $\sigma_h(i) = g_{h+1-i}$ is proven, as we wanted. This completes the proof of statement **(a$_2$)**.

Permutations $\sigma \in S_n$ satisfying the inequalities $\sigma(1) > \sigma(2) > \cdots > \sigma(p)$ and $\sigma(p+1) < \sigma(p+2) < \cdots < \sigma(n)$ for some $p \in \{0, 1, \ldots, n\}$ are known as "V-permutations" (as their plot looks somewhat like the letter "V": first decreasing for a while, then increasing). Can you guess how permutations $\sigma \in S_n$ satisfying $\sigma(1) < \sigma(2) < \cdots < \sigma(p)$ and $\sigma(p+1) > \sigma(p+2) > \cdots > \sigma(n)$ are called?[10]

Exercise 4 is a lemma in the theory of free Lie algebras (see [BleLau92, (10)]).

TODO: Explain how Exercise 4 can also be obtained as a particular case of the formula for a permutation in terms of its Rothe diagram. (See `https://sumidiot.blogspot.com/2008/05/rothe-diagram.html` for now.)

0.5. Lexicographic comparison of permutations

Definition 0.14. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ be a permutation. For any $i \in [n]$, we let $\ell_i(\sigma)$ denote the number of $j \in \{i+1, i+2, \ldots, n\}$ such that $\sigma(i) > \sigma(j)$.

For example, if σ is the permutation of $[5]$ written in one-line notation as $[4, 1, 5, 2, 3]$, then $\ell_1(\sigma) = 3$, $\ell_2(\sigma) = 0$, $\ell_3(\sigma) = 2$, $\ell_4(\sigma) = 0$ and $\ell_5(\sigma) = 0$.

Definition 0.15. Let $n \in \mathbb{N}$. Let $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ be two n-tuples of integers. We say that $(a_1, a_2, \ldots, a_n) <_{\text{lex}} (b_1, b_2, \ldots, b_n)$ if and only if there exists some $k \in [n]$ such that $a_k \neq b_k$, and the **smallest** such k satisfies $a_k < b_k$.

For example, $(4, 1, 2, 5) <_{\text{lex}} (4, 1, 3, 0)$ and $(1, 1, 0, 1) <_{\text{lex}} (2, 0, 0, 0)$. The relation $<_{\text{lex}}$ is usually pronounced "is lexicographically smaller than"; the word "lexicographic" comes from the idea that if numbers were letters, then a "word" $a_1 a_2 \cdots a_n$ would appear earlier in a dictionary than $b_1 b_2 \cdots b_n$ if and only if $(a_1, a_2, \ldots, a_n) <_{\text{lex}} (b_1, b_2, \ldots, b_n)$.

Exercise 5. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ and $\tau \in S_n$. Prove the following:

(a) If

$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) <_{\text{lex}} (\tau(1), \tau(2), \ldots, \tau(n)),$$

then

$$(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) <_{\text{lex}} (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau)).$$

(b) If $(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) = (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau))$, then $\sigma = \tau$.

The solution to Exercise 5 given below is one of those cases where a simple argument becomes insufferably long and dreary as I try to capture it in writing. Apologies for what you are about to see. The proof relies on the following lemma:

[10] *Answer:* They are called "Λ-permutations". Both names "V-permutations" and "Λ-permutations" are due to the shape of the plot when the permutation is plotted in 2D.

Lemma 0.16. Let $n \in \mathbb{N}$. Let $\sigma \in S_n$ and $i \in [n]$. Then:

(a) We have $\ell_i(\sigma) = |[\sigma(i)-1] \setminus \sigma([i])|$.

(b) We have $\ell_i(\sigma) = |[\sigma(i)-1] \setminus \sigma([i-1])|$.

Proof of Lemma 0.16. **(a)** We know that $\ell_i(\sigma)$ is the number of $j \in \{i+1, i+2, \ldots, n\}$ such that $\sigma(i) > \sigma(j)$ (by the definition of $\ell_i(\sigma)$). Hence,

$$\begin{aligned} \ell_i(\sigma) &= (\text{the number of } j \in \{i+1, i+2, \ldots, n\} \text{ such that } \sigma(i) > \sigma(j)) \\ &= |\{j \in \{i+1, i+2, \ldots, n\} \mid \sigma(i) > \sigma(j)\}|. \end{aligned} \tag{17}$$

Define a set A by

$$A = \{j \in \{i+1, i+2, \ldots, n\} \mid \sigma(i) > \sigma(j)\}. \tag{18}$$

Thus,

$$|A| = |\{j \in \{i+1, i+2, \ldots, n\} \mid \sigma(i) > \sigma(j)\}| = \ell_i(\sigma) \tag{19}$$

(by (17)).

Let B be the set $[\sigma(i)-1] \setminus \sigma([i])$.

The map σ is a permutation of $[n]$ (since $\sigma \in S_n$), and thus is invertible, and therefore is injective.

For each $k \in A$, we have $\sigma(k) \in B$ [11]. Hence, we can define a map $\alpha : A \to B$ by

$$(\alpha(k) = \sigma(k) \qquad \text{for each } k \in A).$$

Consider this α.

On the other hand, for each $k \in B$, we have $\sigma^{-1}(k) \in A$ [12]. Hence, we can define a map $\beta : B \to A$ by

$$\left(\beta(k) = \sigma^{-1}(k) \qquad \text{for each } k \in B\right).$$

[11] *Proof.* Let $k \in A$. Thus, $k \in A = \{j \in \{i+1, i+2, \ldots, n\} \mid \sigma(i) > \sigma(j)\}$. In other words, k is an element of $\{i+1, i+2, \ldots, n\}$ and satisfies $\sigma(i) > \sigma(k)$.

From $k \in \{i+1, i+2, \ldots, n\} \subseteq [n]$, we conclude that $\sigma(k)$ is well-defined. Also, $\sigma(k) < \sigma(i)$ (since $\sigma(i) > \sigma(k)$), so that $\sigma(k) \leq \sigma(i) - 1$ (since $\sigma(k)$ and $\sigma(i)$ are integers). Thus, $\sigma(k) \in [\sigma(i)-1]$.

Next, let us prove that $\sigma(k) \notin \sigma([i])$.

Indeed, assume the contrary (for the sake of contradiction). Hence, $\sigma(k) \in \sigma([i])$. In other words, there exists some $j \in [i]$ such that $\sigma(k) = \sigma(j)$. Consider this j. From $\sigma(k) = \sigma(j)$, we obtain $k = j$ (since the map σ is injective). Hence, $k = j \in [i]$. But $k \in \{i+1, i+2, \ldots, n\} = [n] \setminus [i]$, so that $k \notin [i]$. This contradicts $k \in [i]$. This contradiction shows that our assumption was false. Hence, $\sigma(k) \notin \sigma([i])$ is proven.

Combining $\sigma(k) \in [\sigma(i)-1]$ with $\sigma(k) \notin \sigma([i])$, we obtain $\sigma(k) \in [\sigma(i)-1] \setminus \sigma([i]) = B$. Qed.

[12] *Proof.* Let $k \in B$. Thus, $k \in B = [\sigma(i)-1] \setminus \sigma([i])$. In other words, $k \in [\sigma(i)-1]$ and $k \notin \sigma([i])$.

From $k \in [\sigma(i)-1]$, we obtain $1 \leq k \leq \sigma(i) - 1$. Also, $k \in [\sigma(i)-1] \subseteq [n]$, so that $\sigma^{-1}(k)$ is a well-defined element of $[n]$.

We have $\sigma\left(\sigma^{-1}(k)\right) = k \leq \sigma(i) - 1 < \sigma(i)$. In other words, $\sigma(i) > \sigma\left(\sigma^{-1}(k)\right)$.

Next, we claim that $\sigma^{-1}(k) \in \{i+1, i+2, \ldots, n\}$. Indeed, assume the contrary (for the sake of contradiction). Thus, $\sigma^{-1}(k) \notin \{i+1, i+2, \ldots, n\}$. Combining this with $\sigma^{-1}(k) \in [n]$, we obtain

$$\sigma^{-1}(k) \in [n] \setminus \{i+1, i+2, \ldots, n\} = [i].$$

Consider this β.

The maps α and β are mutually inverse (since α is a restriction of σ, whereas β is a restriction of σ^{-1}), and therefore are bijections. Hence, there is a bijection from A to B (namely, α). Thus, $|A| = |B|$.

But (19) yields
$$\ell_i(\sigma) = |A| = |B| = |[\sigma(i) - 1] \setminus \sigma([i])|$$
(since $B = [\sigma(i) - 1] \setminus \sigma([i])$). This proves Lemma 0.16 **(a)**.

(b) If we had $\sigma(i) \in [\sigma(i) - 1]$, then we would have $\sigma(i) \leq \sigma(i) - 1 < \sigma(i)$, which would be absurd. Hence, we have $\sigma(i) \notin [\sigma(i) - 1]$.

But $[i] = \{i\} \cup [i-1]$. Hence,
$$\sigma\left(\underbrace{[i]}_{=\{i\}\cup[i-1]}\right) = \sigma(\{i\} \cup [i-1]) = \underbrace{\sigma(\{i\})}_{=\{\sigma(i)\}} \cup \sigma([i-1]) = \{\sigma(i)\} \cup \sigma([i-1]).$$

Thus,
$$\begin{aligned}
&[\sigma(i) - 1] \setminus \underbrace{\sigma([i])}_{=\{\sigma(i)\}\cup\sigma([i-1])} \\
&= [\sigma(i) - 1] \setminus (\{\sigma(i)\} \cup \sigma([i-1])) \\
&= \underbrace{([\sigma(i) - 1] \setminus \{\sigma(i)\})}_{\substack{=[\sigma(i)-1] \\ (\text{since } \sigma(i)\notin[\sigma(i)-1])}} \setminus \sigma([i-1]) = [\sigma(i) - 1] \setminus \sigma([i-1]).
\end{aligned}$$

Now, Lemma 0.16 **(a)** yields
$$\ell_i(\sigma) = \left|\underbrace{[\sigma(i) - 1] \setminus \sigma([i])}_{=[\sigma(i)-1]\setminus\sigma([i-1])}\right| = |[\sigma(i) - 1] \setminus \sigma([i-1])|.$$

This proves Lemma 0.16 **(b)**. □

Solution to Exercise 5 (sketched). **(a)** Assume that
$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) <_{\text{lex}} (\tau(1), \tau(2), \ldots, \tau(n)).$$

Hence, $k = \sigma\left(\underbrace{\sigma^{-1}(k)}_{\in[i]}\right) \in \sigma([i])$, which contradicts $k \notin \sigma([i])$. This contradiction shows that our assumption was false. Thus, $\sigma^{-1}(k) \in \{i+1, i+2, \ldots, n\}$ is proven.

Now, we know that $\sigma^{-1}(k) \in \{i+1, i+2, \ldots, n\}$ and $\sigma(i) > \sigma(\sigma^{-1}(k))$. In other words, $\sigma^{-1}(k)$ is a $j \in \{i+1, i+2, \ldots, n\}$ satisfying $\sigma(i) > \sigma(j)$. In other words,
$$\sigma^{-1}(k) \in \{j \in \{i+1, i+2, \ldots, n\} \mid \sigma(i) > \sigma(j)\}.$$

In view of (18), this rewrites as $\sigma^{-1}(k) \in A$. Qed.

According to Definition 0.15, this means the following: There exists some $k \in [n]$ such that $\sigma(k) \neq \tau(k)$, and the **smallest** such k satisfies $\sigma(k) < \tau(k)$.

Let i be the smallest such k. Thus, $\sigma(i) < \tau(i)$, but

$$\text{each } k \in [i-1] \text{ satisfies } \sigma(k) = \tau(k) \tag{20}$$

(since i is the **smallest** $k \in [n]$ such that $\sigma(k) \neq \tau(k)$).

Thus,

$$\text{each } k \in [i] \text{ satisfies } \sigma([k-1]) = \tau([k-1]) \tag{21}$$

[13]. Hence,

$$\text{each } k \in [i-1] \text{ satisfies } \ell_k(\sigma) = \ell_k(\tau) \tag{22}$$

[14]. Furthermore,

$$\ell_i(\sigma) < \ell_i(\tau) \tag{23}$$

[15]. Thus, $\ell_i(\sigma) \neq \ell_i(\tau)$. In other words, i is a $k \in [n]$ such that $\ell_k(\sigma) \neq \ell_k(\tau)$. Moreover, (22) shows that i is the **smallest** such k. Thus, the smallest $k \in [n]$ such that $\ell_k(\sigma) \neq \ell_k(\tau)$ satisfies $\ell_k(\sigma) < \ell_k(\tau)$ (because this k is i, and i satisfies (23)).

---

[13]*Proof of (21):* Let $k \in [i]$. Thus, $k \leq i$.

Let $j \in [k-1]$. Thus, $j \leq \underbrace{k}_{\leq i} -1 \leq i-1$, so that $j \in [i-1]$. Hence, (20) (applied to j instead of k) shows that $\sigma(j) = \tau(j)$.

Now, forget that we fixed j. We thus have shown that $\sigma(j) = \tau(j)$ for each $j \in [k-1]$. In other words,

$$(\sigma(1), \sigma(2), \ldots, \sigma(k-1)) = (\tau(1), \tau(2), \ldots, \tau(k-1)).$$

Thus,

$$\{\sigma(1), \sigma(2), \ldots, \sigma(k-1)\} = \{\tau(1), \tau(2), \ldots, \tau(k-1)\}.$$

Now,

$$\sigma\left(\underbrace{[k-1]}_{=\{1,2,\ldots,k-1\}}\right) = \sigma(\{1,2,\ldots,k-1\}) = \{\sigma(1), \sigma(2), \ldots, \sigma(k-1)\}$$
$$= \{\tau(1), \tau(2), \ldots, \tau(k-1)\} = \tau\left(\underbrace{\{1,2,\ldots,k-1\}}_{=[k-1]}\right) = \tau([k-1]).$$

This proves (21).

[14]*Proof of (22):* Let $k \in [i-1]$. Then, Lemma 0.16 **(b)** (applied to k instead of i) yields $\ell_k(\sigma) = |[\sigma(k)-1] \setminus \sigma([k-1])|$. The same argument (applied to τ instead of σ) yields $\ell_k(\tau) = |[\tau(k)-1] \setminus \tau([k-1])|$.

But $k \in [i-1] \subseteq [i]$. Hence, (21) yields $\sigma([k-1]) = \tau([k-1])$. Also, (20) yields $\sigma(k) = \tau(k)$. Hence,

$$\ell_k(\sigma) = \left|\left[\underbrace{\sigma(k)}_{=\tau(k)} -1\right] \setminus \underbrace{\sigma([k-1])}_{=\tau([k-1])}\right| = |[\tau(k)-1] \setminus \tau([k-1])| = \ell_k(\tau).$$

This proves (22).

[15]*Proof of (23):* We have $i \in [i]$. Hence, (21) (applied to $k = i$) yields $\sigma([i-1]) = \tau([i-1])$.

Thus, we have shown that there exists some $k \in [n]$ such that $\ell_k(\sigma) \neq \ell_k(\tau)$, and the **smallest** such k satisfies $\ell_k(\sigma) < \ell_k(\tau)$. But this means precisely that

$$(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) <_{\text{lex}} (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau))$$

(according to Definition 0.15). Hence, Exercise 5 **(a)** is solved.

(b) Assume that

$$(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) = (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau)). \tag{25}$$

We must prove that $\sigma = \tau$.

Indeed, assume the contrary. Thus, $\sigma \neq \tau$. Hence, there exists some $k \in [n]$ satisfying $\sigma(k) \neq \tau(k)$. Therefore, there exists the **smallest** such k. This smallest k must satisfy either $\sigma(k) < \tau(k)$ or $\sigma(k) > \tau(k)$ (because it satisfies $\sigma(k) \neq \tau(k)$). We can WLOG assume that it satisfies $\sigma(k) < \tau(k)$ (because otherwise, we can simply switch the roles of σ and τ). Assume this. Thus, $(\sigma(1), \sigma(2), \ldots, \sigma(n)) <_{\text{lex}} (\tau(1), \tau(2), \ldots, \tau(n))$ (because of Definition 0.15). Hence, Exercise 5 **(a)** shows that $(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) <_{\text{lex}} (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau))$. In other words, there

---

Also, $\sigma(i) < \tau(i)$, so that $\sigma(i) - 1 < \tau(i) - 1$ and therefore $[\sigma(i) - 1] \subseteq [\tau(i) - 1]$.

From $\sigma(i) < \tau(i)$, we also obtain $\sigma(i) \leq \tau(i) - 1$ (since $\sigma(i)$ and $\tau(i)$ are integers), and thus $\sigma(i) \in [\tau(i) - 1]$.

Also, $\sigma(i) \notin \sigma([i-1])$. [*Proof:* Assume the contrary. Thus, $\sigma(i) \in \sigma([i-1])$. In other words, $\sigma(i) = \sigma(j)$ for some $j \in [i-1]$. Consider this j. From $\sigma(i) = \sigma(j)$, we obtain $i = j$ (since σ is injective), so that $i = j \in [i-1]$ and thus $i \leq i - 1 < i$. But this is absurd. Hence, we found a contradiction, so that $\sigma(i) \notin \sigma([i-1])$ is proven.]

If we had $\sigma(i) \in [\sigma(i) - 1]$, then we would have $\sigma(i) \leq \sigma(i) - 1 < \sigma(i)$, which is absurd. Hence, we have $\sigma(i) \notin [\sigma(i) - 1]$. Thus, also $\sigma(i) \notin [\sigma(i) - 1] \setminus \sigma([i-1])$.

Combining $\sigma(i) \in [\tau(i) - 1]$ with $\sigma(i) \notin \sigma([i-1])$, we obtain $\sigma(i) \in [\tau(i) - 1] \setminus \sigma([i-1])$.

Now,

$$\underbrace{[\sigma(i) - 1]}_{\subseteq [\tau(i) - 1]} \setminus \sigma([i-1]) \subseteq [\tau(i) - 1] \setminus \sigma([i-1]). \tag{24}$$

Moreover, the set $[\tau(i) - 1] \setminus \sigma([i-1])$ contains $\sigma(i)$ (since $\sigma(i) \in [\tau(i) - 1] \setminus \sigma([i-1])$), but the set $[\sigma(i) - 1] \setminus \sigma([i-1])$ does not (since $\sigma(i) \notin [\sigma(i) - 1] \setminus \sigma([i-1])$). Thus, these two sets are distinct. In other words, $[\sigma(i) - 1] \setminus \sigma([i-1]) \neq [\tau(i) - 1] \setminus \sigma([i-1])$. Combining this with (24), we conclude that $[\sigma(i) - 1] \setminus \sigma([i-1])$ is a **proper** subset of $[\tau(i) - 1] \setminus \sigma([i-1])$. Thus,

$$|[\sigma(i) - 1] \setminus \sigma([i-1])| < |[\tau(i) - 1] \setminus \sigma([i-1])|$$

(since a proper subset of any finite set must always have smaller size than the latter).

But Lemma 0.16 **(b)** yields $\ell_i(\sigma) = |[\sigma(i) - 1] \setminus \sigma([i-1])|$. The same argument (applied to τ instead of σ) yields $\ell_i(\tau) = |[\tau(i) - 1] \setminus \tau([i-1])|$. Hence,

$$\begin{aligned} \ell_i(\sigma) &= |[\sigma(i) - 1] \setminus \sigma([i-1])| \\ &< \left| [\tau(i) - 1] \setminus \underbrace{\sigma([i-1])}_{= \tau([i-1])} \right| = |[\tau(i) - 1] \setminus \tau([i-1])| = \ell_i(\tau). \end{aligned}$$

This proves (23).

exists some $k \in [n]$ such that $\ell_k(\sigma) \neq \ell_k(\tau)$, and the **smallest** such k satisfies $\ell_k(\sigma) < \ell_k(\tau)$ (according to Definition 0.15).

In particular, there exists some $k \in [n]$ such that $\ell_k(\sigma) \neq \ell_k(\tau)$. In other words, $(\ell_1(\sigma), \ell_2(\sigma), \ldots, \ell_n(\sigma)) \neq (\ell_1(\tau), \ell_2(\tau), \ldots, \ell_n(\tau))$. But this contradicts (25). This contradiction shows that our assumption was false. Hence, $\sigma = \tau$ is proven. This solves Exercise 5 **(b)**. $\square$

0.6. Comparing subsets of $[n]$

If I and J are two finite sets of integers, then we write $I \leq_{\#} J$ if and only if the following two properties hold:

- We have $|I| \geq |J|$.
- For every $r \in \{1, 2, \ldots, |J|\}$, the r-th smallest element of I is $\leq$ to the r-th smallest element of J.

For example, $\{2, 4\} \leq_{\#} \{2, 5\}$ and $\{1, 3\} \leq_{\#} \{2, 4\}$ and $\{1, 3, 5\} \leq_{\#} \{2, 4\}$. (But not $\{1, 3\} \leq_{\#} \{2, 4, 5\}$.)

The relation $\leq_{\#}$ is called the *Gale order* on the powerset of $[n]$.

Exercise 6. Let $n \in \mathbb{N}$. Let I and J be two subsets of $[n]$.

(a) For every subset S of $[n]$ and every $\ell \in [n]$, let $\alpha_S(\ell)$ denote the number of all elements of S that are $\leq \ell$. Prove that $I \leq_{\#} J$ holds if and only if every $\ell \in [n]$ satisfies $\alpha_I(\ell) \geq \alpha_J(\ell)$.

(b) Prove that $I \leq_{\#} J$ if and only if $[n] \setminus J \leq_{\#} [n] \setminus I$.

The following solution is mostly copypasted from [GriRei18, Proof of Proposition 12.75.2], where the exercise serves as a lemma for a combinatorial proof of an identity between Schur polynomials.

Solution to Exercise 6. **(a)** We must prove the equivalence

$$(I \leq_{\#} J) \iff (\text{every } \ell \in [n] \text{ satisfies } \alpha_I(\ell) \geq \alpha_J(\ell)). \tag{26}$$

$\Longrightarrow$: Assume that $I \leq_{\#} J$. In other words, the following two properties hold:

Property α: We have $|I| \geq |J|$.

Property β: For every $r \in \{1, 2, \ldots, |J|\}$, the r-th smallest element of I is $\leq$ to the r-th smallest element of J.

Now, let $\ell \in [n]$. Then, we need to show that $\alpha_I(\ell) \geq \alpha_J(\ell)$. Since this is obvious if $\alpha_J(\ell) = 0$ (because $\alpha_I(\ell) \geq 0$), we can WLOG assume that $\alpha_J(\ell) \neq 0$. Assume this. Thus, $\alpha_J(\ell) \geq 1$. Also, $\alpha_J(\ell) = \left|\underbrace{\{s \in J \mid s \leq \ell\}}_{\subseteq J}\right| \leq |J| \leq |I|$ (since $|I| \geq |J|$).

Hence, both the $\alpha_J(\ell)$-th smallest element of J and the $\alpha_J(\ell)$-th smallest element of I are well-defined.

Since $\alpha_J(\ell) = |\{s \in J \mid s \leq \ell\}|$, we know that the elements of J which are $\leq \ell$ are precisely the $\alpha_J(\ell)$ smallest elements of J. Thus,

$$(\text{the } \alpha_J(\ell)\text{-th smallest element of } J) = (\text{the largest element of } J \text{ which is } \leq \ell).$$

But by Property β (applied to $r = \alpha_J(\ell)$), we have

$$\begin{aligned}(\text{the } \alpha_J(\ell)\text{-th smallest element of } I) &\leq (\text{the } \alpha_J(\ell)\text{-th smallest element of } J) \\ &= (\text{the largest element of } J \text{ which is } \leq \ell) \leq \ell.\end{aligned}$$

Hence, there are at least $\alpha_J(\ell)$ elements of I which are $\leq \ell$ (namely, the $\alpha_J(\ell)$ smallest ones). In other words, $|\{s \in I \mid s \leq \ell\}| \geq \alpha_J(\ell)$. Now, the definition of $\alpha_I(\ell)$ yields $\alpha_I(\ell) = |\{s \in I \mid s \leq \ell\}| \geq \alpha_J(\ell)$. We thus have proven the $\Longrightarrow$ direction of (26).

$\Longleftarrow$: Assume that every $\ell \in [n]$ satisfies $\alpha_I(\ell) \geq \alpha_J(\ell)$. We need to prove that $I \leq_\# J$. In other words, we need to prove that the following two properties hold:

Property α: We have $|I| \geq |J|$.

Property β: For every $r \in \{1, 2, \ldots, |J|\}$, the r-th smallest element of I is $\leq$ to the r-th smallest element of J.

First of all, $\{s \in I \mid s \leq n\} = I$ (since every $s \in I$ satisfies $s \leq n$), and the definition of $\alpha_I(n)$ yields $\alpha_I(n) = \left|\underbrace{\{s \in I \mid s \leq n\}}_{=I}\right| = |I|$. Similarly, $\alpha_J(n) = |J|$. Applying $\alpha_I(\ell) \geq \alpha_J(\ell)$ to $\ell = n$, we obtain $\alpha_I(n) \geq \alpha_J(n)$, so that $|I| = \alpha_I(n) \geq \alpha_J(n) = |J|$, and thus Property α is proven.

Now, let $r \in \{1, 2, \ldots, |J|\}$. The r-th smallest element of I and the r-th smallest element of J are then well-defined (because of $r \leq |J| \leq |I|$). Let ℓ be the r-th smallest element of J. Then, $\{s \in J \mid s \leq \ell\}$ is the set consisting of the r smallest elements of J, so that $|\{s \in J \mid s \leq \ell\}| = r$. Now, the definition of $\alpha_J(\ell)$ yields $\alpha_J(\ell) = |\{s \in J \mid s \leq \ell\}| = r$.

But the definition of $\alpha_I(\ell)$ yields $\alpha_I(\ell) = |\{s \in I \mid s \leq \ell\}|$, so that

$$|\{s \in I \mid s \leq \ell\}| = \alpha_I(\ell) \geq \alpha_J(\ell) = r.$$

In other words, there exist at least r elements of I which are $\leq \ell$. Hence, the r-th smallest element of I must be $\leq \ell$. Since ℓ is the r-th smallest element of J, this rewrites as follows: The r-th smallest element of I is $\leq$ to the r-th smallest element of J. Thus, Property β holds. Now we know that both Properties α and β hold. Hence, $I \leq_\# J$ holds (which, as we know, is equivalent to the conjunction of said properties). This proves the $\Longleftarrow$ direction of (26). Thus, (26) is proven. In other words, Exercise 6 **(a)** is solved.

(b) For every $\ell \in [n]$ and $S \subseteq [n]$, let $\alpha_S(\ell)$ denote the number $|\{s \in S \mid s \leq \ell\}|$. Thus, every $\ell \in [n]$ satisfies

$$\begin{aligned}
\alpha_I(\ell) + \alpha_{[n]\setminus I}(\ell) &= |\{s \in I \mid s \leq \ell\}| + |\{s \in [n] \setminus I \mid s \leq \ell\}| \\
&= \left| \left\{ s \in \underbrace{I \cup ([n] \setminus I)}_{=[n]} \mid s \leq \ell \right\} \right| \qquad \text{(since } I \text{ and } [n] \setminus I \text{ are disjoint)} \\
&= |\{s \in [n] \mid s \leq \ell\}| = |\{1, 2, \ldots, \ell\}| = \ell,
\end{aligned}$$

so that $\alpha_{[n]\setminus I}(\ell) = \ell - \alpha_I(\ell)$. Similarly, every $\ell \in [n]$ satisfies $\alpha_{[n]\setminus J}(\ell) = \ell - \alpha_J(\ell)$.

Applying (26) to $[n] \setminus J$ and $[n] \setminus I$ in lieu of I and J, we obtain the equivalence

$$([n] \setminus J \leq_{\#} [n] \setminus I) \iff \left(\text{every } \ell \in [n] \text{ satisfies } \alpha_{[n]\setminus J}(\ell) \geq \alpha_{[n]\setminus I}(\ell) \right).$$

Hence, we have the following chain of equivalences:

$$\begin{aligned}
& ([n] \setminus J \leq_{\#} [n] \setminus I) \\
\iff & \left(\text{every } \ell \in [n] \text{ satisfies } \underbrace{\alpha_{[n]\setminus J}(\ell)}_{=\ell - \alpha_J(\ell)} \geq \underbrace{\alpha_{[n]\setminus I}(\ell)}_{=\ell - \alpha_I(\ell)} \right) \\
\iff & (\text{every } \ell \in [n] \text{ satisfies } \ell - \alpha_J(\ell) \geq \ell - \alpha_I(\ell)) \\
\iff & (\text{every } \ell \in [n] \text{ satisfies } \alpha_I(\ell) \geq \alpha_J(\ell)) \\
\iff & (I \leq_{\#} J) \qquad \text{(by (26))}.
\end{aligned}$$

This solves Exercise 6 **(b)**. □

Remark 0.17. Recall that we have defined a *Dyck word* as a list w of $2n$ numbers, exactly n of which are 0's while the other n are 1's, and having the property that for each $k \in [2n]$, the number of 0's among the first k entries of w is $\leq$ to the number of 1's among the first k entries of w.

It is not hard to see the connection between the relation $\leq_{\#}$ and Dyck words: Let $w = (w_1, w_2, \ldots, w_{2n}) \in \{0,1\}^{2n}$ be a list of $2n$ numbers, exactly n of which are 0's while the other n are 1's. Then, w is a Dyck word if and only if

$$\{i \in [2n] \mid w_i = 1\} \leq_{\#} \{i \in [2n] \mid w_i = 0\}$$

(in other words, for every $r \in [n]$, the r-th appearance of 1 in w precedes the r-th appearance of 0 in w).

0.7. A rigorous approach to the existence of a cycle decomposition

The purpose of the following exercise is to give a rigorous proof of the fact that any permutation can be decomposed into disjoint cycles.

Exercise 7. Let X be a finite set. Let σ be a permutation of X.

Define a binary relation $\sim$ on the set X as follows: For two elements $x \in X$ and $y \in X$, we set $x \sim y$ if and only if there exists some $k \in \mathbb{N}$ such that $y = \sigma^k(x)$.

(a) Prove that $\sim$ is an equivalence relation.

For any $x \in X$, we let $[x]_\sim$ denote the $\sim$-equivalence class of x.

(b) For any $x \in X$, prove that $[x]_\sim = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right\}$, where $k = |[x]_\sim|$.

(c) For any $\sim$-equivalence class E, let us define c_E to be the map

$$X \to X, \qquad x \mapsto \begin{cases} \sigma(x), & \text{if } x \in E; \\ x, & \text{if } x \notin E \end{cases}.$$

Prove that c_E is a permutation of X.

(d) Prove that if $E = [x]_\sim$ for some $x \in X$, then c_E can be written as $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$, where $k = |[x]_\sim|$. (Don't forget to show that $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)$ are distinct, so that $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$ is well-defined.)

(e) Let $E_1, E_2, \ldots, E_m$ be all $\sim$-equivalence classes (listed without repetitions – that is, $E_i \neq E_j$ whenever $i \neq j$). Prove that

$$\sigma = c_{E_1} \circ c_{E_2} \circ \cdots \circ c_{E_m}.$$

Exercise 7 is mostly an exercise in understanding the definitions and writing up proofs. The first two parts of it are similar to Exercise 6 on homework set #3; thus, our solution below is partly copypasted from the latter (with the necessary changes made).

Our solution relies on a few lemmas:

Lemma 0.18. Let X be a set. Let $f : X \to X$ be any map. Let $x \in X$. Let i and j be two nonnegative integers satisfying $i < j$ and $f^i(x) = f^j(x)$. Then,

$$\left\{f^h(x) \mid h \in \mathbb{N}\right\} = \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}.$$

Proof of Lemma 0.18. We have

$$\begin{aligned} \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\} &= \left\{f^h(x) \mid h \in \underbrace{\{0,1,\ldots,j-1\}}_{\subseteq \mathbb{N}}\right\} \\ &\subseteq \left\{f^h(x) \mid h \in \mathbb{N}\right\}. \end{aligned} \tag{27}$$

On the other hand, we have $i \in \{0,1,\ldots,j-1\}$ (since i is a nonnegative integer satisfying $i < j$), and thus $f^i(x) \in \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}$. Hence, $\left\{f^i(x)\right\} \subseteq$

$\left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}$. Therefore,

$$\begin{aligned}\left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\} &= \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\} \cup \left\{\underbrace{f^i(x)}_{=f^j(x)}\right\} \\ &= \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\} \cup \left\{f^j(x)\right\} \\ &= \left\{f^0(x), f^1(x), \ldots, f^j(x)\right\}. \qquad (28)\end{aligned}$$

Now,

$$f^h(x) \in \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\} \qquad \text{for each } h \in \mathbb{N}. \qquad (29)$$

[*Proof of (29):* We shall prove (29) by induction over h:

Induction base: We have $i < j$, hence $j > i \geq 0$ and thus $j \geq 1$ (since j is an integer). Hence, $0 \in \{0, 1, \ldots, j-1\}$, so that $f^0(x) \in \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}$. In other words, (29) holds for $h = 0$. This completes the induction base.

Induction step: Let $g \in \mathbb{N}$. Assume that (29) holds for $h = g$. We must now show that (29) holds for $h = g + 1$ as well.

We have assumed that (29) holds for $h = g$. In other words, $f^g(x) \in \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}$. In other words, there exists some $k \in \{0, 1, \ldots, j-1\}$ such that $f^g(x) = f^k(x)$. Consider this k.

We have $k \in \{0, 1, \ldots, j-1\}$, so that $k + 1 \in \{1, 2, \ldots, j\} \subseteq \{0, 1, \ldots, j\}$ and therefore

$$f^{k+1}(x) \in \left\{f^0(x), f^1(x), \ldots, f^j(x)\right\} = \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}$$

(by (28)). But

$$f^{g+1}(x) = f\left(\underbrace{f^g(x)}_{=f^k(x)}\right) = f\left(f^k(x)\right) = f^{k+1}(x) \in \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}$$

(as we have just proven). In other words, (29) holds for $h = g + 1$ as well. This completes the induction step. Thus, (29) is proven.]

From (29), we immediately obtain

$$\left\{f^h(x) \mid h \in \mathbb{N}\right\} \subseteq \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}.$$

Combining this with (27), we obtain

$$\left\{f^h(x) \mid h \in \mathbb{N}\right\} = \left\{f^0(x), f^1(x), \ldots, f^{j-1}(x)\right\}.$$

This proves Lemma 0.18. □

Lemma 0.19. Let X be a finite set. Let σ be a permutation of X. Let $x \in X$.

(a) There exists a $j \in \mathbb{N}$ such that $\sigma^j(x) \in \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{j-1}(x)\right\}$. Let p be the smallest such j.

(b) The integer p is positive and satisfies $\sigma^p(x) = x$.

(c) The elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)$ are pairwise distinct.

(d) We have $\left\{\sigma^h(x) \mid h \in \mathbb{N}\right\} = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\}$.

Proof of Lemma 0.19. The map σ is a permutation of X. In other words, σ is a bijection $X \to X$. Hence, σ is injective.

(a) Define an $n \in \mathbb{N}$ by $n = |X|$. (This is well-defined, since X is a finite set.)

The $n+1$ elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^n(x)$ cannot all be distinct, because they all belong to the n-element set X. Hence, at least two of these $n+1$ elements are equal. In other words, there exist two elements u and v of $\{0, 1, \ldots, n\}$ such that $u < v$ and $\sigma^u(x) = \sigma^v(x)$. Consider these u and v.

We have $u \in \{0, 1, \ldots, n\} \subseteq \mathbb{N}$. Thus, $u \in \{0, 1, \ldots, v-1\}$ (since $u < v$). Hence, $\sigma^u(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{v-1}(x)\}$. In view of $\sigma^u(x) = \sigma^v(x)$, this rewrites as $\sigma^v(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{v-1}(x)\}$. Thus, there exists a $j \in \mathbb{N}$ such that $\sigma^j(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{j-1}(x)\}$ (namely, $j = v$). This proves Lemma 0.19 **(a)**.

Now, let us study the p in Lemma 0.19. We have defined p as the smallest $j \in \mathbb{N}$ such that $\sigma^j(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{j-1}(x)\}$. Thus, p is an element of $\mathbb{N}$ and satisfies $\sigma^p(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\}$. Hence, the set $\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\}$ is nonempty (since it contains the element $\sigma^p(x)$). Thus, $p \neq 0$ (because if we had $p = 0$, then the set $\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\}$ would be empty). Hence, p is a positive integer (since $p \in \mathbb{N}$).

(b) We already know that p is positive. It thus remains to show that $\sigma^p(x) = x$.

Indeed, we have $\sigma^p(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\}$. In other words, there exists some $i \in \{0, 1, \ldots, p-1\}$ such that $\sigma^p(x) = \sigma^i(x)$. Consider this i.

Next, we claim that $i = 0$. We shall prove this by contradiction. Indeed, assume the contrary. Thus, $i \neq 0$, so that $i > 0$ (since $i \in \mathbb{N}$). Hence, $\sigma^i(x) = \sigma\left(\sigma^{i-1}(x)\right)$. But the integer p is also positive; hence, $p - 1 \in \mathbb{N}$ and $\sigma^p(x) = \sigma\left(\sigma^{p-1}(x)\right)$. Hence, $\sigma\left(\sigma^{p-1}(x)\right) = \sigma^p(x) = \sigma\left(\sigma^{i-1}(x)\right)$. Since σ is injective, we thus conclude that $\sigma^{p-1}(x) = \sigma^{i-1}(x)$. But $i - 1 \in \mathbb{N}$ (since $i > 0$). From $i \in \{0, 1, \ldots, p-1\}$, we obtain $i - 1 \in \{-1, 0, \ldots, (p-1)-1\}$. Combined with $i - 1 \in \mathbb{N}$, this yields $i - 1 \in \{-1, 0, \ldots, (p-1)-1\} \cap \mathbb{N} = \{0, 1, \ldots, (p-1)-1\}$. Hence, $\sigma^{i-1}(x) \in \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{(p-1)-1}(x)\right\}$. Hence,

$$\sigma^{p-1}(x) = \sigma^{i-1}(x) \in \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{(p-1)-1}(x)\right\}.$$

Thus, $p - 1$ is a $j \in \mathbb{N}$ such that $\sigma^j(x) \in \{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{j-1}(x)\}$ (because $p - 1 \in \mathbb{N}$). But we defined p to be the **smallest** such j. Hence, $p \leq p - 1$. This contradicts $p > p - 1$. This contradiction shows that our assumption was false; hence, we have shown that $i = 0$. Therefore, $\sigma^i(x) = \underbrace{\sigma^0}_{=\mathrm{id}}(x) = \mathrm{id}(x) = x$.

Now, $\sigma^p(x) = \sigma^i(x) = x$. This completes the proof of Lemma 0.19 **(b)**.

(c) Assume the contrary. Thus, two of the elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)$ are equal. In other words, there exist two elements u and v of $\{0, 1, \ldots, p-1\}$ such that $u < v$ and $\sigma^u(x) = \sigma^v(x)$. Consider these u and v. Notice that $v \leq p - 1$ (since $v \in \{0, 1, \ldots, p-1\}$).

From $u \in \{0, 1, \ldots, p-1\}$, we obtain $u \geq 0$. From $u < v$, we obtain $u \leq v - 1$ (since u and v are integers), so that $u \in \{0, 1, \ldots, v-1\}$ (since $u \geq 0$).

Hence, $\sigma^{u}(x) \in \left\{\sigma^{0}(x), \sigma^{1}(x), \ldots, \sigma^{v-1}(x)\right\}$. From $\sigma^{u}(x) = \sigma^{v}(x)$, we obtain $\sigma^{v}(x) = \sigma^{u}(x) \in \left\{\sigma^{0}(x), \sigma^{1}(x), \ldots, \sigma^{v-1}(x)\right\}$. Thus, v is a $j \in \mathbb{N}$ such that $\sigma^{j}(x) \in \left\{\sigma^{0}(x), \sigma^{1}(x), \ldots, \sigma^{j-1}(x)\right\}$ (because $v \in \mathbb{N}$). But we defined p to be the **smallest** such j. Hence, $p \leq v$. This contradicts $v \leq p-1 < p$. This contradiction shows that our assumption was false. Thus, Lemma 0.19 **(c)** is proven.

(d) We have $\sigma^{p}(x) \in \left\{\sigma^{0}(x), \sigma^{1}(x), \ldots, \sigma^{p-1}(x)\right\}$. In other words, there exists some $i \in \{0, 1, \ldots, p-1\}$ such that $\sigma^{p}(x) = \sigma^{i}(x)$. Consider this i. Hence, $i < p$ (since $i \in \{0, 1, \ldots, p-1\}$) and $\sigma^{i}(x) = \sigma^{p}(x)$. Thus, Lemma 0.18 (applied to $f = \sigma$ and $j = p$) yields

$$\left\{\sigma^{h}(x) \mid h \in \mathbb{N}\right\} = \left\{\sigma^{0}(x), \sigma^{1}(x), \ldots, \sigma^{p-1}(x)\right\}.$$

This proves Lemma 0.19 **(d)**. □

Lemma 0.20. Let X be a set. Let $f : X \to X$ be any map. Let $x \in X$. Let $p \in \mathbb{N}$ be such that $f^{p}(x) = x$. Then, $f^{kp}(x) = x$ for each $k \in \mathbb{N}$.

Proof of Lemma 0.20. Lemma 0.20 is intuitively obvious: All it says is that if applying the map f to x a total of p times brings you back to x, then applying the map f to x a total of kp times brings you back to x as well. This intuition can easily be translated into a rigorous argument:

We shall prove Lemma 0.20 by induction over k:

Induction base: We have $f^{0p} = f^{0} = \mathrm{id}_X$, so that $f^{0p}(x) = \mathrm{id}_X(x) = x$. Thus, Lemma 0.20 holds for $k = 0$. This completes the induction base.

Induction step: Let $m \in \mathbb{N}$. Assume that Lemma 0.20 holds for $k = m$. We must prove that Lemma 0.20 holds for $k = m+1$.

Let $x \in X$. Let $p \in \mathbb{N}$ be such that $f^{p}(x) = x$. Then, $f^{mp}(x) = x$ (since Lemma 0.20 holds for $k = m$). But $f^{mp} \circ f^{p} = f^{mp+p} = f^{(m+1)p}$. Hence, $(f^{mp} \circ f^{p})(x) = f^{(m+1)p}(x)$, and therefore

$$f^{(m+1)p}(x) = (f^{mp} \circ f^{p})(x) = f^{mp}\left(\underbrace{f^{p}(x)}_{=x}\right) = f^{mp}(x) = x.$$

In other words, Lemma 0.20 holds for $k = m+1$. This completes the induction step. Thus, Lemma 0.20 is proven. □

Lemma 0.21. Let X be a set. Let $m \in \mathbb{N}$. Let $f_1, f_2, \ldots, f_m$ be m maps from X to X. Let x and y be two elements of X.

Let $i \in [m]$. Assume that $f_i(x) = y$. Assume further that

$$f_j(x) = x \qquad \text{for each } j \in [m] \text{ satisfying } j < i. \tag{30}$$

Assume also that

$$f_j(y) = y \qquad \text{for each } j \in [m] \text{ satisfying } j > i. \tag{31}$$

Then, $(f_m \circ f_{m-1} \circ \cdots \circ f_1)(x) = y$.

Proof of Lemma 0.21. The idea behind this proof is very simple (if we don't insist on being rigorous): Imagine the element x undergoing the maps $f_1, f_2, \ldots, f_m$ in this order; the result is, of course, $(f_m \circ f_{m-1} \circ \cdots \circ f_1)(x)$. But let us look closer at the step-by-step procedure. The element is initially x. Then, the maps $f_1, f_2, \ldots, f_m$ are being applied to it in this order. Up until the map f_i is applied, the element does not change (because of (30)). Then, the map f_i is applied, and the element becomes y (since $f_i(x) = y$). From then on, the maps $f_{i+1}, f_{i+2}, \ldots, f_m$ again leave the element unchanged (due to (31)). Thus, the final result is y. This shows that $(f_m \circ f_{m-1} \circ \cdots \circ f_1)(x) = y$.

Let us now rewrite the above argument in rigorous terms.

We have $i \in [m]$, so that $1 \leq i \leq m$. Now, we claim the following:

Observation 1: We have $(f_g \circ f_{g-1} \circ \cdots \circ f_1)(x) = \begin{cases} x, & \text{if } g < i; \\ y, & \text{if } g \geq i \end{cases}$ for each $g \in \{0, 1, \ldots, m\}$.

[*Proof of Observation 1:* We shall prove Observation 1 by induction on g:

Induction base: We have $0 < i$ (since $i \in [m]$). Thus, $\begin{cases} x, & \text{if } 0 < i; \\ y, & \text{if } 0 \geq i \end{cases} = x$. Comparing this with

$$\underbrace{(f_0 \circ f_{0-1} \circ \cdots \circ f_1)}_{\substack{=(\text{empty composition of maps } X \to X) \\ = \mathrm{id}}}(x) = \mathrm{id}(x) = x,$$

we obtain $(f_0 \circ f_{0-1} \circ \cdots \circ f_1)(x) = \begin{cases} x, & \text{if } 0 < i; \\ y, & \text{if } 0 \geq i \end{cases}$. In other words, Observation 1 holds for $g = 0$. This completes the induction base.

Induction step: Let $h \in \{0, 1, \ldots, m\}$ be positive. Assume that Observation 1 holds for $g = h - 1$. We must then prove that Observation 1 holds for $g = h$.

We have

$$f_h\left(\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases}\right) = \begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases} \tag{32}$$

[16].

[16]*Proof of (32):* We are in one of the following three cases:

Case 1: We have $h < i$.

Case 2: We have $h = i$.

Case 3: We have $h > i$.

Let us first consider Case 1. In this case, we have $h < i$. Thus, $\begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases} = x$.

Applying (30) to $j = h$, we find $f_h(x) = x$ (since $h < i$).

Also, $h - 1 < h < i$. Hence, $\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases} = x$. Applying the map f_h to this equality, we obtain

$$f_h\left(\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases}\right) = f_h(x) = x = \begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases}.$$

Hence, (32) is proven in Case 1.

Let us now consider case 2. In this case, we have $h = i$. Thus, $h \geq i$. Hence, $\begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases} = y$.

But we assumed that Observation 1 holds for $g = h - 1$. In other words, we have

$$\left(f_{h-1} \circ f_{(h-1)-1} \circ \cdots \circ f_1\right)(x) = \begin{cases} x, & \text{if } h - 1 < i; \\ y, & \text{if } h - 1 \geq i \end{cases}.$$

Now,

$$\begin{aligned}
\underbrace{(f_h \circ f_{h-1} \circ \cdots \circ f_1)}_{\substack{= f_h \circ (f_{h-1} \circ f_{h-2} \circ \cdots \circ f_1) \\ = f_h \circ \left(f_{h-1} \circ f_{(h-1)-1} \circ \cdots \circ f_1\right)}} (x) &= \left(f_h \circ \left(f_{h-1} \circ f_{(h-1)-1} \circ \cdots \circ f_1\right)\right)(x) \\
&= f_h \left(\underbrace{\left(f_{h-1} \circ f_{(h-1)-1} \circ \cdots \circ f_1\right)(x)}_{= \begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases}} \right) \\
&= f_h \left(\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases} \right) = \begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases} \qquad (\text{by (32)}) .
\end{aligned}$$

In other words, Observation 1 holds for $g = h$. This completes the induction step. Thus, Observation 1 is proven.]

We can now apply Observation 1 to $g = m$. We thus obtain

$$(f_m \circ f_{m-1} \circ \cdots \circ f_1)(x) = \begin{cases} x, & \text{if } m < i; \\ y, & \text{if } m \geq i \end{cases} = y$$

(since $m \geq i$ (since $i \leq m$)). This proves Lemma 0.21. $\square$

From $h = i$, we obtain $f_h(x) = f_i(x) = y$.

Also, $h - 1 < h = i$. Hence, $\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases} = x$. Applying the map f_h to this equality, we obtain

$$f_h \left(\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases} \right) = f_h(x) = y = \begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases}.$$

Hence, (32) is proven in Case 2.

Let us first consider Case 3. In this case, we have $h > i$. Thus, $h \geq i$, so that $\begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases} = y$.

Applying (31) to $j = h$, we find $f_h(y) = y$ (since $h > i$).

Also, $h > i$, so that $h \geq i + 1$ (since h and i are integers). Thus, $h - 1 \geq i$. Hence, $\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases} = y$. Applying the map f_h to this equality, we obtain

$$f_h \left(\begin{cases} x, & \text{if } h-1 < i; \\ y, & \text{if } h-1 \geq i \end{cases} \right) = f_h(y) = y = \begin{cases} x, & \text{if } h < i; \\ y, & \text{if } h \geq i \end{cases}.$$

Hence, (32) is proven in Case 3.

We have now proven (32) in each of the three Cases 1, 2 and 3 (which are the only cases that can occur). Thus, (32) always holds.

Lemma 0.22. Let X be a set. Let $m \in \mathbb{N}$. Let $g_1, g_2, \ldots, g_m$ be m maps from X to X. Let x and y be two elements of X.

Let $i \in [m]$. Assume that $g_i(x) = y$. Assume further that

$$g_j(x) = x \qquad \text{for each } j \in [m] \text{ satisfying } j > i. \tag{33}$$

Assume also that

$$g_j(y) = y \qquad \text{for each } j \in [m] \text{ satisfying } j < i. \tag{34}$$

Then, $(g_1 \circ g_2 \circ \cdots \circ g_m)(x) = y$.

Proof of Lemma 0.22. Lemma 0.22 follows by applying Lemma 0.21 to $g_m, g_{m-1}, \ldots, g_1$ instead of $f_1, f_2, \ldots, f_m$.

Here is the argument in more detail:

For each $j \in [m]$, we define a map f_j from X to X by $f_j = g_{m+1-j}$.

From $i \in [m]$, we obtain $m + 1 - i \in [m]$. Thus, we can define $i' \in [m]$ by $i' = m + 1 - i$. Consider this i'. From $i' = m + 1 - i$, we obtain $m + 1 - i' = i$. Now, the definition of $f_{i'}$ yields $f_{i'} = g_{m+1-i'} = g_i$ (since $m + 1 - i' = i$). Thus, $f_{i'}(x) = g_i(x) = y$.

Furthermore, $f_j(x) = x$ for each $j \in [m]$ satisfying $j < i'$ [17]. Also, $f_j(y) = y$ for each $j \in [m]$ satisfying $j > i'$ [18]. Hence, Lemma 0.21 (applied to i' instead of i) yields $(f_m \circ f_{m-1} \circ \cdots \circ f_1)(x) = y$.

But each $j \in [m]$ satisfies

$$\begin{aligned} f_{m+1-j} &= g_{m+1-(m+1-j)} \qquad (\text{by the definition of } f_{m+1-j}) \\ &= g_j \qquad (\text{since } m + 1 - (m + 1 - j) = j). \end{aligned}$$

In other words, we have $(f_m, f_{m-1}, \ldots, f_1) = (g_1, g_2, \ldots, g_m)$. Hence, $f_m \circ f_{m-1} \circ \cdots \circ f_1 = g_1 \circ g_2 \circ \cdots \circ g_m$. Hence, $(f_m \circ f_{m-1} \circ \cdots \circ f_1)(x) = (g_1 \circ g_2 \circ \cdots \circ g_m)(x)$. Therefore,

$$(g_1 \circ g_2 \circ \cdots \circ g_m)(x) = (f_m \circ f_{m-1} \circ \cdots \circ f_1)(x) = y.$$

This proves Lemma 0.22. $\square$

Solution to Exercise 7 (sketched). The map σ is a permutation of X, thus a bijection $X \to X$. Hence, in particular, σ is injective.

Before we properly start solving the exercise, let us make some basic observations:

> *Observation 1.* For every $x \in X$, there exists some positive integer p such that $\sigma^p(x) = x$.

[17] *Proof.* Let $j \in [m]$ be such that $j < i'$. Then, $m + 1 - j \in [m]$ (since $j \in [m]$) and $m + 1 - \underbrace{j}_{<i'} > m + 1 - i' = i$. Hence, (33) (applied to $m + 1 - j$ instead of j) yields $g_{m+1-j}(x) = x$. But the definition of f_j yields $f_j = g_{m+1-j}$. Thus, $f_j(x) = g_{m+1-j}(x) = x$. Qed.

[18] *Proof.* Let $j \in [m]$ be such that $j > i'$. Then, $m + 1 - j \in [m]$ (since $j \in [m]$) and $m + 1 - \underbrace{j}_{>i'} < m + 1 - i' = i$. Hence, (34) (applied to $m + 1 - j$ instead of j) yields $g_{m+1-j}(y) = y$. But the definition of f_j yields $f_j = g_{m+1-j}$. Thus, $f_j(y) = g_{m+1-j}(y) = y$. Qed.

[*Proof of Observation 1:* Let $x \in X$. Let $n = |X|$. The $n+1$ elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^n(x)$ cannot all be distinct, because they belong to the n-element set X. Hence, at least two of these $n+1$ elements are equal. In other words, there exist two elements i and j of $\{0, 1, \ldots, n\}$ such that $i < j$ and $\sigma^i(x) = \sigma^j(x)$. Consider these i and j. From $i < j$, we conclude that $j - i$ is a positive integer. Thus, $\sigma^j = \sigma^i \circ \sigma^{j-i}$.

Also, the map σ^i is injective (since the map σ is injective, but any composition of injective maps is injective). Hence, from

$$\sigma^i(x) = \underbrace{\sigma^j}_{=\sigma^i \circ \sigma^{j-i}}(x) = \left(\sigma^i \circ \sigma^{j-i}\right)(x) = \sigma^i\left(\sigma^{j-i}(x)\right),$$

we obtain $x = \sigma^{j-i}(x)$. In other words, $\sigma^{j-i}(x) = x$. Hence, there exists some positive integer p such that $\sigma^p(x) = x$ (namely, $p = j - i$). This proves Observation 1.]

Now, we must show that $\sim$ is an equivalence relation. Indeed, the relation $\sim$ is reflexive[19], symmetric[20] and transitive[21]. In other words, the relation $\sim$ is an equivalence relation. This solves Exercise 7 **(a)**.

[19] *Proof.* Let $x \in X$. We shall show that $x \sim x$.

Indeed, $\sigma^0 = \mathrm{id}_X$, so that $\sigma^0(x) = \mathrm{id}_X(x) = x$. Hence, there exists some $k \in \mathbb{N}$ such that $x = \sigma^k(x)$ (namely, $k = 0$). In other words, $x \sim x$ (by the definition of the relation $\sim$).

Now, forget that we fixed x. We thus have shown that every $x \in X$ satisfies $x \sim x$. In other words, the relation $\sim$ is reflexive.

[20] *Proof.* Let $x \in X$ and $y \in X$ be such that $x \sim y$. We shall show that $y \sim x$.

Indeed, we have $x \sim y$. In other words, there exists some $k \in \mathbb{N}$ such that $y = \sigma^k(x)$ (by the definition of the relation $\sim$). Consider such a k, and denote it by u. Thus, $u \in \mathbb{N}$ satisfies $y = \sigma^u(x)$.

Observation 1 yields that there exists some positive integer p such that $\sigma^p(x) = x$. Consider this p. Hence, Lemma 0.20 (applied to $f = \sigma$ and $k = u$) yields $\sigma^{up}(x) = x$. But p is positive; hence, $p \geq 1$ and thus $up \geq u1 = u$. Hence, $up - u \in \mathbb{N}$. Hence, $\sigma^{up-u} \circ \sigma^u = \sigma^{(up-u)+u} = \sigma^{up}$. Thus, $(\sigma^{up-u} \circ \sigma^u)(x) = \sigma^{up}(x) = x$. Hence, $x = (\sigma^{up-u} \circ \sigma^u)(x) = \sigma^{up-u}\left(\underbrace{\sigma^u(x)}_{=y}\right) = \sigma^{up-u}(y)$. Thus, there exists some $k \in \mathbb{N}$ such that $x = \sigma^k(y)$ (namely, $k = up - u$). In other words, $y \sim x$ (by the definition of the relation $\sim$).

Now, forget that we fixed x and y. We thus have shown that if $x \in X$ and $y \in X$ satisfy $x \sim y$, then $y \sim x$. In other words, the relation $\sim$ is symmetric.

[21] *Proof.* Let $x \in X$, $y \in X$ and $z \in X$ be such that $x \sim y$ and $y \sim z$. We shall show that $x \sim z$.

Indeed, we have $x \sim y$. In other words, there exists some $k \in \mathbb{N}$ such that $y = \sigma^k(x)$ (by the definition of the relation $\sim$). Consider such a k, and denote it by u. Thus, $u \in \mathbb{N}$ satisfies $y = \sigma^u(x)$.

Also, we have $y \sim z$. In other words, there exists some $k \in \mathbb{N}$ such that $z = \sigma^k(y)$ (by the definition of the relation $\sim$). Consider such a k, and denote it by v. Thus, $v \in \mathbb{N}$ satisfies $z = \sigma^v(y)$.

But $\sigma^v \circ \sigma^u = \sigma^{v+u}$. Thus, $(\sigma^v \circ \sigma^u)(x) = \sigma^{v+u}(x)$. In view of $(\sigma^v \circ \sigma^u)(x) = \sigma^v\left(\underbrace{\sigma^u(x)}_{=y}\right) = \sigma^v(y) = z$, this rewrites as $z = \sigma^{v+u}(x)$. Thus, there exists some $k \in \mathbb{N}$ such that $z = \sigma^k(x)$ (namely, $k = v + u$). In other words, $x \sim z$ (by the definition of the relation $\sim$).

Now, forget that we fixed x, y and z. We thus have shown that if $x \in X$, $y \in X$ and $z \in X$ satisfy $x \sim y$ and $y \sim z$, then $x \sim z$. In other words, the relation $\sim$ is transitive.

(b) Let $x \in X$. Lemma 0.19 **(a)** shows that there exists a $j \in \mathbb{N}$ such that $\sigma^j(x) \in \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{j-1}(x)\right\}$. Let p be the smallest such j.

Lemma 0.19 **(b)** shows that the integer p is positive and satisfies $\sigma^p(x) = x$. Lemma 0.19 **(c)** shows that the elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)$ are pairwise distinct. Lemma 0.19 **(d)** shows that $\left\{\sigma^h(x) \mid h \in \mathbb{N}\right\} = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\}$.

Define a set S by $S = \left\{\sigma^h(x) \mid h \in \mathbb{N}\right\}$. Thus,

$$S = \left\{\sigma^h(x) \mid h \in \mathbb{N}\right\} = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\}.$$

The definition of the equivalence class $[x]_\sim$ of x shows that

$$[x]_\sim = \{y \in X \mid y \sim x\}.$$

Now, $[x]_\sim \subseteq S$ [22] and $S \subseteq [x]_\sim$ [23]. Combining these two relations, we obtain $[x]_\sim = S = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\}$.

The p elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)$ are pairwise distinct (as we have seen above). Thus, $\left|\left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\}\right| = p$.

Let $k = |[x]_\sim|$. Then,

$$k = |[x]_\sim| = \left|\left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\}\right| = p.$$

Now,

$$[x]_\sim = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)\right\} = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right\}$$

(since $p = k$). This solves Exercise 7 **(b)**.

(c) Let E be a $\sim$-equivalence class. We must prove that c_E is a permutation of X.

It is clear that c_E is well-defined. Next, we claim that

$$t \in c_E(X) \qquad \text{for each } t \in X. \tag{35}$$

[22] *Proof.* Let $w \in [x]_\sim$. Thus, $w \in [x]_\sim = \{y \in X \mid y \sim x\}$. In other words, w is an element of X and satisfies $w \sim x$.

We have $w \sim x$. Hence, $x \sim w$ (since the relation $\sim$ is symmetric). In other words, there exists some $k \in \mathbb{N}$ such that $w = \sigma^k(x)$ (by the definition of the relation $\sim$). Consider this k. We have $w = \sigma^k(x) \in \left\{\sigma^h(x) \mid h \in \mathbb{N}\right\} = S$.

Now, forget that we fixed w. We thus have shown that $w \in S$ for each $w \in [x]_\sim$. In other words, $[x]_\sim \subseteq S$.

[23] *Proof.* Let $w \in S$. Thus, $w \in S = \left\{\sigma^h(x) \mid h \in \mathbb{N}\right\}$. In other words, $w = \sigma^h(x)$ for some $h \in \mathbb{N}$. Consider this h.

There exists some $k \in \mathbb{N}$ such that $w = \sigma^k(x)$ (namely, $k = h$). In other words, $x \sim w$ (by the definition of the relation $\sim$). Thus, $w \sim x$ (since the relation $\sim$ is symmetric).

Hence, w is an element of X and satisfies $w \sim x$. In other words, $w \in \{y \in X \mid y \sim x\}$. In view of $[x]_\sim = \{y \in X \mid y \sim x\}$, this rewrites as $w \in [x]_\sim$.

Now, forget that we fixed w. We thus have shown that $w \in [x]_\sim$ for each $w \in S$. In other words, $S \subseteq [x]_\sim$.

[*Proof of (35):* Let $t \in X$. We must prove that $t \in c_E(X)$.

We are in one of the following two cases:

Case 1: We have $t \in E$.

Case 2: We have $t \notin E$.

Let us first consider Case 1. In this case, we have $t \in E$. But E is an $\sim$-equivalence class. Hence, E is an $\sim$-equivalence class containing t (since $t \in E$). In other words, $E = [t]_\sim$ (since the only $\sim$-equivalence class containing t is $[t]_\sim$).

Recall that σ is a permutation of X. Hence, an element $\sigma^{-1}(t)$ of X is well-defined. Denote this element by z. Thus, $z = \sigma^{-1}(t)$. Hence, $\sigma(z) = t$.

We have $\underbrace{\sigma^1}_{=\sigma}(z) = \sigma(z) = t$, so that $t = \sigma^1(z)$. Hence, there exists some $k \in \mathbb{N}$ such that $t = \sigma^k(z)$ (namely, $k = 1$). In other words, $z \sim t$ (by the definition of the relation $\sim$). Hence, z is an element of X and satisfies $z \sim t$. In other words, $z \in \{y \in X \mid y \sim t\}$. But $E = [t]_\sim = \{y \in X \mid y \sim t\}$ (by the definition of the equivalence class $[t]_\sim$). Hence, $z \in \{y \in X \mid y \sim t\} = E$. The definition of c_E yields

$$c_E(z) = \begin{cases} \sigma(z), & \text{if } z \in E; \\ z, & \text{if } z \notin E \end{cases} = \sigma(z) \qquad (\text{since } z \in E)$$
$$= t.$$

Hence, $t = c_E\left(\underbrace{z}_{\in X}\right) \in c_E(X)$. Thus, we have proven $t \in c_E(X)$ in Case 1.

Let us now consider Case 2. In this case, we have $t \notin E$. The definition of c_E yields

$$c_E(t) = \begin{cases} \sigma(t), & \text{if } t \in E; \\ t, & \text{if } t \notin E \end{cases} = t \qquad (\text{since } t \notin E).$$

Hence, $t = c_E\left(\underbrace{t}_{\in X}\right) \in c_E(X)$. Thus, we have proven $t \in c_E(X)$ in Case 2.

We have now proven $t \in c_E(X)$ in each of the two Cases 1 and 2. Hence, $t \in c_E(X)$ is proven. This proves (35).]

Now, (35) shows that $X \subseteq c_E(X)$. In other words, the map c_E is surjective. Thus, c_E is a surjective map between two finite sets of the same size (namely, X and X), and therefore must be bijective (since any surjective map between two finite sets of the same size is bijective). In other words, c_E is a bijection $X \to X$, therefore a permutation of X. This solves Exercise 7 **(c)**.

(d) Let $x \in X$ be such that $E = [x]_\sim$. Let $k = |[x]_\sim|$. We must prove that c_E can be written as $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$ (and in particular, we must prove that $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)$ are distinct, so that $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$ is well-defined).

Lemma 0.19 **(a)** shows that there exists a $j \in \mathbb{N}$ such that $\sigma^j(x) \in \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{j-1}(x)\right\}$. Let p be the smallest such j. As in the solution to Exercise 7 **(b)** (which we have given above), we can see that $k = p$.

Lemma 0.19 **(b)** shows that the integer p is positive and satisfies $\sigma^p(x) = x$. In view of $k = p$, this rewrites as follows: The integer k is positive and satisfies $\sigma^k(x) = x$. Since the integer k is positive, we have $1 \in [k]$.

Lemma 0.19 **(c)** shows that the elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{p-1}(x)$ are pairwise distinct. In view of $k = p$, this rewrites as follows: The elements $\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)$

are pairwise distinct. Hence, the permutation $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$ is well-defined (since k is a positive integer).

Exercise 7 **(b)** shows that $[x]_\sim = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right\}$. Hence,

$$\left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right\} = [x]_\sim = E \tag{36}$$

(since $E = [x]_\sim$).

It remains to prove that c_E can be written as $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$.

We define a k-tuple $(i_1, i_2, \ldots, i_k)$ of elements of X by

$$(i_1, i_2, \ldots, i_k) = \left(\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right). \tag{37}$$

Thus,

$$i_u = \sigma^{u-1}(x) \qquad \text{for each } u \in [k]. \tag{38}$$

Applying this to $u = 1$, we obtain $i_1 = \sigma^{1-1}(x)$ (since $1 \in [k]$).

Also, from (37), we obtain

$$\{i_1, i_2, \ldots, i_k\} = \left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right\} = E \tag{39}$$

(by (36)).

We also let i_{k+1} stand for i_1. Thus, $i_{k+1} = i_1 = \underbrace{\sigma^{1-1}}_{=\sigma^0=\mathrm{id}}(x) = \mathrm{id}(x) = x = \sigma^k(x)$ (since $\sigma^k(x) = x$). Therefore, we see that

$$i_u = \sigma^{u-1}(x) \qquad \text{for each } u \in [k+1]. \tag{40}$$

[*Proof of (40):* Let $u \in [k+1]$. We must prove that $i_u = \sigma^{u-1}(x)$. If $u \in [k]$, then this follows from (38). Hence, for the rest of this proof, we WLOG assume that $u \notin [k]$. Combining $u \in [k+1]$ with $u \notin [k]$, we obtain $u \in [k+1] \setminus [k] = \{k+1\}$, so that $u = k+1$. Thus, $i_u = i_{k+1} = \sigma^k(x) = \sigma^{u-1}(x)$ (since $k = u-1$ (since $u = k+1$)). This proves (40).]

We have $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)} = \operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ (since $\left(\sigma^0(x), \sigma^1(x), \ldots, \sigma^{k-1}(x)\right) = (i_1, i_2, \ldots, i_k)$). But the definition of $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ yields

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = \begin{cases} i_{j+1}, & \text{if } p = i_j \text{ for some } j \in \{1, 2, \ldots, k\}; \\ p, & \text{otherwise} \end{cases} \tag{41}$$

for every $p \in X$.

Now, we claim that

$$\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = c_E(p) \qquad \text{for each } p \in X. \tag{42}$$

[*Proof of (42):* Let $p \in X$. We must prove the equality (42).

We are in one of the following two cases:

Case 1: We have $p = i_j$ for some $j \in \{1, 2, \ldots, k\}$.

Case 2: We don't have $\left(p = i_j \text{ for some } j \in \{1,2,\ldots,k\}\right)$.

Let us first consider Case 1. In this case, we have $p = i_j$ for some $j \in \{1,2,\ldots,k\}$. Consider this j. Thus, (41) simplifies to $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = i_{j+1}$.

We have $j \in \{1,2,\ldots,k\} = [k]$. Hence, (38) (applied to $u = j$) yields $i_j = \sigma^{j-1}(x)$. Hence, $\sigma^{j-1}(x) = i_j = p$.

But $j \in \{1,2,\ldots,k\}$, so that $j+1 \in \{2,3,\ldots,k+1\} \subseteq [k+1]$. Hence, (40) (applied to $u = j+1$) yields

$$i_{j+1} = \underbrace{\sigma^{(j+1)-1}}_{=\sigma^j}(x) = \sigma^j(x) = \sigma\left(\underbrace{\sigma^{j-1}(x)}_{=p}\right) = \sigma(p). \tag{43}$$

However,

$$\begin{aligned} p = i_j &\in \{i_1,i_2,\ldots,i_k\} \qquad (\text{since } j \in \{1,2,\ldots,k\}) \\ &= E \qquad (\text{by (39)}). \end{aligned}$$

The definition of c_E now shows that

$$\begin{aligned} c_E(p) &= \begin{cases} \sigma(p), & \text{if } p \in E; \\ p, & \text{if } p \notin E \end{cases} = \sigma(p) \qquad (\text{since } p \in E) \\ &= i_{j+1} \qquad (\text{by (43)}). \end{aligned}$$

Comparing this with $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = i_{j+1}$, we obtain $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = c_E(p)$. Thus, (42) is proven in Case 1.

Let us now consider Case 2. In this case, we don't have $\left(p = i_j \text{ for some } j \in \{1,2,\ldots,k\}\right)$. Thus, (41) simplifies to $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = p$.

But we don't have $\left(p = i_j \text{ for some } j \in \{1,2,\ldots,k\}\right)$. In other words, $p \notin \{i_1,i_2,\ldots,i_k\}$. In view of (39), this rewrites as $p \notin E$. The definition of c_E now shows that

$$c_E(p) = \begin{cases} \sigma(p), & \text{if } p \in E; \\ p, & \text{if } p \notin E \end{cases} = p \qquad (\text{since } p \notin E).$$

Comparing this with $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = p$, we obtain $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}(p) = c_E(p)$. Thus, (42) is proven in Case 2.

We have now proven (42) in each of the two Cases 1 and 2. This completes the proof of (42).]

The equality (42) shows that $\operatorname{cyc}_{i_1,i_2,\ldots,i_k} = c_E$ (since both $\operatorname{cyc}_{i_1,i_2,\ldots,i_k}$ and c_E are maps $X \to X$). Thus,

$$c_E = \operatorname{cyc}_{i_1,i_2,\ldots,i_k} = \operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$$

(by (37)). In other words, c_E can be written as $\operatorname{cyc}_{\sigma^0(x),\sigma^1(x),\ldots,\sigma^{k-1}(x)}$. This concludes the solution to Exercise 7 **(d)**.

(e) Recall that the $\sim$-equivalence classes form a set partition of the set X (in fact, this holds for the equivalence classes of any equivalence relation on X). Thus, each element of X belongs to exactly one $\sim$-equivalence class. Since $E_1, E_2, \ldots, E_m$ are all the $\sim$-equivalence classes (listed without repetition), we can rewrite this fact as follows: Each element of X belongs to exactly one of the sets $E_1, E_2, \ldots, E_m$. Thus, the sets $E_1, E_2, \ldots, E_m$ are disjoint. In other words, if i and j are two distinct elements of $[m]$, then

$$E_i \cap E_j = \varnothing. \tag{44}$$

Now, fix $x \in X$. Define $y \in X$ by $y = \sigma(x)$. We are going to show that $(c_{E_1} \circ c_{E_2} \circ \cdots \circ c_{E_m})(x) = y$.

The element x of X belongs to exactly one of the sets $E_1, E_2, \ldots, E_m$ (since each element of X belongs to exactly one of the sets $E_1, E_2, \ldots, E_m$). In other words, there is exactly one $i \in [m]$ such that $x \in E_i$. Consider this i.

Hence, i is the **only** element $j \in [m]$ such that $x \in E_j$. Therefore, every $j \in [m]$ distinct from i must satisfy

$$x \notin E_j. \tag{45}$$

We have $x \sim y$ [24]. Thus, $y \in [x]_\sim$. But recall that E_i is a $\sim$-equivalence class (since $E_1, E_2, \ldots, E_m$ are all the $\sim$-equivalence classes) and contains x (since $x \in E_i$). Hence, E_i is the $\sim$-equivalence class of x. In other words, $E_i = [x]_\sim$. Hence, $y \in [x]_\sim = E_i$. Hence, every $j \in [m]$ distinct from i must satisfy

$$y \notin E_j \tag{46}$$

[25].

We have $\sigma_{E_i}(x) = y$ [26]. Furthermore, $\sigma_{E_j}(x) = x$ for each $j \in [m]$ satisfying $j > i$ [27]. Also, $\sigma_{E_j}(y) = y$ for each $j \in [m]$ satisfying $j < i$ [28]. Therefore, Lemma 0.22 (applied to $g_j = \sigma_{E_j}$) shows that $(c_{E_1} \circ c_{E_2} \circ \cdots \circ c_{E_m})(x) = y = \sigma(x)$.

---

[24] *Proof.* We have $y = \underbrace{\sigma}_{=\sigma^1}(x) = \sigma^1(x)$. Thus, there exists some $k \in \mathbb{N}$ such that $y = \sigma^k(x)$ (namely, $k = 1$). In other words, $x \sim y$ (since $x \sim y$ if and only if there exists some $k \in \mathbb{N}$ such that $y = \sigma^k(x)$).

[25] *Proof of (46):* Fix $j \in [m]$ distinct from i. We must show that $y \notin E_j$.

Assume the contrary. Thus, $y \in E_j$. Combining this with $y \in E_i$, we find $y \in E_i \cap E_j$. Therefore, the set $E_i \cap E_j$ is nonempty (namely, it contains y). But j is distinct from i. Hence, (44) yields $E_i \cap E_j = \varnothing$. This contradicts the fact that the set $E_i \cap E_j$ is nonempty. This contradiction shows that our assumption was false, qed.

[26] *Proof.* The definition of σ_{E_i} yields

$$\sigma_{E_i}(x) = \begin{cases} \sigma(x), & \text{if } x \in E_i; \\ x, & \text{if } x \notin E_i \end{cases} = \sigma(x) \qquad (\text{since } x \in E_i)$$
$$= y.$$

[27] *Proof.* Let $j \in [m]$ be such that $j > i$. Thus, j is distinct from i (since $j > i$). Hence, (45) shows that $x \notin E_j$. Now, the definition of σ_{E_j} yields

$$\sigma_{E_j}(x) = \begin{cases} \sigma(x), & \text{if } x \in E_j; \\ x, & \text{if } x \notin E_j \end{cases} = x \qquad (\text{since } x \notin E_j).$$

Qed.

[28] *Proof.* Let $j \in [m]$ be such that $j < i$. Thus, j is distinct from i (since $j < i$). Hence, (46) shows that $y \notin E_j$. Now, the definition of σ_{E_j} yields

$$\sigma_{E_j}(y) = \begin{cases} \sigma(y), & \text{if } y \in E_j; \\ y, & \text{if } y \notin E_j \end{cases} = y \qquad (\text{since } y \notin E_j).$$

Now, forget that we fixed x. We thus have shown that $\left(c_{E_1} \circ c_{E_2} \circ \cdots \circ c_{E_m}\right)(x) = \sigma(x)$ for each $x \in X$. In other words, $c_{E_1} \circ c_{E_2} \circ \cdots \circ c_{E_m} = \sigma$. This solves Exercise 7 **(e)**.

References

[BleLau92] D. Blessenohl and H. Laue, *Algebraic combinatorics related to the free Lie algebra*, Séminaire Lotharingien de Combinatoire **29**, page B29e (1992). `https://eudml.org/doc/121542`

[Comtet74] Louis Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, D. Reidel Publishing Company, 1974.

[DanRot78] Ottavio D'Antona, Gian-Carlo Rota, *Two Rings Connected with the Inclusion-Exclusion Principle*, Journal of Combinatorial Theory, Series A, Volume 24, Issue 3, May 1978, pp. 395–402.

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

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

[GriRei18] Darij Grinberg, Victor Reiner, *Hopf algebras in Combinatorics*, version of 11 May 2018, `arXiv:1409.8356v5`. See also `http://www.cip.ifi.lmu.de/~grinberg/algebra/HopfComb-sols.pdf` for a version that gets updated.

$\square$

Qed.