Math 222 Fall 2022, Lecture 27: Permutations

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

4. Permutations

4.3. The cycle decomposition of a permutation

Last time (Lecture 26), we studied permutations and introduced a few ways to denote them. The cycle digraph, in particular, has revealed a few patterns: It always seems to consist of a bunch of cycles that are disjoint (i.e., have no nodes in common).

Let us make this precise and draw some consequences.

4.3.1. Orbits and the relation $\overset{\sigma}{\sim}$

Definition 4.3.1. Let X be a finite set. Let σ be a permutation of X.

Consider the relation $\overset{\sigma}{\sim}$ on the set X defined as follows: For any $i, j \in X$, we set

$$\left(i \overset{\sigma}{\sim} j\right) \iff \left(i = \sigma^k(j) \text{ for some } k \in \mathbb{N}\right).$$

We will soon see that this relation $\overset{\sigma}{\sim}$ is an equivalence relation.[1] But first, let us analyze its equivalence classes. Recall that if $\sim$ is a relation on some set X, and if a is an element of X, then the $\sim$-equivalence class of a is defined to be the set

$$[a]_{\sim} := \{b \in X \mid b \sim a\}.$$

If $\sim$ is the relation $\overset{\sigma}{\sim}$ from Definition 4.3.1, then these $\sim$-equivalence classes have a special name:

Definition 4.3.2. Let X be a finite set. Let σ be a permutation of X. Then, the equivalence classes of the relation $\overset{\sigma}{\sim}$ are called the **orbits** of σ.

Example 4.3.3. Let σ be the permutation of $[10]$ whose OLN is

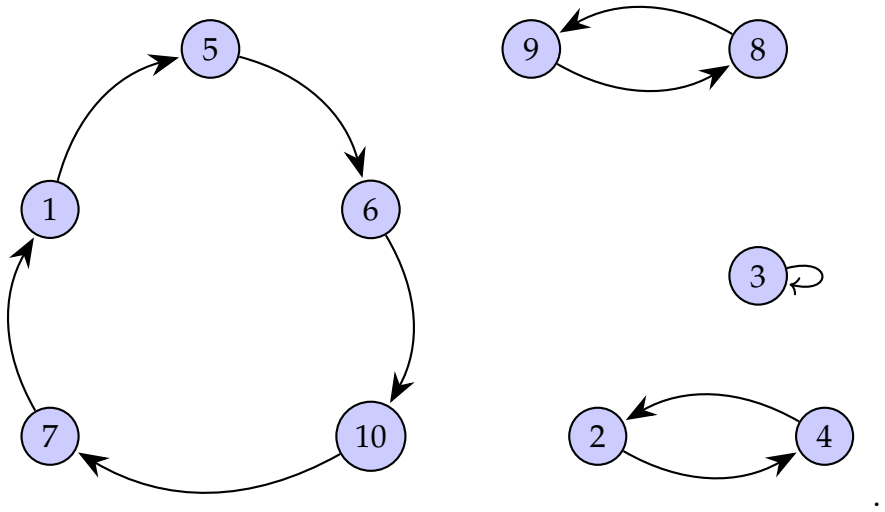
$$5\ 4\ 3\ 2\ 6\ (10)\ 1\ 9\ 8\ 7.$$

[1] Note that this would not be the case if X were allowed to be infinite! For example, the map

$$\sigma : \mathbb{Z} \to \mathbb{Z},\ i \mapsto i+1$$

is a permutation of $\mathbb{Z}$, but the relation $\overset{\sigma}{\sim}$ is not symmetric (since $2 \overset{\sigma}{\sim} 1$ but not $1 \overset{\sigma}{\sim} 2$) and therefore not an equivalence relation.

(We have put the element 10 in parentheses to make its place clearer.) For example, $\sigma(5) = 6$ and $\sigma(6) = 10$. As we recall (Example 4.1.5 in Lecture 26), the cycle digraph of σ is



.

The orbits of σ are

$$\{1,5,6,10,7\}, \qquad \{8,9\}, \qquad \{3\}, \qquad \{2,4\}.$$

Of course, $\{1,5,6,7,10\}$ is an orbit as well, but that's just the same orbit as $\{1,5,6,10,7\}$, so it doesn't count as a fifth orbit.

As we see on this example, the orbits of a permutation σ correspond to the cycles on the cycle digraph of σ. More precisely, each cycle on the cycle digraph of σ results in an orbit, which consists of all nodes on this cycle. However, the orbit is just a set, so it doesn't "remember" the order of the nodes on the cycle; it only remembers which nodes are on the cycle.

This all holds in general, not just in Example 4.3.3. To convince ourselves of this, we shall prove the following proposition, which describes the structure of a single orbit $[a]_\sim$ of σ:

Proposition 4.3.4 (structure of an orbit of σ)**.** Let X be a finite set. Let σ be a permutation of X.

Let $\sim$ be the relation $\overset{\sigma}{\sim}$. Let $a \in X$. Then:

(a) We have

$$[a]_\sim = \left\{ \sigma^k(a) \mid k \in \mathbb{N} \right\} \tag{1}$$

$$= \left\{ \sigma^0(a), \sigma^1(a), \sigma^2(a), \ldots \right\}. \tag{2}$$

In other words, the orbit $[a]_\sim$ of σ consists of all elements of X that can be obtained from a by repeatedly applying the map σ some number of times.

("Some number" allows for the number 0, so the element $a = \sigma^0(a)$ is included as well.)

(b) There exists some positive integer m such that

$$\begin{aligned} \sigma^m(a) &= a \qquad \text{and} \\ [a]_\sim &= \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)\right\} \qquad \text{and} \\ |[a]_\sim| &= m. \end{aligned}$$

Proof. The definition of $[a]_\sim$ yields

$$\begin{aligned} [a]_\sim &= \{b \in X \mid b \sim a\} \\ &= \left\{b \in X \mid b \overset{\sigma}{\sim} a\right\} \qquad \left(\text{since } \sim \text{ is the same as } \overset{\sigma}{\sim}\right) \\ &= \left\{b \in X \mid b = \sigma^k(a) \text{ for some } k \in \mathbb{N}\right\} \\ &\qquad \left(\begin{array}{c} \text{since "}b \overset{\sigma}{\sim} a\text{" means "}b = \sigma^k(a) \text{ for some } k \in \mathbb{N}\text{"} \\ \text{(by Definition 4.3.1)} \end{array}\right) \\ &= \left\{\sigma^k(a) \mid k \in \mathbb{N}\right\} \\ &= \left\{\sigma^0(a),\ \sigma^1(a),\ \sigma^2(a),\ \ldots\right\}. \end{aligned}$$

This proves Proposition 4.3.4 **(a)**.

Now, consider the map

$$\begin{aligned} \{0, 1, \ldots, |X|\} &\to X, \\ i &\mapsto \sigma^i(a). \end{aligned}$$

If this map was injective, then we would have $|\{0, 1, \ldots, |X|\}| \leq |X|$ (by the Pigeonhole Principle for Injections[2]), which would contradict $|\{0, 1, \ldots, |X|\}| = |X| + 1 > |X|$. Thus, this map cannot be injective. In other words, there exist two elements i and j of $\{0, 1, \ldots, |X|\}$ such that $i < j$ and $\sigma^i(a) = \sigma^j(a)$. Such two elements i and j will be called **twins**. Let i and j be two twins with the smallest possible value of j. Thus,

$$i, j \in \{0, 1, \ldots, |X|\} \qquad \text{and} \qquad i < j \qquad \text{and} \qquad \sigma^i(a) = \sigma^j(a).$$

Moreover,

$$\text{if } u \text{ and } v \text{ are any two twins, then } v \geq j \tag{3}$$

(since i and j are two twins with **smallest possible** j).

Note that $i < j$, thus $j > i \geq 0$. Hence, j is a positive integer.

[2] i.e., Theorem 2.4.5 **(a)** in Lecture 16

The map σ is a permutation, thus a bijection. Hence, it has an inverse σ^{-1}.

Now, we claim that $i = 0$. Indeed, assume the contrary. Thus, $i \geq 1$, so that $j \geq 1$ as well (since $i < j$). Applying σ^{-1} to both sides of the equality $\sigma^i(a) = \sigma^j(a)$, we obtain $\sigma^{i-1}(a) = \sigma^{j-1}(a)$ (since $\sigma^{-1}\left(\sigma^k(a)\right) = \sigma^{k-1}(a)$ for each $k \in \mathbb{Z}$). Moreover, both $i-1$ and $j-1$ belong to $\{0, 1, \ldots, |X|\}$ (since $i \geq 1$ and $j \geq 1$) and satisfy $i - 1 < j - 1$ (since $i < j$). Hence, $i-1$ and $j-1$ are two elements of $\{0, 1, \ldots, |X|\}$ such that $i - 1 < j - 1$ and $\sigma^{i-1}(a) = \sigma^{j-1}(a)$. In other words, $i-1$ and $j-1$ are twins. Thus, (3) (applied to $u = i-1$ and $v = j-1$) yields $j - 1 \geq j$. But this is absurd. This contradiction shows that our assumption was false, and therefore $i = 0$ is proved.

Hence, $\sigma^i(a) = \sigma^0(a) = a$, so that

$$a = \sigma^i(a) = \sigma^j(a). \tag{4}$$

In other words,

$$\sigma^j(a) = a. \tag{5}$$

Hence, for each $k \in \mathbb{Z}$, we have

$$\sigma^k(a) = \sigma^{k-j}\left(\underbrace{\sigma^j(a)}_{=a}\right) = \sigma^{k-j}(a). \tag{6}$$

Thus, for each $k \in \mathbb{Z}$, we have

$$\begin{aligned}
\sigma^k(a) &= \sigma^{k-j}(a) && \text{(by (6))} \\
&= \sigma^{k-2j}(a) && \text{(by (6), applied to } k-j \text{ instead of } k) \\
&= \sigma^{k-3j}(a) && \text{(by (6), applied to } k-2j \text{ instead of } k) \\
&= \cdots .
\end{aligned}$$

In other words, the value $\sigma^k(a)$ does not change if we subtract j from k any number of times.

Therefore, if $k \in \mathbb{N}$, then[3]

$$\begin{aligned}
\sigma^k(a) &= \sigma^{k\%j}(a) && \left(\begin{array}{c}\text{since } k\%j \text{ can be obtained from } k \\ \text{by subtracting } j \text{ some number of times}\end{array}\right) \\
&\in \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)\right\}
\end{aligned}$$

(since $k\%j$ is a remainder upon division by j, and therefore must be one of the j numbers $0, 1, \ldots, j-1$). In other words,

$$\left\{\sigma^k(a) \mid k \in \mathbb{N}\right\} \subseteq \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)\right\}.$$

[3]Here, we let $k\%j$ denote the remainder obtained when k is divided by j.

Combining this inclusion with the (obvious) inclusion

$$\left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)\right\} \subseteq \left\{\sigma^k(a) \mid k \in \mathbb{N}\right\},$$

we obtain

$$\left\{\sigma^k(a) \mid k \in \mathbb{N}\right\} = \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)\right\}.$$

In view of (1), we can rewrite this as

$$[a]_\sim = \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)\right\}. \tag{7}$$

Moreover, we claim that the j elements $\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)$ are distinct. Indeed, if two of them were equal (say, $\sigma^u(a) = \sigma^v(a)$ for some $u, v \in \{0,1,\ldots,j-1\}$ satisfying $u < v$), then we would get a contradiction to (3) (since u and v would be twins, and thus (3) would yield $v \geq j$, but this would contradict $v \leq j-1 < j$). Thus, the j elements $\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)$ are distinct. Therefore,

$$\left|\left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{j-1}(a)\right\}\right| = j.$$

In view of (7), we can rewrite this as

$$|[a]_\sim| = j. \tag{8}$$

We have now proved the three equalities (5), (7) and (8). Since j is a positive integer, we thus conclude that there exists some positive integer m such that

$$\begin{aligned} \sigma^m(a) &= a \qquad \text{and} \\ [a]_\sim &= \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)\right\} \qquad \text{and} \\ |[a]_\sim| &= m \end{aligned}$$

(namely, $m = j$). This proves Proposition 4.3.4 **(b)**. □

Proposition 4.3.5. Let X be a finite set. Let σ be a permutation of X. Then, the relation $\overset{\sigma}{\sim}$ is an equivalence relation.

Proof. We must prove that it satisfies three axioms: reflexivity, symmetry and transitivity.

- Reflexivity is easy: For any $a \in X$, we have $a \overset{\sigma}{\sim} a$, since $a = \sigma^k(a)$ for some $k \in \mathbb{N}$ (namely, for $k = 0$).

- Transitivity is fairly easy as well: Let $a, b, c \in X$ be such that $a \overset{\sigma}{\sim} b$ and $b \overset{\sigma}{\sim} c$. From $a \overset{\sigma}{\sim} b$, we obtain $a = \sigma^i(b)$ for some $i \in \mathbb{N}$. From $b \overset{\sigma}{\sim} c$, we obtain $b = \sigma^j(c)$ for some $j \in \mathbb{N}$. Consider these i and j. Now,

$$a = \sigma^i\left(\underbrace{b}_{=\sigma^j(c)}\right) = \sigma^i\left(\sigma^j(c)\right) = \sigma^{i+j}(c).$$

  Hence, $a = \sigma^k(c)$ for some $k \in \mathbb{N}$ (namely, for $k = i + j$). In other words, $a \overset{\sigma}{\sim} c$. This proves transitivity.

- Let us now prove symmetry. Indeed, let $a, b \in X$ be such that $a \overset{\sigma}{\sim} b$. Denoting the relation $\overset{\sigma}{\sim}$ by $\sim$, we thus have $a \sim b$. Hence, $a \in [b]_\sim$ (since the $\sim$-equivalence class $[b]_\sim$ is defined to consist of all $c \in X$ that satisfy $c \sim b$).

  However, Proposition 4.3.4 **(b)** (applied to b instead of a) yields that there exists some positive integer m such that

$$\begin{aligned} &\sigma^m(b) = b \qquad \text{and} \\ &[b]_\sim = \left\{\sigma^0(b), \sigma^1(b), \ldots, \sigma^{m-1}(b)\right\} \qquad \text{and} \\ &|[b]_\sim| = m. \end{aligned}$$

  Consider this m. From $a \in [b]_\sim = \left\{\sigma^0(b), \sigma^1(b), \ldots, \sigma^{m-1}(b)\right\}$, we obtain that $a = \sigma^i(b)$ for some $i \in \{0, 1, \ldots, m-1\}$. Consider this i. Then, $i \leq m - 1 < m$, so that $m - i \in \mathbb{N}$. Now, $\sigma^{m-i}\left(\underbrace{a}_{=\sigma^i(b)}\right) = \sigma^{m-i}\left(\sigma^i(b)\right) = \sigma^m(b) = b$, so that $b = \sigma^{m-i}(a)$. Hence, $b = \sigma^k(a)$ for some $k \in \mathbb{N}$ (namely, for $k = m - i$). In other words, $b \overset{\sigma}{\sim} a$. This proves symmetry.

Altogether, we have now proved all three axioms, so that Proposition 4.3.5 is proven. □

Proposition 4.3.5 tells us that $\overset{\sigma}{\sim}$ is an equivalence relation, so that the orbits of σ "behave well":

Corollary 4.3.6. Let X be a finite set. Let σ be a permutation of X. Then, the set of all orbits of σ is a set partition of X. In particular, each element of X belongs to precisely one orbit of σ.

Proof. The orbits of σ are the $\overset{\sigma}{\sim}$-equivalence classes (by their definition). Since $\overset{\sigma}{\sim}$ is an equivalence relation, each element of X belongs to exactly one $\overset{\sigma}{\sim}$-equivalence class (namely, its own)[4]. In other words, each element of X belongs

[4]Here, we have used Theorem 3.3.11 from Lecture 24.

to precisely one orbit of σ. Hence, the orbits of σ are disjoint, and their union is X. Moreover, each orbit of σ is nonempty (since the relation $\overset{\sigma}{\sim}$ satisfies the reflexivity axiom, and thus each of its equivalence classes $[a]_{\overset{\sigma}{\sim}}$ contains at least the element a). Therefore, the set of all orbits of σ is a set partition of X. Thus, Corollary 4.3.6 is proved. □

For future use, we note the following:

Remark 4.3.7. Let X be a finite set. Let σ be a permutation of X. Let $a \in X$. Then, a is a fixed point of σ if and only if the orbit of σ containing a is a 1-element set.

Proof. Let $\sim$ be the relation $\overset{\sigma}{\sim}$. Then, the orbit of σ containing a is the $\sim$-equivalence class $[a]_\sim$. Thus, we need to prove that a is a fixed point of σ if and only if $[a]_\sim$ is a 1-element set. Let us prove the "$\Longrightarrow$" and "$\Longleftarrow$" directions of this statement separately:

$\Longrightarrow$: Assume that a is a fixed point of σ. Then, $\sigma(a) = a$. Hence, by induction, we see that $\sigma^k(a) = a$ for each $k \in \mathbb{N}$. However, (1) yields

$$[a]_\sim = \left\{ \underbrace{\sigma^k(a)}_{=a} \mid k \in \mathbb{N} \right\} = \{a \mid k \in \mathbb{N}\} = \{a\}.$$

Thus, $[a]_\sim$ is a 1-element set. This proves the "$\Longrightarrow$" direction.

$\Longleftarrow$: Assume that $[a]_\sim$ is a 1-element set. Thus, $[a]_\sim = \{a\}$ (since $[a]_\sim$ contains a).

Now, $\sigma(a) \overset{\sigma}{\sim} a$, since we have $\sigma(a) = \sigma^k(a)$ for some $k \in \mathbb{N}$ (namely, for $k = 1$). Hence, $\sigma(a) \in [a]_\sim = \{a\}$. In other words, $\sigma(a) = a$. In other words, a is a fixed point of σ. This proves the "$\Longleftarrow$" direction. □

4.3.2. The DCD (= disjoint cycle decomposition) of a permutation

The orbits of a permutation σ do not (in general) determine σ. For example, the 3-cycles $\mathrm{cyc}_{1,2,3}$ and $\mathrm{cyc}_{1,3,2}$ have the same orbits (since an orbit is just a set, and we have $\{1,2,3\} = \{1,3,2\}$). In order to properly encode the information from the cycle digraph, we need to also remember how a permutation acts on each orbit.

Looking back at Example 4.3.3, we see that if we restrict a permutation σ to a given orbit of σ, then it becomes a cycle (more precisely, an m-cycle, where m is the size of the orbit). This follows almost immediately from Proposition 4.3.4 **(b)**:

Proposition 4.3.8 (structure of an orbit of σ, part 2)**.** Let X be a finite set. Let σ be a permutation of X.

Let $\sim$ be the relation $\overset{\sigma}{\sim}$. Let $a \in X$. Let m be the positive integer whose existence is claimed in Proposition 4.3.4 **(b)**. Then, the permutation σ and the m-cycle

$$\operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)} \in S_X$$

are equal on the orbit $[a]_\sim$. In other words, for any $b \in [a]_\sim$, we have $\sigma(b) = \operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)}(b)$.

Proof. Let $b \in [a]_\sim$. We must prove that $\sigma(b) = \operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)}(b)$.

The properties of m stated in Proposition 4.3.4 **(b)** yield $\sigma^m(a) = a$ and $[a]_\sim = \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)\right\}$.

We have $b \in [a]_\sim = \left\{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)\right\}$ and therefore $b = \sigma^k(a)$ for some $k \in \{0, 1, \ldots, m-1\}$. Consider this k. From $b = \sigma^k(a)$, we obtain

$$\sigma(b) = \sigma\left(\sigma^k(a)\right) = \sigma^{k+1}(a). \tag{9}$$

On the other hand,

$$\begin{aligned}
&\operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)}\left(\underbrace{b}_{=\sigma^k(a)}\right) \\
&= \operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)}\left(\sigma^k(a)\right) \\
&= \begin{cases} \sigma^{k+1}(a), & \text{if } k < m-1; \\ \sigma^0(a), & \text{if } k = m-1 \end{cases}
\end{aligned} \tag{10}$$

(by the definition of the m-cycle $\operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)}$).

We claim that the right hand sides of the equalities (10) and (9) are equal. Indeed, they are clearly equal in the case when $k < m-1$ (since both of them are $\sigma^{k+1}(a)$ in this case). However, they are also equal in the case when $k = m-1$ (because in this case, we have $k+1 = m$ and thus $\sigma^{k+1}(a) = \sigma^m(a) = a = \sigma^0(a)$, so that $\sigma^0(a) = \sigma^{k+1}(a)$). Thus, we know that the right hand sides of the equalities (10) and (9) are always equal. Thus, so are the left hand sides. In other words,

$$\operatorname{cyc}_{\sigma^0(a),\ \sigma^1(a),\ \ldots,\ \sigma^{m-1}(a)}(b) = \sigma(b).$$

This completes the proof of Proposition 4.3.8. □

Proposition 4.3.8 describes how a permutation σ of a finite set X acts on a given orbit of σ. However, by understanding a permutation σ on each of its orbits, we gain an understanding of the whole permutation σ on the entire set X. This is best illustrated on an example:

Example 4.3.9. Let $X = [10]$, and let σ be the permutation in Example 4.3.3. Its orbits are $\{1,5,6,10,7\}$, $\{8,9\}$, $\{3\}$ and $\{2,4\}$. From Proposition 4.3.8, we know that σ equals the cycles $\mathrm{cyc}_{1,5,6,10,7}$, $\mathrm{cyc}_{8,9}$, cyc_3 and $\mathrm{cyc}_{2,4}$ on these respective orbits. We claim that the whole permutation σ can therefore be written as

$$\sigma = \mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_{8,9} \circ \mathrm{cyc}_3 \circ \mathrm{cyc}_{2,4}\,. \tag{11}$$

Why is this true?

For example, let us apply both sides of (11) to the element $8 \in X$. The left hand side clearly yields $\sigma(8) = 9$. Applying the right hand side, we obtain

$$\begin{aligned}
&\left(\mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_{8,9} \circ \mathrm{cyc}_3 \circ \mathrm{cyc}_{2,4}\right)(8) \\
&= \mathrm{cyc}_{1,5,6,10,7}\left(\mathrm{cyc}_{8,9}\left(\mathrm{cyc}_3\left(\underbrace{\mathrm{cyc}_{2,4}(8)}_{=8}\right)\right)\right) \\
&= \mathrm{cyc}_{1,5,6,10,7}\left(\mathrm{cyc}_{8,9}\left(\underbrace{\mathrm{cyc}_3(8)}_{=8}\right)\right) \\
&= \mathrm{cyc}_{1,5,6,10,7}\left(\underbrace{\mathrm{cyc}_{8,9}(8)}_{=9}\right) \\
&= \mathrm{cyc}_{1,5,6,10,7}(9) \\
&= 9.
\end{aligned}$$

We note that this calculation followed a very simple pattern: We started with the element 8 and successively applied the cycles $\mathrm{cyc}_{2,4}$, cyc_3, $\mathrm{cyc}_{8,9}$ and $\mathrm{cyc}_{1,5,6,10,7}$ to it. The first two of these cycles left 8 unchanged (since $8 \notin \{2,4\}$ and $8 \notin \{3\}$). The next cycle (that is, $\mathrm{cyc}_{8,9}$) sent 8 to 9. Finally, the last cycle ($\mathrm{cyc}_{1,5,6,10,7}$) left 9 unchanged again (since $9 \notin \{1,5,6,10,7\}$). The reason for this simple behavior is that each of the four cycles comes from an orbit of σ, but our element 8 lies in only one orbit, and thus is moved by only one cycle, whereas the remaining cycles (whether they are applied before or after its move) leave it unchanged.

Of course, there was nothing special about the element 8 that we used here. We can likewise see that for any $x \in X$, we have

$$\sigma(x) = \left(\mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_{8,9} \circ \mathrm{cyc}_3 \circ \mathrm{cyc}_{2,4}\right)(x)\,.$$

In other words, $\sigma = \mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_{8,9} \circ \mathrm{cyc}_3 \circ \mathrm{cyc}_{2,4}$. Thus, (11) is proved.

Note that each of the four cycles on the right hand side of (11) can be rewritten by cyclically rotating its subscripts: For instance, $\mathrm{cyc}_{1,5,6,10,7}$ can be rewritten as $\mathrm{cyc}_{5,6,10,7,1}$ or as $\mathrm{cyc}_{6,10,7,1,5}$ or in two more ways. Furthermore,

the four cycles on the right hand side of (11) can be swapped at will: For instance, we have

$$\begin{aligned}
\sigma &= \underbrace{\mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_{8,9}}_{=\mathrm{cyc}_{8,9} \circ \mathrm{cyc}_{1,5,6,10,7}} \circ \mathrm{cyc}_3 \circ \mathrm{cyc}_{2,4} \\
&= \mathrm{cyc}_{8,9} \circ \underbrace{\mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_3}_{=\mathrm{cyc}_3 \circ \mathrm{cyc}_{1,5,6,10,7}} \circ \mathrm{cyc}_{2,4} \\
&= \mathrm{cyc}_{8,9} \circ \mathrm{cyc}_3 \circ \underbrace{\mathrm{cyc}_{1,5,6,10,7} \circ \mathrm{cyc}_{2,4}}_{=\mathrm{cyc}_{2,4} \circ \mathrm{cyc}_{1,5,6,10,7}} \\
&= \mathrm{cyc}_{8,9} \circ \mathrm{cyc}_3 \circ \mathrm{cyc}_{2,4} \circ \mathrm{cyc}_{1,5,6,10,7} = \cdots .
\end{aligned}$$

The reason for this is the following (very easily proved) fact:

Proposition 4.3.10. Let X be a set. Let $b_1, b_2, \ldots, b_i, c_1, c_2, \ldots, c_j$ be distinct elements of X. Then,

$$\mathrm{cyc}_{b_1,b_2,\ldots,b_i} \circ \mathrm{cyc}_{c_1,c_2,\ldots,c_j} = \mathrm{cyc}_{c_1,c_2,\ldots,c_j} \circ \mathrm{cyc}_{b_1,b_2,\ldots,b_i} .$$

Proof. Easy and left to the reader. □

Generalizing the reasoning from Example 4.3.9, we find the following:

Theorem 4.3.11 (disjoint cycle decomposition of a permutation)**.** Let X be a finite set. Let σ be a permutation of X. Then:

(a) There is a list

$$\begin{pmatrix} \left(a_{1,1}, a_{1,2}, \ldots, a_{1,m_1}\right), \\ \left(a_{2,1}, a_{2,2}, \ldots, a_{2,m_2}\right), \\ \ldots, \\ \left(a_{k,1}, a_{k,2}, \ldots, a_{k,m_k}\right) \end{pmatrix}$$

of nonempty lists of elements of X such that:

- Each element of X appears exactly once in the composite list

$$\begin{pmatrix} a_{1,1}, a_{1,2}, \ldots, a_{1,m_1}, \\ a_{2,1}, a_{2,2}, \ldots, a_{2,m_2}, \\ \ldots, \\ a_{k,1}, a_{k,2}, \ldots, a_{k,m_k} \end{pmatrix}.$$

- We have

$$\sigma = \operatorname{cyc}_{a_{1,1},a_{1,2},\ldots,a_{1,m_1}} \circ \operatorname{cyc}_{a_{2,1},a_{2,2},\ldots,a_{2,m_2}} \circ \cdots \circ \operatorname{cyc}_{a_{k,1},a_{k,2},\ldots,a_{k,m_k}}.$$

Such a list is called a **disjoint cycle decomposition** (short: **DCD**) of σ. Its entries (which themselves are lists of elements of X) are called the **cycles** of σ.

(b) Any two DCDs of σ can be obtained from one another by (repeatedly) swapping the cycles with each other and rotating each cycle (i.e., replacing $\left(a_{i,1}, a_{i,2}, \ldots, a_{i,m_i}\right)$ by $\left(a_{i,2}, a_{i,3}, \ldots, a_{i,m_i}, a_{i,1}\right)$).

(c) Now assume that X is a set of integers (or, more generally, any totally ordered set). Then, there is a unique DCD

$$\begin{pmatrix} \left(a_{1,1}, a_{1,2}, \ldots, a_{1,m_1}\right), \\ \left(a_{2,1}, a_{2,2}, \ldots, a_{2,m_2}\right), \\ \ldots, \\ \left(a_{k,1}, a_{k,2}, \ldots, a_{k,m_k}\right) \end{pmatrix}$$

of σ that satisfies the following two additional requirements:

- We have $a_{i,1} \leq a_{i,p}$ for each $i \in [k]$ and each $p \in [m_i]$ (that is, each cycle in the DCD is written with its smallest entry first).
- We have $a_{1,1} > a_{2,1} > \cdots > a_{k,1}$ (that is, the cycles appear in the DCD in the order of decreasing first entries).

Proof of Theorem 4.3.11 (sketched). **(a)** Let $\sim$ be the relation $\overset{\sigma}{\sim}$.

Let $X_1, X_2, \ldots, X_k$ be the orbits of σ (listed with no repetition).

Let $i \in [k]$. Then, the set X_i is an orbit of σ, that is, a $\sim$-equivalence class (by the definition of an orbit). Hence, it can be written in the form $X_i = [x_i]_\sim$ for some $x_i \in X$. Consider this x_i. Proposition 4.3.4 **(b)** (applied to $a = x_i$) yields that there exists some positive integer m_i such that

$$\sigma^{m_i}(x_i) = x_i \qquad \text{and} \tag{12}$$

$$[x_i]_\sim = \left\{\sigma^0(x_i),\ \sigma^1(x_i),\ \ldots,\ \sigma^{m_i-1}(x_i)\right\} \qquad \text{and} \tag{13}$$

$$|[x_i]_\sim| = m_i. \tag{14}$$

Consider this m_i.

Proposition 4.3.8 (applied to $a = x_i$ and $m = m_i$) yields that the permutation σ and the m_i-cycle

$$\operatorname{cyc}_{\sigma^0(x_i),\ \sigma^1(x_i),\ \ldots,\ \sigma^{m_i-1}(x_i)} \in S_X$$

are equal on the orbit $[x_i]_\sim = X_i$.

Forget that we fixed i. Thus, for each $i \in [k]$, we have found an $x_i \in X$ and a positive integer m_i satisfying (12), (13) and (14) and with the property that the permutation σ and the m_i-cycle

$$\operatorname{cyc}_{\sigma^0(x_i),\, \sigma^1(x_i),\, \ldots,\, \sigma^{m_i-1}(x_i)} \in S_X$$

are equal on the orbit $[x_i]_\sim = X_i$. Hence, each element of X appears exactly once in the list

$$\begin{aligned}
&(\sigma^0(x_1),\ \sigma^1(x_1),\ \ldots,\ \sigma^{m_1-1}(x_1), \\
&\quad \sigma^0(x_2),\ \sigma^1(x_2),\ \ldots,\ \sigma^{m_2-1}(x_2), \\
&\quad \ldots, \\
&\quad \sigma^0(x_k),\ \sigma^1(x_k),\ \ldots,\ \sigma^{m_k-1}(x_k)\ )
\end{aligned}$$

(by (12) and (13), since each element of X belongs to exactly one of the orbits $X_1, X_2, \ldots, X_k$), and furthermore, we have

$$\begin{aligned}
\sigma = \operatorname{cyc}_{\sigma^0(x_1),\, \sigma^1(x_1),\, \ldots,\, \sigma^{m_1-1}(x_1)} &\circ \operatorname{cyc}_{\sigma^0(x_2),\, \sigma^1(x_2),\, \ldots,\, \sigma^{m_2-1}(x_2)} \\
&\circ \cdots \circ \operatorname{cyc}_{\sigma^0(x_k),\, \sigma^1(x_k),\, \ldots,\, \sigma^{m_k-1}(x_k)}
\end{aligned}$$

(this can be proved by the same reasoning as we used in Example 4.3.9 to prove (11)). Thus, the list

$$\begin{aligned}
&\Big(\big(\sigma^0(x_1),\ \sigma^1(x_1),\ \ldots,\ \sigma^{m_1-1}(x_1)\big), \\
&\quad \big(\sigma^0(x_2),\ \sigma^1(x_2),\ \ldots,\ \sigma^{m_2-1}(x_2)\big), \\
&\quad \ldots, \\
&\quad \big(\sigma^0(x_k),\ \sigma^1(x_k),\ \ldots,\ \sigma^{m_k-1}(x_k)\big)\ \Big)
\end{aligned}$$

is a DCD of σ. This proves that a DCD of σ exists. In other words, Theorem 4.3.11 **(a)** is proven.

(For alternative proofs of Theorem 4.3.11 **(a)**, see (e.g.) [21s, proof of Theorem 5.5.2 **(a)**], [17f-hw7s, Exercise 7 **(e)** and **(d)**], [Goodma15, Theorem 1.5.3], [Bourba74, Chapter I, §5.7, Proposition 7], or `https://proofwiki.org/wiki/Existence_and_Uniqueness_of_Cycle_Decomposition` (but see Remark 4.3.12 further below for a certain disagreement about the definition of a DCD).)

(b) See [Goodma15, Theorem 1.5.3] or [Bourba74, Chapter I, §5.7, Proposition 7]. The idea is fairly simple: Let

$$\begin{aligned}
&\Big(\big(a_{1,1}, a_{1,2}, \ldots, a_{1,m_1}\big), \\
&\quad \big(a_{2,1}, a_{2,2}, \ldots, a_{2,m_2}\big), \\
&\quad \ldots, \\
&\quad \big(a_{k,1}, a_{k,2}, \ldots, a_{k,m_k}\big)\ \Big)
\end{aligned}$$

be a DCD of σ. Then, for each $i \in X$, the cycle of this DCD that contains i is uniquely determined by σ and i up to cyclic rotation (indeed, it is a rotated version of the list $\left(i, \sigma\left(i\right), \sigma^2\left(i\right), \ldots, \sigma^{r-1}\left(i\right)\right)$, where r is the smallest positive integer satisfying $\sigma^r\left(i\right) = i$). Therefore, all cycles of this DCD are uniquely determined by σ up to cyclic rotation and up to the relative order in which these cycles appear in the DCD. But this is precisely the claim of Theorem 4.3.11 **(b)**.

(c) Clearly, such a DCD of σ exists: Indeed, we can start with an arbitrary DCD of σ (this exists because of Theorem 4.3.11 **(a)**), and then transform it using cyclic rotations and swaps into a form that satisfies the two additional requirements of Theorem 4.3.11 **(c).** Namely, we

- first cyclically rotate each cycle to ensure that it begins with its smallest entry, and
- then we swap the cycles appropriately to ensure that they appear in the order of decreasing first entries.

It remains to prove that a DCD of σ satisfying the two additional requirements of Theorem 4.3.11 **(c)** is unique. But this follows easily from Theorem 4.3.11 **(b)**: Indeed, Theorem 4.3.11 **(b)** shows that any two such DCDs can be transformed into one another by rotating each cycle and swapping the cycles; however, the two additional requirements uniquely determine the first entry of each cycle and also the order of the cycles, and thus there is no freedom left for two different DCDs to fit the bill. □

Remark 4.3.12. Some authors omit 1-element cycles from the DCD of a permutation, since we have $\mathrm{cyc}_i = \mathrm{id}$ for each $i \in X$. However, if you do this, then the condition "Each element of X appears exactly once in the composite list" in the definition of a DCD has to be replaced by "Each element of X appears at most once in the composite list".

4.3.3. Composing with a transposition

Before we move on with counting questions, we shall prove a lemma that we will soon need. The lemma is about how a transposition $t_{i,j}$ affects the orbits of a permutation σ when it is multiplied onto σ from the left – i.e., about how the orbits of $t_{i,j} \circ \sigma$ differ from those of σ.

Recall the notion of a transposition $t_{i,j}$, defined in Definition 4.2.1 (Lecture 26): If i and j are two distinct elements of a set X, then the transposition $t_{i,j}$ is the permutation of X that swaps i with j, while leaving all other elements of X unchanged. We recall that this permutation $t_{i,j}$ satisfies $t_{i,j} \circ t_{i,j} = \mathrm{id}$ (since swapping i with j twice results in every element of X returning where it was).

Lemma 4.3.13. Let X be a finite set. Let i and j be two distinct elements of X. Let $\sigma \in S_X$ be a permutation, and let $\tau = t_{i,j} \circ \sigma$. Then:

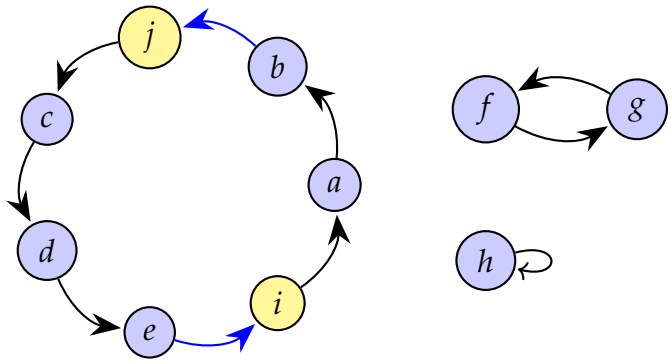
(a) If $i \overset{\sigma}{\sim} j$, then the permutation τ has 1 more orbit than σ.

(b) If $i \overset{\sigma}{\sim} j$, then we don't have $i \overset{\tau}{\sim} j$.

(c) If we don't have $i \overset{\sigma}{\sim} j$, then the permutation τ has 1 fewer orbit than σ.

(d) If we don't have $i \overset{\sigma}{\sim} j$, then we have $i \overset{\tau}{\sim} j$.

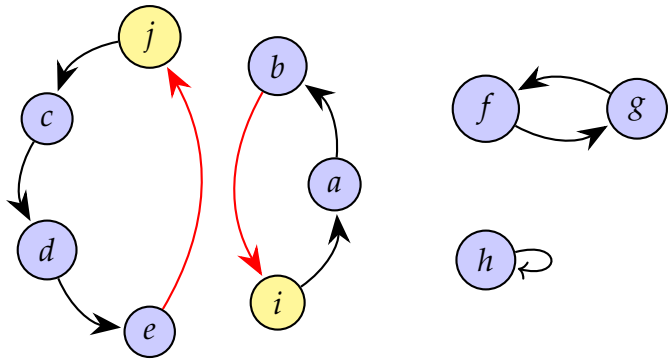
Example 4.3.14. Let X be a 10-element set $\{a, b, c, \ldots, j\}$, and let $\sigma \in S_X$ be the permutation whose cycle digraph looks as follows:



.

Note that i and j visibly belong to the same orbit of σ, so that $i \overset{\sigma}{\sim} j$ (and, specifically, $j = \sigma^3(i)$).

Let $\tau = t_{i,j} \circ \sigma$. This permutation τ has the following cycle digraph:



.

Observe that the cycle digraphs of σ and τ differ only in two of their arcs: namely, the arcs that end in the nodes i and j. (These are the two arcs that we colored blue in the cycle digraph of σ and red in the cycle digraph of τ.) The change from σ to τ causes these two arcs to swap their targets (= ending points). All other arcs of the cycle digraph of σ (colored black) appear equally in the cycle digraph of τ. By comparing the two cycle digraphs, we see that the cycle digraph of τ contains 1 more cycle than that of σ. In other words, the permutation τ has 1 more orbit than σ. This confirms Lemma 4.3.13 **(a)** in our example. Furthermore, i and j lie in different cycles of the cycle digraph

of τ, thus belong to different orbits of τ. In other words, we don't have $i \overset{\tau}{\sim} j$. This confirms Lemma 4.3.13 **(b)** in our example. Essentially, by composing $t_{i,j}$ with σ, we have "broken up" the cycle of σ that contained both i and j into two smaller cycles, one of which contains i while the other contains j.

The same example can also be used to illustrate parts **(c)** and **(d)** of Lemma 4.3.13, once we swap σ with τ. Indeed, from $\tau = t_{i,j} \circ \sigma$, we obtain $\sigma = t_{i,j} \circ \tau$ (because $t_{i,j} \circ \underbrace{\tau}_{=t_{i,j} \circ \sigma} = \underbrace{t_{i,j} \circ t_{i,j}}_{=\mathrm{id}} \circ \sigma = \mathrm{id} \circ \sigma = \sigma$), and therefore the relationship between σ and τ is mutual. As we know that we don't have $i \overset{\tau}{\sim} j$, we thus see that the assumptions of Lemma 4.3.13 **(c)** and **(d)** are satisfied if we swap σ with τ. The conclusions, too, are satisfied, as we see from the cycle digraphs. Essentially, when we compose $t_{i,j}$ with τ, the two cycles containing i and j get "merged" into a common cycle.

Proof of Lemma 4.3.13 (sketched). We generalize and formalize what we already saw in Example 4.3.14. If a and b are two elements of X, then the notation "$a \overset{\sigma}{\mapsto} b$" shall mean "$\sigma(a) = b$", whereas the notation "$a \overset{\tau}{\mapsto} b$" shall mean "$\tau(a) = b$". We note that if k elements $a_1, a_2, \ldots, a_k$ of X (with $k \geq 1$) form a circular chain of the form

$$a_1 \overset{\tau}{\mapsto} a_2 \overset{\tau}{\mapsto} \cdots \overset{\tau}{\mapsto} a_k \overset{\tau}{\mapsto} a_1, \tag{15}$$

then the set $\{a_1, a_2, \ldots, a_k\}$ must be an orbit of τ (since (15) shows that $\{a_1, a_2, \ldots, a_k\} = \left\{\tau^0(a_1),\ \tau^1(a_1),\ \tau^2(a_1),\ \ldots\right\}$, but we know from (2) (applied to τ and a_1 instead of σ and a) that the set $\left\{\tau^0(a_1),\ \tau^1(a_1),\ \tau^2(a_1),\ \ldots\right\}$ is an orbit of τ).

(a) Assume that $i \overset{\sigma}{\sim} j$. Thus, i and j belong to the $\overset{\sigma}{\sim}$-equivalence class, i.e., to the same orbit of σ. Let X_1 be this orbit, and let $X_2, X_3, \ldots, X_k$ be all the remaining orbits of σ (listed without repetition). Thus, σ has k orbits. We shall now understand what the orbits of τ are.

Let $\sim$ be the relation $\overset{\sigma}{\sim}$. Thus, $i \sim j$ (since $i \overset{\sigma}{\sim} j$). Proposition 4.3.4 **(b)** (applied to $a = i$) yields that there exists some positive integer m such that

$$\begin{aligned} \sigma^m(i) &= i \qquad \text{and} \\ [i]_\sim &= \left\{\sigma^0(i),\ \sigma^1(i),\ \ldots,\ \sigma^{m-1}(i)\right\} \qquad \text{and} \\ |[i]_\sim| &= m. \end{aligned}$$

Consider this m. From $i \sim j$, we obtain $j \in [i]_\sim = \left\{\sigma^0(i),\ \sigma^1(i),\ \ldots,\ \sigma^{m-1}(i)\right\}$. In other words, $j = \sigma^p(i)$ for some $p \in \{0, 1, \ldots, m-1\}$. Consider this p. Thus, $\sigma^p(i) = j \neq i = \sigma^0(i)$, so that $p \neq 0$. Hence, $p \in [m-1]$ (since $p \in \{0, 1, \ldots, m-1\}$ but $p \neq 0$).

Recall that X_1 is the orbit of σ that contains i. In other words, $X_1 = [i]_\sim$ (since $[i]_\sim$ is the orbit of σ that contains i). Therefore,

$$X_1 = [i]_\sim = \left\{\sigma^0(i),\ \sigma^1(i),\ \ldots,\ \sigma^{m-1}(i)\right\} = \left\{i,\ \sigma^1(i),\ \sigma^2(i),\ \ldots,\ \sigma^{m-1}(i)\right\}$$

(since $\sigma^0(i) = i$). The action of the map σ on the elements of X_1 looks as follows:

$$i \overset{\sigma}{\mapsto} \sigma^1(i) \overset{\sigma}{\mapsto} \sigma^2(i) \overset{\sigma}{\mapsto} \cdots \overset{\sigma}{\mapsto} \sigma^{m-1}(i) \overset{\sigma}{\mapsto} i$$

(since $\sigma\left(\sigma^{m-1}(i)\right) = \sigma^m(i) = i$). Since $j = \sigma^p(i)$, we can rewrite the element $\sigma^p(i)$ on this circular chain as j, so that the chain becomes

$$\begin{aligned} & i \overset{\sigma}{\mapsto} \sigma^1(i) \overset{\sigma}{\mapsto} \sigma^2(i) \overset{\sigma}{\mapsto} \cdots \overset{\sigma}{\mapsto} \sigma^{p-1}(i) \\ & \overset{\sigma}{\mapsto} j \overset{\sigma}{\mapsto} \sigma^{p+1}(i) \overset{\sigma}{\mapsto} \sigma^{p+2}(i) \overset{\sigma}{\mapsto} \cdots \overset{\sigma}{\mapsto} \sigma^{m-1}(i) \overset{\sigma}{\mapsto} i. \end{aligned} \tag{16}$$

However, the transposition $t_{i,j}$ swaps i with j while leaving all other elements unchanged. Thus, the permutation $\tau = t_{i,j} \circ \sigma$ differs from σ only in two things:

- The element of X that is sent to i by σ is instead sent to j by τ (because $t_{i,j}$ sends i to j).
- The element of X that is sent to j by σ is instead sent to i by τ (because $t_{i,j}$ sends j to i).
- All other elements of X are sent by τ to the same value that they are sent to by σ.

Hence, when we replace σ by τ, then the two arrows $\sigma^{p-1}(i) \overset{\sigma}{\mapsto} j$ and $\sigma^{m-1}(i) \overset{\sigma}{\mapsto} i$ in the circular chain (16) turn into $\sigma^{p-1}(i) \overset{\tau}{\mapsto} i$ and $\sigma^{m-1}(i) \overset{\tau}{\mapsto} j$, whereas all the other arrows remain unchanged. Hence, the circular chain (16) turns into two separate circular chains

$$\begin{aligned} & i \overset{\tau}{\mapsto} \sigma^1(i) \overset{\tau}{\mapsto} \sigma^2(i) \overset{\tau}{\mapsto} \cdots \overset{\tau}{\mapsto} \sigma^{p-1}(i) \overset{\tau}{\mapsto} i \qquad \text{and} \\ & j \overset{\tau}{\mapsto} \sigma^{p+1}(i) \overset{\tau}{\mapsto} \sigma^{p+2}(i) \overset{\tau}{\mapsto} \cdots \overset{\tau}{\mapsto} \sigma^{m-1}(i) \overset{\tau}{\mapsto} j. \end{aligned}$$

Hence, in place of the orbit X_1 of σ, the permutation τ has two disjoint orbits

$$\begin{aligned} X_1' &:= \left\{ i,\ \sigma^1(i),\ \sigma^2(i),\ \ldots,\ \sigma^{p-1}(i) \right\} \qquad \text{and} \\ X_1'' &:= \left\{ j,\ \sigma^{p+1}(i),\ \sigma^{p+2}(i),\ \ldots,\ \sigma^{m-1}(i) \right\}. \end{aligned}$$

All the remaining orbits $X_2, X_3, \ldots, X_k$ of σ remain orbits of τ (because the map τ sends them to the same values that the map σ sends them to[5]).

To summarize: we have shown that the orbit X_1 of σ splits into two orbits X_1' and X_1'' of τ (in the sense that X_1' and X_1'' are two disjoint orbits of τ, and their union is X_1),

[5] *Proof.* Let x be an element of one of the orbits $X_2, X_3, \ldots, X_k$. We must prove that $\tau(x) = \sigma(x)$.

The orbits $X_1, X_2, \ldots, X_k$ of σ are disjoint. Hence, x does not belong to X_1 (since x belongs to one of the orbits $X_2, X_3, \ldots, X_k$).

We have $\sigma(x) \overset{\sigma}{\sim} x$ (since $\sigma(x) = \sigma^k(x)$ for $k = 1$). Thus, the elements $\sigma(x)$ and x belong to the same $\overset{\sigma}{\sim}$-equivalence class, i.e., to the same orbit of σ. This shows that the element $\sigma(x)$ does not belong to the orbit X_1 (since x does not belong to X_1). Hence, $\sigma(x)$ equals neither i nor j (since both i and j belong to X_1). Therefore, $t_{i,j}(\sigma(x)) = \sigma(x)$ (since the transposition $t_{i,j}$ leaves every element other than i and j fixed).

Now, $\tau = t_{i,j} \circ \sigma$, so that $\tau(x) = \left(t_{i,j} \circ \sigma\right)(x) = t_{i,j}(\sigma(x)) = \sigma(x)$, qed.

whereas all the remaining orbits $X_2, X_3, \ldots, X_k$ of σ remain orbits of τ. Therefore, τ has 1 more orbit than σ. This proves Lemma 4.3.13 **(a)**.

(b) We continue with the notations from our above proof of Lemma 4.3.13 **(a)**. The two orbits X_1' and X_1'' of τ are disjoint, thus different, and furthermore they contain i and j, respectively. Hence, i and j belong to different orbits of τ. In other words, i and j belong to different $\overset{\tau}{\sim}$-equivalence classes. Hence, we don't have $i \overset{\tau}{\sim} j$. This proves Lemma 4.3.13 **(b)**.

(c) Assume that we don't have $i \overset{\sigma}{\sim} j$. Thus, i and j belong to different $\overset{\sigma}{\sim}$-equivalence classes, i.e., to different orbits of σ. Let X_1 and X_2 be these two orbits, so that $i \in X_1$ and $j \in X_2$. Let $X_3, X_4, \ldots, X_k$ be all the remaining orbits of σ (listed without repetition). Thus, σ has k orbits. We shall now understand what the orbits of τ are.

Let $\sim$ be the relation $\overset{\sigma}{\sim}$. Proposition 4.3.4 **(b)** (applied to $a = i$) yields that there exists some positive integer m such that

$$\begin{aligned} \sigma^m(i) &= i \qquad \text{and} \\ [i]_\sim &= \left\{\sigma^0(i),\ \sigma^1(i),\ \ldots,\ \sigma^{m-1}(i)\right\} \qquad \text{and} \\ |[i]_\sim| &= m. \end{aligned}$$

Consider this m. Proposition 4.3.4 **(b)** (applied to $a = j$) yields that there exists some positive integer n (we cannot call it m now) such that

$$\begin{aligned} \sigma^n(j) &= j \qquad \text{and} \\ [j]_\sim &= \left\{\sigma^0(j),\ \sigma^1(j),\ \ldots,\ \sigma^{n-1}(j)\right\} \qquad \text{and} \\ |[j]_\sim| &= n. \end{aligned}$$

Consider this n.

Recall that X_1 is the orbit of σ that contains i. In other words, $X_1 = [i]_\sim$ (since $[i]_\sim$ is the orbit of σ that contains i). Therefore,

$$X_1 = [i]_\sim = \left\{\sigma^0(i),\ \sigma^1(i),\ \ldots,\ \sigma^{m-1}(i)\right\} = \left\{i,\ \sigma^1(i),\ \sigma^2(i),\ \ldots,\ \sigma^{m-1}(i)\right\}$$

(since $\sigma^0(i) = i$). The action of the map σ on the elements of X_1 looks as follows:

$$i \overset{\sigma}{\mapsto} \sigma^1(i) \overset{\sigma}{\mapsto} \sigma^2(i) \overset{\sigma}{\mapsto} \cdots \overset{\sigma}{\mapsto} \sigma^{m-1}(i) \overset{\sigma}{\mapsto} i \tag{17}$$

(since $\sigma\left(\sigma^{m-1}(i)\right) = \sigma^m(i) = i$). Similarly, the action of σ on the elements of X_2 looks as follows:

$$j \overset{\sigma}{\mapsto} \sigma^1(j) \overset{\sigma}{\mapsto} \sigma^2(j) \overset{\sigma}{\mapsto} \cdots \overset{\sigma}{\mapsto} \sigma^{n-1}(j) \overset{\sigma}{\mapsto} j. \tag{18}$$

However, the transposition $t_{i,j}$ swaps i with j while leaving all other elements unchanged. Thus, the permutation $\tau = t_{i,j} \circ \sigma$ differs from σ only in two things:

- The element of X that is sent to i by σ is instead sent to j by τ (because $t_{i,j}$ sends i to j).
- The element of X that is sent to j by σ is instead sent to i by τ (because $t_{i,j}$ sends j to i).

- All other elements of X are sent by τ to the same value that they are sent to by σ.

Hence, when we replace σ by τ, then the two arrows $\sigma^{m-1}(i) \overset{\sigma}{\mapsto} i$ and $\sigma^{n-1}(j) \overset{\sigma}{\mapsto} j$ in the circular chains (17) and (18) turn into $\sigma^{m-1}(i) \overset{\tau}{\mapsto} j$ and $\sigma^{n-1}(j) \overset{\tau}{\mapsto} i$, whereas all the other arrows remain unchanged. Hence, these two circular chains get merged into one single circular chain

$$i \overset{\tau}{\mapsto} \sigma^1(i) \overset{\tau}{\mapsto} \sigma^2(i) \overset{\tau}{\mapsto} \cdots \overset{\tau}{\mapsto} \sigma^{m-1}(i) \overset{\tau}{\mapsto} j \overset{\tau}{\mapsto} \sigma^1(j) \overset{\tau}{\mapsto} \sigma^2(j) \overset{\tau}{\mapsto} \cdots \overset{\tau}{\mapsto} \sigma^{n-1}(j) \overset{\tau}{\mapsto} i.$$

Hence, in place of the orbits X_1 and X_2 of σ, the permutation τ has a single orbit

$$X_1 \cup X_2 = \left\{ i,\ \sigma^1(i),\ \sigma^2(i),\ \ldots,\ \sigma^{m-1}(i),\ j,\ \sigma^1(j),\ \ldots,\ \sigma^{n-1}(j) \right\}.$$

All the remaining orbits $X_3, X_4, \ldots, X_k$ of σ remain orbits of τ (because the map τ sends them to the same values that the map σ sends them to[6]).

To summarize: we have shown that the orbits X_1 and X_2 of σ are merged into one single orbit $X_1 \cup X_2$ of τ (this is just saying that $X_1 \cup X_2$ is an orbit of τ), whereas all the remaining orbits $X_3, X_4, \ldots, X_k$ of σ remain orbits of τ. Therefore, τ has 1 fewer orbit than σ. This proves Lemma 4.3.13 **(c)**.

(d) We continue with the notations from our above proof of Lemma 4.3.13 **(c)**. Both i and j are contained in the orbit $X_1 \cup X_2$ of τ. Hence, i and j belong to the same orbit of τ. In other words, i and j belong to the same $\overset{\tau}{\sim}$-equivalence class. Hence, $i \overset{\tau}{\sim} j$. This proves Lemma 4.3.13 **(d)**. □

4.3.4. Stirling numbers of the first kind

Recall that S_n denotes the symmetric group $S_{[n]}$ (for any $n \in \mathbb{N}$).

Let us now count permutations with a given # of orbits (or, equivalently, cycles in their DCD):

Definition 4.3.15. Let $n \in \mathbb{N}$ and $k \in \mathbb{Z}$. Then, the **unsigned Stirling number of the first kind** $\begin{bmatrix} n \\ k \end{bmatrix}$ is defined to be the # of all permutations $\sigma \in S_n$ that have exactly k orbits (i.e., that have exactly k cycles in their DCD, provided that we **don't** omit the 1-element cycles).

The **signed Stirling number of the first kind** $s(n,k)$ is defined to be $(-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}$.

Example 4.3.16. We have $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 11$, because there are exactly 11 permutations $\sigma \in S_4$ that have exactly 2 orbits. The OLNs of these 11 permutations are

[6] This is proved similarly to the analogous claim in our above proof of Lemma 4.3.13 **(c)**.

(written without parentheses and without commas)

$$1342,\quad 1423,\quad 2143,\quad 2314,\quad 2431,\quad 3124,$$
$$3241,\quad 3412,\quad 4132,\quad 4213,\quad 4321.$$

Here are DCDs of two of them:

$$1342 = \mathrm{cyc}_{2,3,4} \circ \mathrm{cyc}_1; \qquad 2143 = \mathrm{cyc}_{1,2} \circ \mathrm{cyc}_{3,4}\,.$$

Eight of the 11 permutations we listed have a DCD of the form $\mathrm{cyc}_{a,b,c} \circ \mathrm{cyc}_d$ (and thus have an orbit of size 3 and an orbit of size 1); the other three have a DCD of the form $\mathrm{cyc}_{a,b} \circ \mathrm{cyc}_{c,d}$ (and thus have two orbits of size 2).

Convention 4.3.17. We will abbreviate "Stirling numbers of the first kind" as "**Stirling-1 numbers**", and "Stirling numbers of the second kind" as "**Stirling-2 numbers**".

Let us compute some "easy" values of $\begin{bmatrix} n \\ k \end{bmatrix}$:

Proposition 4.3.18. **(a)** We have $\begin{bmatrix} n \\ k \end{bmatrix} = 0$ for any $n, k \in \mathbb{N}$ satisfying $k > n$.

(b) We have $\begin{bmatrix} n \\ n \end{bmatrix} = 1$ for any $n \in \mathbb{N}$.

(c) We have $\begin{bmatrix} n \\ 1 \end{bmatrix} = (n-1)!$ for any $n > 0$.

(d) We have $\begin{bmatrix} n \\ 0 \end{bmatrix} = [n = 0]$ for any $n \in \mathbb{N}$.

(e) We have $\begin{bmatrix} 0 \\ k \end{bmatrix} = [k = 0]$ for any $k \in \mathbb{Z}$.

(f) We have $\begin{bmatrix} n \\ k \end{bmatrix} = 0$ for any $n \in \mathbb{N}$ and any negative $k \in \mathbb{Z}$.

Proof (sketched). Part **(e)** is obvious (there is only one permutation $\sigma \in S_0$, namely the identity map $\mathrm{id}_{[0]}$, and it has 0 orbits). So is part **(f)** (since a permutation cannot have a negative # of orbits). Thus, it remains to prove parts **(a)**, **(b)**, **(c)** and **(d)**. To that purpose, we fix $n \in \mathbb{N}$.

(a) If $\sigma \in S_n$ is any permutation, then the orbits of σ are disjoint nonempty subsets of $[n]$ (by Proposition 4.3.6), and thus there cannot be more than n of them (since $[n]$ has only n elements, and each of the orbits of σ must contain at least one of these n elements). In other words, a permutation $\sigma \in S_n$ cannot have k orbits for $k > n$. In other words, $\begin{bmatrix} n \\ k \end{bmatrix} = 0$ for any $k \in \mathbb{N}$ satisfying $k > n$. This proves Proposition 4.3.18 **(a)**.

(b) We start as in part **(a)**: If $\sigma \in S_n$ is any permutation, then the orbits of σ are disjoint nonempty subsets of $[n]$. Hence, if there are n of these orbits, then each of them must be a 1-element set (otherwise, their sizes would sum up to a number larger than n, which is impossible for disjoint subsets of $[n]$). However, this means that each element of $[n]$ is a fixed point of σ (by Remark 4.3.7), and therefore σ is the identity permutation $\mathrm{id}_{[n]}$.

Thus, we have shown that the only permutation $\sigma \in S_n$ that has n orbits is $\mathrm{id}_{[n]}$. Hence, $\begin{bmatrix} n \\ n \end{bmatrix} = 1$. This proves Proposition 4.3.18 **(b)**.

(c) Proposition 4.3.8 shows that if a permutation $\sigma \in S_n$ has only 1 orbit, then it is an n-cycle (because its one orbit must necessarily be the whole set $[n]$). Conversely, a permutation $\sigma \in S_n$ that is an n-cycle must have only 1 orbit (namely, the set $[n]$). Thus, we conclude that the permutations $\sigma \in S_n$ that have only 1 orbit are precisely the n-cycles in S_n. Hence, their # is the # of n-cycles in S_n. In other words,

$$\begin{bmatrix} n \\ 1 \end{bmatrix} = (\#\text{ of } n\text{-cycles in } S_n)\,.$$

However, Exercise 1 in Lecture 26 (applied to $X = [n]$ and $k = n$) yields that

$$\begin{aligned}
(\#\text{ of } n\text{-cycles in } S_n) &= \begin{cases} 1, & \text{if } n = 1; \\ (n-1)! \cdot \dbinom{n}{n}, & \text{if } n > 1 \end{cases} \\
&= \begin{cases} 1, & \text{if } n = 1; \\ (n-1)!, & \text{if } n > 1 \end{cases} \qquad \left(\text{since } \dbinom{n}{n} = 1\right) \\
&= (n-1)! \qquad (\text{since } 1 = (n-1)! \text{ when } n = 1)\,.
\end{aligned}$$

Thus,

$$\begin{bmatrix} n \\ 1 \end{bmatrix} = (\#\text{ of } n\text{-cycles in } S_n) = (n-1)!.$$

This proves Proposition 4.3.18 **(c)**.

(d) If $\sigma \in S_n$ is any permutation, then the orbits of σ must cover the entire set $[n]$ (by Proposition 4.3.6). Hence, σ cannot have 0 orbits, unless the set $[n]$ is empty, i.e., unless $n = 0$. Thus, $\begin{bmatrix} n \\ 0 \end{bmatrix} = 0$ for any $n > 0$. Combining this with the obvious fact that $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$, we conclude that Proposition 4.3.18 **(d)** holds. □

Let us now further explore the Stirling-1 numbers. Unlike the Stirling-2 numbers, they cannot be computed by a simple explicit formula, even allowing for a summation sign. But they satisfy the following recursion:

Proposition 4.3.19 (recurrence relation for the Stirling-1 numbers)**.** For any positive integer n and any integer k, we have

$$\begin{bmatrix} n \\ k \end{bmatrix} = (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}.$$

Proof sketch. Let n be a positive integer. Let k be any integer. We have

$$\begin{aligned} \begin{bmatrix} n \\ k \end{bmatrix} &= (\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits}) \\ &= \sum_{i=1}^{n} (\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i) \end{aligned}$$

(by the sum rule). Now, we claim the following:

Claim 1: We have

$$(\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = n) = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}.$$

Claim 2: For any $i \in [n-1]$, we have

$$(\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i) = \begin{bmatrix} n-1 \\ k \end{bmatrix}.$$

Once Claim 1 and Claim 2 are proved, we will conclude that

$$\begin{aligned} \begin{bmatrix} n \\ k \end{bmatrix} &= \sum_{i=1}^{n} (\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i) \\ &= \sum_{i=1}^{n-1} \underbrace{(\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i)}_{\substack{= \begin{bmatrix} n-1 \\ k \end{bmatrix} \\ \text{(by Claim 2)}}} \\ &\qquad + \underbrace{(\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = n)}_{\substack{= \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} \\ \text{(by Claim 1)}}} \\ &= \underbrace{\sum_{i=1}^{n-1} \begin{bmatrix} n-1 \\ k \end{bmatrix}}_{= (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix}} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} = (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}, \end{aligned}$$

and thus Proposition 4.3.19 will be proved. So it remains to verify Claim 1 and Claim 2.

Proof of Claim 1: If a permutation $\sigma \in S_n$ has k orbits and satisfies $\sigma(n) = n$, then n is a fixed point of σ, and thus the 1-element set $\{n\}$ is itself an orbit of σ (by Remark 4.3.7). Hence, by restricting such a permutation σ to the subset $[n-1]$, we obtain a permutation $\sigma\mid_{[n-1]}$ of $[n-1]$ that has $k-1$ orbits (note that this is well-defined, because $\sigma(n) = n$ ensures that $\sigma(i) \in [n-1]$ for each $i \in [n-1]$). This yields a bijection

$$\begin{aligned}&\text{from } \{\text{permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = n\}\\ &\text{to } \{\text{permutations } \tau \in S_{n-1} \text{ with } k-1 \text{ orbits}\},\end{aligned}$$

which sends each permutation σ to its restriction $\sigma\mid_{[n-1]}\colon [n-1] \to [n-1]$. (The inverse map of this bijection simply extends a permutation $\tau \in S_{n-1}$ to $[n]$ by setting $\tau(n) := n$. This extension creates a new orbit $\{n\}$, thus raising the # of orbits from $k-1$ to k.)

Thus, the bijection principle yields

$$\begin{aligned}&(\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = n)\\ &= (\#\text{ of permutations } \tau \in S_{n-1} \text{ with } k-1 \text{ orbits}) = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}\end{aligned}$$

(by the definition of $\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}$). This proves Claim 1.

Proof of Claim 2: Let $i \in [n-1]$. Thus, $i \neq n$. Consider the transposition $t_{i,n}$, which swaps i with n. It satisfies $t_{i,n} \circ t_{i,n} = \mathrm{id}$ (since swapping i with n twice puts all elements back where they started). Now, if $\sigma \in S_n$ is a permutation with k orbits and with $\sigma(n) = i$, then the permutation $t_{i,n} \circ \sigma$ has $k+1$ orbits (this follows easily from Lemma 4.3.13 **(a)**[7]) and satisfies $(t_{i,n} \circ \sigma)(n) = n$ [8]. Hence, we obtain a map

$$\begin{aligned}&\text{from } \{\text{permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i\}\\ &\text{to } \{\text{permutations } \sigma \in S_n \text{ with } k+1 \text{ orbits and with } \sigma(n) = n\},\end{aligned}$$

which sends each σ to $t_{i,n} \circ \sigma$. Conversely, if $\sigma \in S_n$ is a permutation with $k+1$ orbits and with $\sigma(n) = n$, then the permutation $t_{i,n} \circ \sigma$ has k orbits (this follows easily from

[7] Here is the proof in more detail: Let $\sigma \in S_n$ be a permutation with k orbits and with $\sigma(n) = i$. Set $\tau := t_{i,n} \circ \sigma$.

We have $i = \sigma(n) = \sigma^1(n)$, so that $i \overset{\sigma}{\sim} n$. Hence, Lemma 4.3.13 **(a)** (applied to $j = n$) yields that the permutation τ has 1 more orbit than σ. Thus, τ has $k+1$ orbits (since σ has k orbits). In other words, $t_{i,n} \circ \sigma$ has $k+1$ orbits (since $\tau = t_{i,n} \circ \sigma$). Qed.

[8] *Proof.* We have $(t_{i,n} \circ \sigma)(n) = t_{i,n}\left(\underbrace{\sigma(n)}_{=i}\right) = t_{i,n}(i) = n$ (by the definition of $t_{i,n}$).

Lemma 4.3.13 **(c)**[9]) and satisfies $(t_{i,n} \circ \sigma)(n) = i$ [10]. Hence, we obtain a map

$$\begin{aligned}&\text{from } \{\text{permutations } \sigma \in S_n \text{ with } k+1 \text{ orbits and with } \sigma(n) = n\} \\ &\text{to } \{\text{permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i\},\end{aligned}$$

which sends each σ to $t_{i,n} \circ \sigma$.

Thus, we have found two maps between the two sets

$$\begin{aligned}&\{\text{permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i\} \\ &\text{and } \{\text{permutations } \sigma \in S_n \text{ with } k+1 \text{ orbits and with } \sigma(n) = n\}\end{aligned}$$

(one map in either direction). These two maps are mutually inverse, since any permutation $\sigma \in S_n$ satisfies $t_{i,n} \circ (t_{i,n} \circ \sigma) = \underbrace{t_{i,n} \circ t_{i,n}}_{=\mathrm{id}} \circ \sigma = \mathrm{id} \circ \sigma = \sigma$. Hence, they are bijections. The bijection principle thus yields

$$\begin{aligned}&(\#\text{ of permutations } \sigma \in S_n \text{ with } k \text{ orbits and with } \sigma(n) = i) \\ &= (\#\text{ of permutations } \sigma \in S_n \text{ with } k+1 \text{ orbits and with } \sigma(n) = n) \\ &= \begin{bmatrix} n-1 \\ (k+1)-1 \end{bmatrix} \qquad (\text{by Claim 1, applied to } k+1 \text{ instead of } k) \\ &= \begin{bmatrix} n-1 \\ k \end{bmatrix}.\end{aligned}$$

This proves Claim 2.

Having proved both Claim 1 and Claim 2, we have now completed our proof of Proposition 4.3.19. □

[9] Here is the proof in more detail: Let $\sigma \in S_n$ be a permutation with $k+1$ orbits and with $\sigma(n) = n$. Set $\tau := t_{i,n} \circ \sigma$.

We have $\sigma(n) = n$, so that n is a fixed point of σ. Hence, Remark 4.3.7 (applied to $X = [n]$ and $a = n$) shows that the orbit of σ that contains n is a 1-element set. This orbit must therefore be $\{n\}$ (since n belongs to this orbit), and thus does not contain i (since $i \neq n$). Therefore, we don't have $i \overset{\sigma}{\sim} n$. Hence, Lemma 4.3.13 **(c)** (applied to $j = n$) yields that the permutation τ has 1 fewer orbit than σ. Thus, τ has k orbits (since σ has $k+1$ orbits). In other words, $t_{i,n} \circ \sigma$ has k orbits (since $\tau = t_{i,n} \circ \sigma$). Qed.

[10] *Proof.* We have $(t_{i,n} \circ \sigma)(n) = t_{i,n}\left(\underbrace{\sigma(n)}_{=n}\right) = t_{i,n}(n) = i$ (by the definition of $t_{i,n}$).

Using Proposition 4.3.19, it is easy to construct a table of Stirling-1 numbers:

$\begin{bmatrix} n \\ k \end{bmatrix}$	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
$k=0$	1	0	0	0	0	0	0	0
$k=1$	0	1	1	2	6	24	120	720
$k=2$	0	0	1	3	11	50	274	1764
$k=3$	0	0	0	1	6	35	225	1624
$k=4$	0	0	0	0	1	10	85	735
$k=5$	0	0	0	0	0	1	15	175
$k=6$	0	0	0	0	0	0	1	21
$k=7$	0	0	0	0	0	0	0	1

(see the Wikipedia page for more).

It is time to answer the question that you will surely have asked by now, namely: What do the Stirling-1 numbers have to do with the Stirling-2 numbers? Both types of numbers have been invented by James Stirling in 1730, but a mere coincidence like this does not usually lead to "first kind"/"second kind" nomenclature. There is a deeper connection between the two – a type of duality.

To best understand this duality, we need a little bit of linear algebra. We consider polynomials in one indeterminate X (with rational coefficients). For each $k \in \mathbb{N}$, we consider the **k-th falling factorial**

$$X^{\underline{k}} = X\left(X-1\right)\left(X-2\right)\cdots\left(X-k+1\right);$$

this is a degree-k polynomial in X. Evaluating this polynomial at a number n will, of course, yield the falling factorial $n^{\underline{k}}$ we introduced in Definition 2.4.2 (Lecture 16).

It is obvious that each polynomial in X can be written as a linear combination of the powers $X^0, X^1, X^2, \ldots$. It is a bit less obvious that each polynomial in X can be written as a linear combination of the falling factorials $X^{\underline{0}}, X^{\underline{1}}, X^{\underline{2}}, \ldots$. The easiest way to see this is by writing each power X^n as such a combination. This can be done explicitly, and the coefficients are the Stirling-2 numbers:

Theorem 4.3.20. Let $n \in \mathbb{N}$. Then,

$$X^n = \sum_{k=0}^{n} \left\{ n \atop k \right\} X^{\underline{k}}. \tag{19}$$

Proof sketch. By the "polynomial identity trick" (Corollary 2.2.5 in Lecture 14), it suffices to show that

$$x^n = \sum_{k=0}^{n} \left\{ n \atop k \right\} x^{\underline{k}} \qquad \text{for each } x \in \mathbb{N}.$$

So let us do this. Fix $x \in \mathbb{N}$. Then, Theorem 2.5.1 in Lecture 17 (applied to $k = x$ and $m = n$) yields

$$x^n = \sum_{i=0}^{n} \operatorname{sur}(n,i) \cdot \binom{x}{i} = \sum_{k=0}^{n} \operatorname{sur}(n,k) \cdot \underbrace{\binom{x}{k}}_{= \frac{x(x-1)(x-2)\cdots(x-k+1)}{k!} = \frac{x^{\underline{k}}}{k!} \text{ (since } x(x-1)(x-2)\cdots(x-k+1) = x^{\underline{k}})}$$

$$= \sum_{k=0}^{n} \underbrace{\operatorname{sur}(n,k) \cdot \frac{x^{\underline{k}}}{k!}}_{= \frac{\operatorname{sur}(n,k)}{k!} \cdot x^{\underline{k}}} = \sum_{k=0}^{n} \underbrace{\frac{\operatorname{sur}(n,k)}{k!}}_{= \left\{ n \atop k \right\} \text{ (by the definition of the Stirling-2 numbers)}} \cdot x^{\underline{k}} = \sum_{k=0}^{n} \left\{ n \atop k \right\} x^{\underline{k}},$$

qed. □

So Theorem 4.3.20 expands any power X^n as a linear combination of falling factorials $X^{\underline{k}}$. The coefficients are the Stirling-2 numbers $\left\{ n \atop k \right\}$.

What about the converse direction – i.e., how can we expand a falling factorial $X^{\underline{n}}$ as a linear combination of the powers X^k ? In other words, what are the coefficients of $X^{\underline{n}} = X(X-1)(X-2)\cdots(X-n+1)$?

Here, as it turns out, the coefficients will be the signed Stirling-1 numbers:

Theorem 4.3.21. Let $n \in \mathbb{N}$. Then,

$$\begin{aligned} X^{\underline{n}} &= X(X-1)(X-2)\cdots(X-n+1) \\ &= \sum_{k=0}^{n} s(n,k) X^k \\ &= \sum_{k=0}^{n} (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} X^k. \end{aligned} \tag{20}$$

Equivalently (by substituting $-X$ for X), this becomes

$$X(X+1)(X+2)\cdots(X+n-1) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} X^k.$$

Proof sketch. There is a nice combinatorial proof (see, e.g., [Galvin17, proof of (34)]). Three other proofs can be found in [Stanle11, Proposition 1.3.7] (in which $\begin{bmatrix} n \\ k \end{bmatrix}$ is denoted by $c(n,k)$).

The easiest proof at this point is just a straightforward induction on n, similar to our proof of the binomial formula (Theorem 1.3.19 in Lecture 7). Use Proposition 4.3.19 in the induction step. □

A curious consequence of Theorems 4.3.21 and 4.3.20 is the following relation between the two kinds of Stirling numbers:

Theorem 4.3.22. For any $i, j \in \mathbb{N}$, we have

$$\sum_{k=0}^{i} s(i,k) \left\{ {k \atop j} \right\} = \sum_{k=0}^{i} \left\{ {i \atop k} \right\} s(k,j) = [i = j].$$

Proof sketch. (Some familiarity with vector spaces is required here. See [Loehr17, Theorem 2.64] for the details of this argument. See also `https://math.stackexchange.com/questions/42018/stirling-numbers-and-inverse-matrices` for a generalization.)

Fix $n \in \mathbb{N}$. Consider the $\mathbb{Q}$-vector space P_n of all polynomials of degree $\leq n-1$ (including the zero polynomial) in one indeterminate X (with rational coefficients). This vector space P_n has two important bases:

- the **power basis** $\left(X^0, X^1, \ldots, X^{n-1}\right)$, which consists of the first n powers of X;
- the **falling factorial basis** $\left(X^{\underline{0}}, X^{\underline{1}}, \ldots, X^{\underline{n-1}}\right)$.

(The power basis is clearly a basis. The falling factorial basis can easily be seen to be a basis, since it expands triangularily in the power basis – i.e., each falling factorial $X^{\underline{k}}$ can be written as X^k plus a linear combination of lower powers $X^{k-1}, X^{k-2}, \ldots, X^0$.)

For each $i \in \{0, 1, \ldots, n-1\}$, we have

$$\begin{aligned} X^i &= \sum_{k=0}^{i} \left\{ {i \atop k} \right\} X^{\underline{k}} \qquad \text{(by (19), applied to } i \text{ instead of } n\text{)} \\ &= \sum_{k=0}^{n-1} \left\{ {i \atop k} \right\} X^{\underline{k}} \end{aligned}$$

(here, we have extended the range of the summation from $k \in \{0, 1, \ldots, i\}$ to $k \in \{0, 1, \ldots, n-1\}$, which does not change the value because $\left\{ {i \atop k} \right\} = 0$ for all $k > i$). This shows that the $n \times n$-matrix[11]

$$S_2 := \left(\left\{ {i \atop j} \right\} \right)_{0 \leq i,j \leq n-1}$$

[11] This is an $n \times n$-matrix, but we number its rows and its columns by $0, 1, \ldots, n-1$ rather than by $1, 2, \ldots, n$.

is the transition matrix (= change-of-basis matrix) from the falling factorial basis to the power basis. Likewise, using the formula (20), we can see that the $n \times n$-matrix

$$S_1 := \left(s\left(i,j\right)\right)_{0 \leq i,j \leq n-1}$$

is the transition matrix from the power basis to the falling factorial basis. However, it is well-known that the transition matrices between two bases of a vector space are always mutually inverse (see, e.g., [LaNaSc16, §10.2, $RS = I$]). Thus, we conclude that the two matrices

$$S_1 = \left(s\left(i,j\right)\right)_{0 \leq i,j \leq n-1} \qquad \text{and} \qquad S_2 = \left(\left\{ {i \atop j} \right\}\right)_{0 \leq i,j \leq n-1}$$

are mutually inverse. In other words, they satisfy $S_1 S_2 = I_n$ and $S_2 S_1 = I_n$ (where I_n denotes the $n \times n$-identity matrix). Explicitly, this means that for any $i, j \in \{0, 1, \ldots, n-1\}$, the (i,j)-th entries of both products $S_1 S_2$ and $S_2 S_1$ equal the (i,j)-th entry of I_n, which of course is $[i = j]$. In other words, for any $i, j \in \{0, 1, \ldots, n-1\}$, we have the equality

$$\sum_{k=0}^{n-1} s\left(i,k\right) \left\{ {k \atop j} \right\} = \sum_{k=0}^{n-1} \left\{ {i \atop k} \right\} s\left(k,j\right) = \left[i = j\right].$$

Since both numbers $s\left(i,k\right)$ and $\left\{ {i \atop k} \right\}$ are 0 for $k > i$, we can replace the upper limits of both summation signs by i, thus rewriting this equality as follows:

$$\sum_{k=0}^{i} s\left(i,k\right) \left\{ {k \atop j} \right\} = \sum_{k=0}^{i} \left\{ {i \atop k} \right\} s\left(k,j\right) = \left[i = j\right].$$

Since we can choose n arbitrarily high, we thus conclude that this equality holds for all $i, j \in \mathbb{N}$. This proves Theorem 4.3.22.

Alternatively, Theorem 4.3.22 can also be proved by induction. See `https://proofwiki.org/wiki/First_Inversion_Formula_for_Stirling_Numbers` and `https://proofwiki.org/wiki/Second_Inversion_Formula_for_Stirling_Numbers` for such a proof. See also [BenQui03, Identity 194] for a combinatorial proof of $\sum_{k=0}^{i} s\left(i,k\right) \left\{ {k \atop j} \right\} = \left[i = j\right]$. □

References

[17f-hw7s] Darij Grinberg, *UMN Fall 2017 Math 4990 homework set #7 with solutions*, `http://www.cip.ifi.lmu.de/~grinberg/t/17f/hw7os.pdf`

[21s] Darij Grinberg, *Algebraic Combinatorics (Drexel Spring 2021 Math 701 lecture notes)*, 4 May 2021. `https://www.cip.ifi.lmu.de/~grinberg/t/21s/lecs.pdf`

[BenQui03] Arthur T. Benjamin and Jennifer J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, Dolciani Mathematical Expositions **27**, The Mathematical Association of America, 2003.

[Bourba74] Nicolas Bourbaki, *Algebra I: Chapters 1–3*, Addison-Wesley 1974.

[Galvin17] David Galvin, *Basic discrete mathematics*, 13 December 2017.
`http://www-users.math.umn.edu/~dgrinber/comb/60610lectures2017-Galvin.pdf`
(The URL might change, and the text may get updated. In order to reliably obtain the version of 13 December 2017, you can use the archive.org Wayback Machine: `https://web.archive.org/web/20180205122609/http://www-users.math.umn.edu/~dgrinber/comb/60610lectures2017-Galvin.pdf` .)

[Goodma15] Frederick M. Goodman, *Algebra: Abstract and Concrete*, edition 2.6, 1 May 2015.
`http://homepage.math.uiowa.edu/~goodman/algebrabook.dir/book.2.6.pdf` .

[LaNaSc16] Isaiah Lankham, Bruno Nachtergaele, Anne Schilling, *Linear Algebra As an Introduction to Abstract Mathematics*, 2016.
`https://www.math.ucdavis.edu/~anne/linear_algebra/mat67_course_notes.pdf`

[Loehr17] Nicholas A. Loehr, *Combinatorics*, 2nd edition, CRC Press 2017.

[Stanle11] Richard P. Stanley, *Enumerative Combinatorics, volume 1*, Second edition, version of 15 July 2011. Available at `http://math.mit.edu/~rstan/ec/` .
See `http://math.mit.edu/~rstan/ec/` for errata.