1 Computations in the symmetric group

Recall that, given a set $X$, the set $S_X$ of all bijections from $X$ to itself (or, more briefly, permutations of $X$) is group under function composition. In particular, for each $n \in \mathbb{N}$, the symmetric group $S_n$ is the group of permutations of the set $\{1, \ldots, n\}$, with the group operation equal to function composition. Thus $S_n$ is a group with $n!$ elements, and it is not abelian if $n \geq 3$. If $X$ is a finite set with $\#(X) = n$, then any labeling of the elements of $X$ as $x_1, \ldots, x_n$ defines an isomorphism from $S_X$ to $S_n$ (see the homework for a more precise statement). We will write elements of $S_n$ using Greek letters $\sigma, \tau, \rho, \ldots$, we denote the identity function by 1, and we just use the multiplication symbol $\cdot$ or juxtaposition to denote composition (instead of the symbol $\circ$). Recall however that functions work from right to left: thus $\sigma \tau(i) = \sigma(\tau(i))$, in other words evaluate first $\tau$ at $i$, and then evaluate $\sigma$ on this. We say that $\sigma$ moves $i$ if $\sigma(i) \neq i$, and that $\sigma$ fixes $i$ if $\sigma(i) = i$.

There are many interesting subgroups of $S_n$. For example, the subset $H_n$ defined by

$$H_n = \{ \sigma \in S_n : \sigma(n) = n \}$$

is easily checked to be a subgroup of $S_n$ isomorphic to $S_{n-1}$ (see the homework for a generalization of this). If $n = n_1 + n_2$ for two positive integers $n_1, n_2$ then the subset

$$H = \{ \sigma \in S_n : \sigma(\{1, \ldots, n_1\}) = \{1, \ldots, n_1\} \}$$

is also a subgroup of $S_n$. Note that, if $\sigma \in H$, then automatically

$$\sigma(\{n_1 + 1, \ldots, n_2\}) = \{n_1 + 1, \ldots, n_2\},$$

and in fact it is easy to check that $H$ is isomorphic to $S_{n_1} \times S_{n_2}$. There are many other subgroups of $S_n$. For example, the dihedral group $D_n$, the group of symmetries of a regular $n$-gon in the plane, is (isomorphic to) a subgroup of $S_n$ by looking at the permutation of the vertices of the $n$-gon.
Note that \( #(D_n) = 2^n \), and hence that \( D_n \) is a proper subgroup of \( S_n \) if \( n \geq 4 \). We shall see (Cayley’s theorem) that, if \( G \) is a finite group, then there exists an \( n \) such that \( G \) is isomorphic to a subgroup of \( S_n \) (and in fact one can take \( n = #(G) \)). Thus the groups \( S_n \) are as complicated as all possible finite groups.

To describe a function \( \sigma : \{1, \ldots, n\} \to \{1, \ldots, n\} \), not necessarily a permutation, we can give a table of its values, recording \( i \) and then \( \sigma(i) \), as follows:

<table>
<thead>
<tr>
<th>( i )</th>
<th>( \sigma(i) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \sigma(1) )</td>
</tr>
<tr>
<td>2</td>
<td>( \sigma(2) )</td>
</tr>
<tr>
<td>\ldots</td>
<td>|</td>
</tr>
<tr>
<td>( n )</td>
<td>( \sigma(n) )</td>
</tr>
</tbody>
</table>

Of course, we could describe the same information by a \( 2 \times n \) matrix:

\[
\begin{pmatrix}
1 & 2 & \ldots & n \\
\sigma(1) & \sigma(2) & \ldots & \sigma(n)
\end{pmatrix}
\]

The condition that \( \sigma \) is a permutation is then the statement that the integers \( 1, \ldots, n \) each occur \textbf{exactly once} in the second row of the matrix. For example, if \( \sigma \) is given by the matrix

\[
\begin{pmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
2 & 4 & 8 & 1 & 5 & 7 & 3 & 6
\end{pmatrix}
\]

then \( \sigma(1) = 2 \) and \( \sigma(5) = 5 \).

This notation is however cumbersome and not well suited to calculation. We describe a much more efficient way to write down elements of \( S_n \) by first writing down some special ones, called \textit{cycles}, and then showing that every element of \( S_n \) can be factored into a product of cycles, in an essentially unique way if we are careful.

**Definition 1.1.** Let \( \{a_1, \ldots, a_k\} \) be a subset of \( \{1, \ldots, n\} \) with exactly \( k \) elements; equivalently, \( a_1, \ldots, a_k \) are distinct. We denote by \( (a_1, \ldots, a_k) \) the following element of \( S_n \):

\[
(a_1, \ldots, a_k)(a_i) = a_{i+1}, \quad \text{if } i < k;
\]

\[
(a_1, \ldots, a_k)(a_k) = a_1;
\]

\[
(a_1, \ldots, a_k)(j) = j, \quad \text{if } j \neq a_i \text{ for any } i.
\]

We call \( (a_1, \ldots, a_k) \) a \textit{k-cycle} or \textit{cycle of length k}. For \( k > 1 \), we define the support \( \text{Supp}(a_1, \ldots, a_k) \) to be the set \( \{a_1, \ldots, a_k\} \). Note that, again for \( k > 1 \), the \( k \)-cycle \( (a_1, \ldots, a_k) \) moves \( i \iff i \in \text{Supp}(a_1, \ldots, a_k) \). Two cycles \( (a_1, \ldots, a_k) \) and \( (b_1, \ldots, b_\ell) \) are \textit{disjoint} if

\[
\text{Supp}(a_1, \ldots, a_k) \cap \text{Supp}(b_1, \ldots, b_\ell) = \emptyset,
\]

i.e. the sets \( \{a_1, \ldots, a_k\} \) and \( \{b_1, \ldots, b_\ell\} \) are disjoint.
Remark 1.2. 1) A 1-cycle \((a_1)\) is the identity function 1, no matter what \(a_1\) is. For this reason, we will generally only consider cycles of length at least 2. If \(\sigma = (a_1, \ldots, a_k)\) with \(k \geq 2\), then \(\sigma\) is never the identity, since \(\sigma(a_1) = a_2 \neq a_1\).

2) A 2-cycle \((a_1, a_2)\) is also called a transposition. It is the unique permutation of \(\{1, \ldots, n\}\) which switches \(a_1\) and \(a_2\) and leaves all other elements alone.

3) From the description in 2), it is clear that \((a_1, a_2) = (a_2, a_1)\). But for \(k \geq 3\), the order of the elements \(a_1, \ldots, a_k\) is important: for example \(\sigma_1 = (1, 3, 2) \neq \sigma_2 = (1, 2, 3)\), because \(\sigma_1(1) = 3\) but \(\sigma_2(1) = 2\).

4) However, there are a few ways to change the order without changing the element of \(S_n\): clearly \((a_1, a_2, \ldots, a_k) = (a_2, a_3, \ldots, a_k, a_1) = (a_3, a_4, \ldots, a_k, a_1, a_2) = \cdots = (a_k, a_1, \ldots, a_{k-2}, a_{k-1})\).

In other words, you can start the cycle anywhere, at \(a_i\), say, but then you have to list the elements in order: the next one must be \(a_{i+1}\), and so on, with the understanding that once you reach \(a_k\), the next one has to be \(a_1\), then \(a_2\), and then so on up to \(a_{i-1}\). Clearly, this the only way you can change the order. By convention, we often start with the smallest \(a_i\). Of course, after that, there is no constraint on the sizes of the consecutive members of the cycle.

5) It is easy to see that the inverse \((a_1, a_2, \ldots, a_k)^{-1} = (a_k, a_{k-1}, \ldots, a_1)\). In other words, the inverse of the \(k\)-cycle \((a_1, a_2, \ldots, a_k)\) is the \(k\)-cycle where the elements are written in the opposite order. In particular, for a transposition \((a_1, a_2)\), \((a_1, a_2)^{-1} = (a_2, a_1) = (a_1, a_2)\), i.e. a transposition has order 2.

6) Generalizing the last line of 5), it is easy to see that the order of a \(k\)-cycle \(\sigma = (a_1, a_2, \ldots, a_k)\) is exactly \(k\). In fact, \(\sigma(a_i) = a_{i+1}\), so that \(\sigma^r(a_i) = a_{i+r}\), with the understanding that the addition \(i + r\) is to be taken mod \(k\), but using the representatives 1, \ldots, \(k\) for addition mod \(k\) instead of the more usual ones (in this course) of 0, \ldots, \(k - 1\). In particular, we see that \(\sigma^r(a_i) = a_i\) for all \(i \iff r\) is a multiple of \(k\), and since \(\sigma^r(j) = j\) for \(j \neq a_i\), we see that \(k\) is the smallest positive integer \(r\) such that \(\sigma^r = 1\).

Note however that, if \(\sigma\) is a \(k\)-cycle, its powers \(\sigma^r\) need not be \(k\)-cycles. For example,

\[
(1, 2, 3, 4)^2 = (1, 3)(2, 4),
\]

and \((1, 3)(2, 4)\) is not a \(k\)-cycle for any \(k\).
7) Suppose that \( \sigma_1 = (a_1, \ldots, a_k) \) and \( \sigma_2 = (b_1, \ldots, b_\ell) \) are two disjoint cycles. Then it is easy to see that \( \sigma_1 \sigma_2 = \sigma_2 \sigma_1 \), i.e. “disjoint cycles commute.”

To check this we have to check that, for all \( j \in \{1, \ldots, n\} \), \( \sigma_1 \sigma_2(j) = \sigma_2 \sigma_1(j) \). First, if \( j = a_i \) for some \( i \), then \( \sigma_1(a_i) = a_{i+1} \) (with our usual conventions on adding mod \( k \)) but \( \sigma_2(a_i) = a_i \) since \( a_i \neq b_r \) for any \( r \). For the same reason, \( \sigma_2(a_{i+1}) = a_{i+1} \). Thus, for all \( i \) with \( 1 \leq i \leq k \),

\[
\sigma_1 \sigma_2(a_i) = \sigma_1(\sigma_2(a_i)) = \sigma_1(a_i) = a_{i+1},
\]

whereas

\[
\sigma_2 \sigma_1(a_i) = \sigma_2(\sigma_1(a_i)) = \sigma_2(a_{i+1}) = a_{i+1} = \sigma_1 \sigma_2(a_i).
\]

Similarly, \( \sigma_1 \sigma_2(b_r) = b_{r+1} = \sigma_2 \sigma_1(b_r) \) for all \( r, 1 \leq r \leq \ell \). Finally, if \( j \) is not an \( a_i \) or a \( b_r \) for any \( i, r \), then

\[
\sigma_1 \sigma_2(j) = \sigma_1(\sigma_2(j)) = \sigma_1(j) = j,
\]

and similarly \( \sigma_2 \sigma_1(j) = j \). Thus, for all possible \( j \in \{1, \ldots, n\} \), \( \sigma_1 \sigma_2(j) = \sigma_2 \sigma_1(j) \), hence \( \sigma_1 \sigma_2 = \sigma_2 \sigma_1 \).

Note that non-disjoint cycles might or might not commute. For example,

\[
(1, 2)(1, 3) = (1, 3, 2) \neq (1, 2, 3) = (1, 3)(1, 2),
\]

whereas

\[
(1, 2, 3, 4, 5)(1, 3, 5, 2, 4) = (1, 4, 2, 5, 3) = (1, 3, 5, 2, 4)(1, 2, 3, 4, 5).
\]

8) Finally, we have the beautiful formula: if \( \sigma \in S_n \) is an arbitrary element (not necessarily a \( k \)-cycle), and \( (a_1, \ldots, a_k) \) is a \( k \)-cycle, then

\[
\sigma \cdot (a_1, \ldots, a_k) \cdot \sigma^{-1} = (\sigma(a_1), \ldots, \sigma(a_k)).
\]

In other words, \( \sigma \cdot (a_1, \ldots, a_k) \cdot \sigma^{-1} \) is again a \( k \)-cycle, but it is the \( k \)-cycle where the elements \( a_1, \ldots, a_k \) have been “renamed” by \( \sigma \).

To prove this formula, it suffices to check that

\[
\sigma \cdot (a_1, \ldots, a_k) \cdot \sigma^{-1}(j) = (\sigma(a_1), \ldots, \sigma(a_k))(j)
\]

for every \( j \in \{1, \ldots, n\} \). First, if \( j \) is of the form \( \sigma(a_i) \) for some \( i \), then by definition \( (\sigma(a_1), \ldots, \sigma(a_k))(\sigma(a_i)) = \sigma(a_{i+1}) \), with the usual remark that if \( i = k \) then we interpret \( k + 1 \) as 1. On the other hand,

\[
\sigma \cdot (a_1, \ldots, a_k) \cdot \sigma^{-1}(\sigma(a_i)) = \sigma \cdot (a_1, \ldots, a_k)(\sigma^{-1}(\sigma(a_i)))
\]

\[
= \sigma \cdot (a_1, \ldots, a_k)(a_i) = \sigma((a_1, \ldots, a_k)(a_i)) = \sigma(a_{i+1}).
\]
Thus both sides agree if $j = \sigma(a_i)$ for some $i$. On the other hand, if $j \neq \sigma(a_i)$ for any $i$, then $(\sigma(a_1), \ldots, \sigma(a_k))(j) = j$. But $j \neq \sigma(a_i) \iff \sigma^{-1}(j) \neq a_i$, so

$$
\sigma \cdot (a_1, \ldots, a_k) \cdot \sigma^{-1}(j) = \sigma \cdot (a_1, \ldots, a_k)(\sigma^{-1}(j))
$$

$$
= \sigma((a_1, \ldots, a_k)(\sigma^{-1}(j))) = \sigma(\sigma^{-1}(j)) = j.
$$

Hence $\sigma \cdot (a_1, \ldots, a_k) \cdot \sigma^{-1}(j) = (\sigma(a_1), \ldots, \sigma(a_k))(j)$ for every $j \in \{1, \ldots, n\}$, proving the formula.

In the discussion above, we have seen numerous computations with cycles, and in particular examples where the product of two cycles either is or is not again a cycle. More generally, we have the following factorization result for elements of $S_n$:

**Theorem 1.3.** Let $\sigma \in S_n$, $\sigma \neq 1$. Then $\sigma$ is a product of disjoint cycles of lengths at least 2. Moreover, the terms in the product are unique up to order.

**Remark 1.4.** 1) Just as with factoring natural numbers into primes, a single cycle is a “product” of cycles (it is a product of one cycle).

2) We could also allow 1 in this framework, with the convention that 1 is the empty product, i.e. the “product” of no cycles.

3) Since disjoint cycles commute, we can always change the order in a product of disjoint cycles and the answer will be the same. However, the theorem says that is the only ambiguity possible.

**Example 1.5.** For the group $S_4$, the only possibilities for products of disjoint cycles are: 1) the identity 1; 2) a transposition, i.e. a 2-cycle; 3) a 3-cycle; 4) a 4-cycle; 5) a product of two disjoint 2-cycles. There is just one identity 1. The number of transpositions is $\binom{4}{2} = 6$. For a 3-cycle $(a_1, a_2, a_3)$, there are 4 choice for $a_1$, then 3 choices for $a_2$, then 2 choices for $a_3$, giving a total of $4 \cdot 3 \cdot 2 = 24$ choices for the ordered triple $(a_1, a_2, a_3)$. However, as an element of $S_4$, $(a_1, a_2, a_3) = (a_2, a_3, a_1) = (a_3, a_1, a_2)$, so the total number of different elements of $S_4$ which are 3-cycles is $24/3 = 8$. Likewise, the total number of 4-cycles, viewed as elements of $S_4$, is $4 \cdot 3 \cdot 2 \cdot 1 / 4 = 6$. Finally, to count the number of products $(a, b)(c, d)$ of two disjoint 2-cycles, note that, as above, there are 6 choices for $(a, b)$. The choice of $(a, b)$ then determines $c$ and $d$ since $\{c, d\} = \{1, 2, 3, 4\} \setminus \{a, b\}$. But since $(a, b)(c, d) = (c, d)(a, b)$, we should divide the total number by 2, to get $6/2 = 3$ elements.
of $S_4$ which can be written as a product of two disjoint 2-cycles. As a check, adding up the various possibilities gives $1 + 6 + 8 + 6 + 3 = 24$, as expected.

In this notation, $D_4$ is the subgroup of $S_4$ given by

$$D_4 = \{1, (1,2,3,4), (1,3)(2,4), (1,4,3,2), (2,4), (1,3), (1,2)(3,4), (1,4)(2,3)\}.$$ 

The proof of the theorem gives a procedure which is far easier to understand and implement in practice than it is to explain in the abstract. For example, suppose we are given a concrete permutation corresponding to a $2 \times n$ matrix, for example

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 6 & 5 & 7 & 1 & 4 & 9 & 8 & 2 \end{pmatrix},$$

here is how to write $\sigma$ as a product of disjoint cycles of lengths at least 2: beginning with 1, we see that $\sigma(1) = 3$, $\sigma(3) = 5$, $\sigma(5) = 1$, so we have returned to where we started. Write down the cycle $(1,3,5)$. Searching for elements not in the support of $(1,3,5)$, we see that the first such is 2. Then $\sigma(2) = 6$, $\sigma(6) = 4$, $\sigma(4) = 7$, $\sigma(7) = 9$, $\sigma(9) = 2$ so that we are again where we started. Write down the cycle $(2,6,4,7,9)$. Search again for elements not in the support of $(1,3,5)$ or $(2,6,4,7,9)$. The only remaining element is 8, and $\sigma(8) = 8$. As we omit 1-cycles, we are left with the factorization

$$\sigma = (1,3,5)(2,6,4,7,9) = (2,6,4,7,9)(1,3,5).$$

To describe this procedure in general, we introduce a new idea.

**Definition 1.6.** Let $\sigma \in S_n$. We define a relation $\sim_\sigma$ on $\{1, \ldots, n\}$ as follows: for $i, j \in \{1, \ldots, n\}$, $i \sim_\sigma j$ if there exists an $r \in \mathbb{Z}$ such that $\sigma^r(i) = j$. The orbit of $i$ under $\sigma$ is the set

$$O_\sigma(i) = \{\sigma^r(i) : r \in \mathbb{Z}\}.$$

Before giving examples, we note the following:

**Proposition 1.7.** The relation $\sim_\sigma$ is an equivalence relation and the equivalence class $[i]$ of $i$ for $\sim_\sigma$ is the orbit $O_\sigma(i)$.

**Proof.** Clearly $i = \sigma^0(i)$, so that $i \sim_\sigma i$. Hence $\sim_\sigma$ is reflexive. If $i \sim_\sigma j$, then by definition there exists an $r \in \mathbb{Z}$ such that $\sigma^r(i) = j$. Thus $i = \sigma^{-r}(j)$, so that $j \sim_\sigma i$ and $\sim_\sigma$ is symmetric. To see that it is transitive, suppose that $i \sim_\sigma j$ and $j \sim_\sigma k$. Thus there exist $r, s \in \mathbb{Z}$ such that $\sigma^r(i) = j$ and $\sigma^s(j) = k$. Then $\sigma^{r+s}(i) = \sigma^s(\sigma^r(i)) = \sigma^s(j) = k$. Thus $i \sim_\sigma k$ and so $\sim_\sigma$ is transitive, hence an equivalence relation. By definition, the equivalence class $[i]$ is $O_\sigma(i)$.

6
By general facts about equivalence relations, the orbits $O_\sigma(i)$ partition \{1, \ldots, n\} into disjoint subsets, in the sense that every integer $j \in \{1, \ldots, n\}$ is in exactly one orbit $O_\sigma(i)$. Note that $O_\sigma(i) = \{i\} \iff \sigma$ fixes $i$. For example, the identity permutation has the $n$ orbits $\{1\}, \{2\}, \ldots, \{n\}$. A $k$-cycle $(a_1, \ldots, a_k)$ with $k \geq 2$ has one orbit $\{a_1, \ldots, a_k\}$ with $k$ elements, and the remaining orbits are one element orbits $\{i\}$ for the $i \in \{1, \ldots, n\}$ such that $i \neq a_r$ for any $r$. There are then $n-k$ one element orbits and one orbit with $k$ elements, for a total of $n-k+1$ orbits. For another example, given the $\sigma \in S_9$ described above, $\sigma = (1,3,5)(2,6,4,7,9)$, the orbits of $\sigma$ are \{1,3,5\}, \{2,4,6,7,9\}, and \{8\}. More generally, if $\sigma$ is a product of disjoint cycles $\rho_1 \cdots \rho_t$ of lengths at least 2, then the orbits $O_\sigma(i)$ are the supports of the cycles $\rho_i$ in addition to one element orbits for each $i \in \{1, \ldots, n\}$ which is not in the support of any $\rho_i$.

Note that, while $\sigma$ determines the orbits, the orbits do not completely determine $\sigma$. For example, $\sigma' = (1,5,3)(2,9,7,6,4)$ has the same set of orbits as $\sigma$. On the other hand, the orbits do determine the “shape” of $\sigma$, in other words they tell us in this case that $\sigma$ is a product of a disjoint 3-cycle and a 5-cycle, and they tell us that the support of the 3-cycle is \{1,3,5\} and the support of the 5-cycle is \{2,4,6,7,9\}.

Let us make two remarks about the equivalence relation $\sim_\sigma$ above. First, since $S_n$ is finite, every element $\sigma$ has finite order, say $\sigma^d = 1$ with $d > 0$. Then $\sigma^{d-1} = \sigma^{-1}$, so we can write every negative power of $\sigma$ as a positive power as well. Thus

$$O_\sigma(i) = \{\sigma^r(i) : r \in \mathbb{N}\}.$$ 

Second, given $i \in \{1, \ldots, n\}$, the orbit of $i$ is the set $\{i, \sigma(i), \sigma^2(i), \ldots\}$ by the above discussion. Inductively, set $a_1 = i$ and define $a_{k+1} = \sigma(a_k)$, so that $a_k = \sigma^{k-1}(i)$. Thus the orbit $O_\sigma(i)$ is just $\{a_1, a_2, \ldots\}$, with $\sigma(a_k) = a_{k+1}$. Note that the set $\{a_1, a_2, \ldots\}$ is finite, so there have to exist $r, s \geq 1$ with $r \neq s$ and $a_r = a_s$. Let $k \in \mathbb{N}$ be the smallest integer $\geq 1$ such that $a_{k+1} = a_\ell$ for some $\ell$ with $1 \leq \ell \leq k$, i.e. $a_{k+1}$ is the first term in the sequence which is equal to a previous term. Equivalently, $k$ is the largest positive integer such that $a_1, \ldots, a_k$ are all distinct. We claim that $a_{k+1} = a_1 = i$, in other words that the sequence starts to repeat only when we come back to the starting point $a_1 = i$. For suppose that $a_{k+1} = a_\ell$ with $1 \leq \ell \leq k$. Then $\sigma^k(i) = \sigma^{\ell-1}(i)$, so that $\sigma^{k-\ell+1}(i) = i$. Hence $a_{k-\ell+2} = a_1$, in other words there is a repetition at stage $k-\ell+2$. Since $k+1$ was the smallest positive integer greater than 1 for which some repetition occurs, $k-\ell+2 \geq k+1$, so that $\ell \leq 1$. But also $\ell \geq 1$, so that $\ell = 1$ and $a_{k+1} = a_1$, as claimed. Once we have $a_{k+1} = a_1$, i.e. $\sigma^k(i) = i$, then $a_{k+2} = \sigma(i) = a_2$, $a_{k+3} = a_3$, \ldots. In
general, given \( r \in \mathbb{Z} \), writing \( r = kq + \ell \) with \( 0 \leq \ell \leq k - 1 \), it follows that

\[
\sigma^\ell(i) = \sigma^\ell(\sigma^{kq}(i)) = \sigma^\ell(\sigma^{k}(\sigma^{q}(i))) = \sigma^\ell(i) = a_{\ell+1},
\]

where \( 1 \leq \ell + 1 \leq k \), and hence that

\[
O_\sigma(i) = \{i, \sigma(i), \sigma^2(i), \ldots, \sigma^{k-1}(i)\} = \{a_1, \ldots, a_k\}.
\]

Using the above, let us show that an arbitrary \( \sigma \in S_n \) can be factored into a product of disjoint cycles as in the statement of Theorem 1.3. First, given \( i \in \{1, \ldots, n\} \), either \( \sigma(i) = i \), which happens \( \iff O_\sigma(i) = \{i\} \), or \( O_\sigma(i) \) has at least 2 elements. In this second case, write \( O_\sigma(i) = \{a_1, \ldots, a_k\} \) as above, where \( k = \#(O_\sigma(i)) \) and \( \sigma(a_j) = a_{j+1} \) for \( 1 \leq j \leq k - 1 \) and \( \sigma(a_k) = a_1 \). Thus, if \( \rho \) is the \( k \)-cycle \( (a_1, \ldots, a_k) \), then \( \rho(a_j) = a_{j+1} = \sigma(a_j) \) for all \( a_j \), and \( \rho(r) = r \) if \( r \notin \{a_1, \ldots, a_k\} = O_\sigma(i) = \text{Supp} \rho \).

Now list the orbits of \( \sigma \) as \( O_1, \ldots, O_N \), say, where \( O_1, \ldots, O_M \) are the orbits with at least two elements and \( O_{M+1}, \ldots, O_N \) are the one-element orbits. For each orbit \( O_r \) with \( r \leq M \), we have found a cycle \( \rho_r \) of length at least 2, with \( \text{Supp} \rho_r = O_r \), \( \rho_r(i) = \sigma(i) \) if \( i \in O_r \), and \( \rho_r(i) = i \) if \( i \notin O_r \). Then the cycles \( \rho_1, \ldots, \rho_M \) are disjoint, and by inspection \( \sigma = \rho_1 \cdots \rho_M \).

This proves that \( \sigma \) is a product of disjoint cycles, and the uniqueness up to order is easy to check from the construction.

The proof shows the following:

**Corollary 1.8.** Let \( \sigma \in S_n \), and suppose that the orbits of \( \sigma \) are \( O_1, \ldots, O_N \), where \( \#(O_i) = k_i \geq 2 \) if \( i \leq M \) and \( \#(O_i) = 1 \) if \( i > M \). Then

\[
\sigma = \rho_1 \cdots \rho_M,
\]

where each \( \rho_i \) is a \( k_i \)-cycle, \( \text{Supp} \rho_i = O_i \), and hence, if \( i \neq j \), \( \rho_i \) and \( \rho_j \) are disjoint.

2 The sign of a permutation

We now look for further ways to factor elements of \( S_n \), not in general uniquely. First, we note that the \( k \)-cycle \( (1, \ldots, k) \) is a product of transpositions:

\[
(1, \ldots, k) = (1, k)(1, k-1) \cdots (1, 3)(1, 2),
\]

since the right hand side sends 1 to 2, 2 to 1 and then to 3, 3 to 1 and then to 4, and so on, and finally \( k \) to 1. There is nothing special about choosing the cycle \( (1, \ldots, k) \): if \( a_1, \ldots, a_k \) are \( k \) distinct elements of \( \{1, \ldots, n\} \), then

\[
(a_1, \ldots, a_k) = (a_1, a_k)(a_1, a_{k-1}) \cdots (a_1, a_3)(a_1, a_2).
\]
Hence every $k$-cycle is a product of $k - 1$ transpositions. Since every permutation $\sigma \in S_n$ is a product of $k$-cycles and every $k$-cycle is a product of transpositions, we conclude:

**Theorem 2.1.** Every element $\sigma \in S_n$ is a product of transpositions. \qed

Intuitively, to permute a set with $n$ elements, it is enough to successively switch two at a time. This says that the \binom{n}{2} transpositions $(i, j)$ with $i < j$ generate the group $S_n$. In fact, as in the homework, $S_n$ can be generated by 2 elements, although in many ways the most natural generating set is given by the $n - 1$ elements $(1, 2), (2, 3), \ldots, (n - 1, n)$.

It is easy to see that there is in general no way to write a permutation uniquely as a product of transpositions. For example, we can always insert the identity which is a product $(i, j)(i, j)$ of a transposition with itself. For another example, corresponding to the fact that $(1, 2, 3) = (2, 3, 1) = (3, 1, 2)$ and the above recipe for writing a 3-cycle as a product of transpositions, we have

$$(1, 3)(1, 2) = (1, 2)(2, 3) = (2, 3)(1, 3).$$

If this product is taking place in $S_n$, $n \geq 5$, then we have many more ways of writing $(1, 2, 3)$ as a product of transpositions, for example $(1, 2, 3) = (4, 5)(1, 3)(2, 4)(2, 4)(1, 2)(4, 5)$.

Despite the lack of uniqueness, there is one feature that all of the ways of writing a permutation as a product of transpositions have in common:

**Theorem 2.2.** Let $\sigma \in S_n$, and suppose that $\sigma = \tau_1 \cdots \tau_k = \rho_1 \cdots \rho_\ell$, where the $\tau_i$ and $\rho_j$ are all transpositions. Then $k \equiv \ell \pmod{2}$. In other words, a given element $\sigma$ of $S_n$ can be written as an even number of transpositions or as an odd number of transpositions, but not both.

Thus every element of $S_n$ has a well-defined parity, i.e. is either even or odd, depending on whether it can be written as an even number of transpositions or as an odd number of transpositions. We will describe three different proofs of the theorem; each one is instructive.

**First Proof.** Consider the positive integer

$$N_n = \prod_{1 \leq i < j \leq n} (j - i),$$

the product of all differences of pairs of distinct positive integers between 1 and $n$, where we always take the positive difference. Here the exact value
of $N_n$ is not important (exercise: show that $N_n = (n - 1)!((n - 2)! \cdots 2!)$); clearly, $N_n$ is just a large positive integer, and the main point is that it is nonzero since no factor is zero. Given $\sigma \in S_n$, consider what happens when we consider instead the product $\prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i))$. This is again a product of all possible differences of pairs of distinct positive integers between 1 and $n$, but since $\sigma$ mixes up the order we don’t always subtract the smaller of the pair from the larger. Thus, there exists a sign, i.e. there exists an element $\varepsilon(\sigma) \in \{\pm 1\}$, such that

$$\prod_{1 \leq i < j \leq n} (\sigma(j) - \sigma(i)) = \varepsilon(\sigma)N_n = \varepsilon(\sigma) \prod_{1 \leq i < j \leq n} (j - i).$$

Note that $\varepsilon(\sigma)$ just depends on $\sigma$, not on how $\sigma$ is expressed as a product of transpositions. The two main facts we need about $\varepsilon(\sigma)$ are

1. For all $\sigma_1, \sigma_2 \in S_n$, $\varepsilon(\sigma_1 \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2)$ (i.e. $\varepsilon$ is multiplicative); and
2. If $\tau = (a, b)$ is a transposition, then $\varepsilon(\tau) = -1$.

Assuming (1) and (2), let us finish the proof. If $\sigma = \tau_1 \cdots \tau_k$ is a product of transpositions, then using (1) repeatedly and then (2), we see that

$$\varepsilon(\sigma) = \varepsilon(\tau_1 \cdots \tau_k) = \varepsilon(\tau_1) \cdots \varepsilon(\tau_k) = (-1)^k.$$ 

Hence $\varepsilon(\sigma) = 1$ if $\sigma$ is a product of an even number of transpositions and $\varepsilon(\sigma) = -1$ if $\sigma$ is a product of an even number of transpositions. In particular, if also $\sigma = \rho_1 \cdots \rho_\ell$ where the $\rho_i$ are transpositions, then $\varepsilon(\sigma) = (-1)^\ell = (-1)^k$, and hence $k \equiv \ell \pmod 2$.

We will not write down the proof of (1); it follows from carefully looking at how $\varepsilon(\sigma)$ is defined. However, we will prove (2). Suppose that $\tau = (a, b)$ is a transposition, where we may assume that $a < b$. We examine for which pairs $i < j$, $1 \leq i < j \leq n$, the difference $\tau(j) - \tau(i)$ is negative. Note that, if neither $i$ nor $j$ is $a$ or $b$, then $\tau(j) - \tau(i) = j - i > 0$. Thus we may assume that at least one of $i$, $j$ is either $a$ or $b$. For the moment, we assume that $(i, j) \neq (a, b)$, i.e. either $j = a$ or $i = b$, but not both. If $j = a$, then $i < a$, hence $i \neq b$, and

$$\tau(a) - \tau(i) = b - i > a - i > 0.$$ 

Thus the difference of the pairs $a - i$ don’t change sign. Similarly, if $i = b$ and $j > b$, then $\tau(j) - \tau(b) = j - a > j - b > 0$. So the only possible sign changes come from differences $j - a$, with $j > a$, or $b - i$ with $i < b$. Consider first the case of $j - a$. If $j > b$, then $\tau(j) - \tau(a) = j - b > 0$. If $b > j > a,$
then \( \tau(j) - \tau(a) = j - b < 0 \), and there are \( b - a - 1 \) such \( j \), so these contribute a factor of \((-1)^{b-a-1}\). Likewise, when we look at the difference \( b - i \), then \( \tau(b) - \tau(i) = a - i \). The sign is unchanged if \( i < a \) but becomes negative if \( a < i < b \). Again, there are \( b - a - 1 \) such \( i \), so these contribute another factor of \((-1)^{b-a-1}\) which cancels out the first factor of \((-1)^{b-a-1}\).

The only remaining pair of integers that we have not yet considered is the pair where \( i = a \) and \( j = b \). Here \( \tau(b) - \tau(a) = a - b = -(b - a) \), so there is one remaining factor of \(-1\) which is not canceled out. Thus, the total number of sign changes is

\[
\varepsilon(\tau) = (-1)^{b-a-1}(-1)^{b-a-1}(-1) = -1.
\]

Finally, we note a fact which is useful in Modern Algebra II: it didn’t matter that we chose the integers between 1 and \( n \) and looked at their differences. We could have started with any sequence \( t_1, \ldots, t_n \) of, say, real numbers such that the \( t_i \) are all distinct and looked at the product

\[
\prod_{1 \leq i < j \leq n} (t_i - t_j).
\]

Comparing this product with the product where we permute the pairs, i.e.

\[
\prod_{1 \leq i < j \leq n} (t_{\sigma(i)} - t_{\sigma(j)}),
\]

the above analysis shows that

\[
\prod_{1 \leq i < j \leq n} (t_{\sigma(i)} - t_{\sigma(j)}) = \varepsilon(\sigma) \prod_{1 \leq i < j \leq n} (t_i - t_j),
\]

where \( \varepsilon(\sigma) \in \{\pm1\} \) is the sign factor introduced above.

**Second Proof.** This proof uses basic properties of determinants, which are closely connected with permutations and signs, so one has to be careful not to make a circular argument. For \( \sigma \in S_n \), we will define \( \varepsilon(\sigma) \in \{\pm1\} \) which has the properties (1) and (2) from the first proof, thus showing that, if \( \sigma = \tau_1 \cdots \tau_k \) is a product of \( k \) transpositions, then \( \varepsilon(\sigma) = (-1)^k \); the argument then concludes as in the first proof. To define \( \varepsilon(\sigma) \), we first associate to \( \sigma \) an \( n \times n \) matrix \( P(\sigma) \). Recall from linear algebra that an \( n \times n \) matrix \( P(\sigma) \) is the same thing as a linear map \( \mathbb{R}^n \to \mathbb{R}^n \), which we also denote by \( P(\sigma) \), and that such a linear map is specified by its values on the standard basis \( e_1, \ldots, e_n \); in fact, the value \( P(\sigma)(e_i) \), written as a column vector, is the \( i \)th column of \( P(\sigma) \). So define \( P(\sigma)(e_i) = e_{\sigma(i)} \). Then \( P(\sigma) \) is a permutation matrix: each row and each column have the property that all of the entries except for one of them are 0, and the nonzero entry is 1. Clearly \( P(1) = I \). A calculation shows that

\[
P(\sigma_1 \sigma_2)(e_i) = e_{\sigma_1 \sigma_2(i)};
\]

\[
P(\sigma_1)P(\sigma_2)(e_i) = P(\sigma_1)(e_{\sigma_2(i)}) = e_{\sigma_1 \sigma_2(i)}.
\]
Hence $P(\sigma_1 \sigma_2)$ and $P(\sigma_1)P(\sigma_2)$ have the same value on each basis vector, hence on all vectors, and thus $P(\sigma_1 \sigma_2) = P(\sigma_1)P(\sigma_2)$. Now, to convert the matrix $P(\sigma)$ to a number, we have the determinant $\det$. Define $\varepsilon(\sigma) = \det P(\sigma)$. As it stands, $\varepsilon(\sigma)$ is just a real number. However, using the multiplicative property of the determinant, for all $\sigma_1, \sigma_2 \in S_n$, we have

$$\varepsilon(\sigma_1 \sigma_2) = \det(P(\sigma_1 \sigma_2)) = \det(P(\sigma_1)P(\sigma_2)) = \det P(\sigma_1) \det P(\sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2).$$

Clearly $\varepsilon(1) = \det I = 1$. Moreover, if $\tau = (i, j)$ is a transposition, then $P(\tau)$ is obtained from $I$ by switching the $i$th and $j$th columns of $I$. Another well-known property of the determinant says that, in this case, $\det P(\tau) = -\det I = -1$. Putting this together, we see that, if $\sigma = \tau_1 \cdots \tau_k$ is a product of transpositions, then, as in the first proof,

$$\varepsilon(\sigma) = \varepsilon(\tau_1 \cdots \tau_k) = \varepsilon(\tau_1) \cdots \varepsilon(\tau_k) = (-1)^k,$$

and hence $k$ is either always even or always odd.

**Third Proof.** In this proof, we don’t define the function $\varepsilon$ directly, but just use information about the orbits of $\sigma$ to see that $\sigma$ cannot simultaneously be the product of an even and an odd number of transpositions. For any $\sigma \in S_n$, define $N(\sigma)$ to be the number of orbits of $\sigma$, including the one-element orbits. For example, as we saw earlier, if $\sigma$ is an $r$-cycle, then $N(\sigma) = n - r + 1$. The main point of the third proof is then:

**Claim 2.3.** If $\sigma \in S_n$ is a product of $k$ transpositions, i.e. $\sigma = \tau_1 \cdots \tau_k$ where each $\tau_i$ is a transposition, then

$$N(\sigma) \equiv n - k \pmod 2.$$

To see that Claim 2.3 implies the theorem, suppose that $\sigma = \tau_1 \cdots \tau_k = \rho_1 \cdots \rho_\ell$, where the $\tau_i$ and $\rho_j$ are transpositions. Then we have both that $N(\sigma) \equiv n - k \pmod 2$ and that $N(\sigma) \equiv n - \ell \pmod 2$, hence $k \equiv \ell \pmod 2$.

**Proof of Claim 2.3.** We prove Claim 2.3 by induction on $k$. If $k = 1$, then $\sigma$ is a transposition, hence a 2-cycle, and hence $N(\sigma) = n - 2 + 1 = n - 1$ by the discussion before the statement of the claim, so the congruence in the claim is in fact an equality. By induction, suppose that the claim is true for all products of $k$ transpositions, and consider a product of $k + 1$ transpositions, say $\sigma = \tau_1 \cdots \tau_k \tau_{k+1}$. If we set $\rho = \tau_1 \cdots \tau_k$, then by the inductive hypothesis $N(\rho) \equiv n - k \pmod 2$, and we must show that $N(\rho \tau_{k+1}) \equiv n - k - 1 \pmod 2$. Clearly, it is enough (since $1 \equiv -1 \pmod 2$) to show the following:
**Claim 2.4.** If \( \rho \in S_n \) and \( \tau \in S_n \) is a transposition, then

\[
N(\rho \tau) \equiv N(\rho) + 1 \pmod{2}.
\]

*In fact, \( N(\rho \tau) = N(\rho) \pm 1 \).*

**Proof of Claim 2.4.** Label the orbits of \( \rho \) (including the one-element orbits) as \( O_1, \ldots, O_N \), where \( N = N(\rho) \). As usual, we can assume that \( O_1, \ldots, O_M \) have at least two elements and that \( O_{M+1}, \ldots, O_N \) are the one-element orbits. We must show that \( \rho \tau \) has \( N \pm 1 \) orbits. Since \( \tau \) is a transposition, we can write \( \tau = (a, b) \) for some \( a, b \in \{1, \ldots, n\}, \ a \neq b \).

**Case I.** There is some orbit \( O_i \) of \( \rho \), necessarily with at least two elements, such that \( a, b \in O_i \). In this case, we claim that \( N(\rho \tau) = N(\rho) + 1 \). For each \( j \leq M \), let \( O_j \) correspond to the \( k_j \)-cycle \( \mu_j \), so that \( \rho = \mu_1 \cdots \mu_M \), where the \( \mu_j \) are disjoint and hence commute. For the cycle \( \mu_i \), we know that \( a, b \in \text{Supp} \mu_i = O_i \), and \( a, b \) are not in the support of any \( \mu_j, j \neq i \). Then \( (a, b) \) commutes with all of the \( \mu_j, j \neq i \), and hence \( \rho \tau = \mu_1 \cdots \mu_M \tau = \mu_1 \cdots (\mu_i \tau) \cdots \mu_M \).

If we write \( \mu_i = (i_1, \ldots, i_k) \), then one of the \( i_r \), say \( i_t \), is \( a \), and another, say \( i_s \), is \( b \), and after switching \( a \) and \( b \) we can assume that \( t < s \). Thus we can write \( \mu_i = (i_1, \ldots, i_{t-1}, a, i_{t+1}, \ldots, i_{s-1}, b, i_{s+1}, \ldots, i_k) \). Of course, it is possible that \( t = 1 \) and \( s = 2 \), say. Now consider the product

\[
\mu_i \tau = (i_1, \ldots, i_{t-1}, a, i_{t+1}, \ldots, i_{s-1}, b, i_{s+1}, \ldots, i_k)(a, b).
\]

We see that \( a \) is sent to \( i_{s+1} \), \( i_{s+1} \) is sent to \( i_{s+2} \), \ldots, \( i_{k-1} \) to \( i_k \), \( i_k \) to \( i_1 \), \ldots, and finally \( i_t \) is sent back to \( a \). As for \( b \), it is sent to \( i_{t+1} \), \( i_{t+1} \) is sent to \( i_{t+2} \), \ldots, and finally \( i_{s-1} \) is sent back to \( b \). In other words,

\[
\mu_i \tau = (i_1, \ldots, i_{t-1}, a, i_{t+1}, \ldots, i_s, b, i_{s+1}, \ldots, i_k)(a, b)
\]

\[
= (a, i_{s+1}, \ldots, i_k, i_1, \ldots, i_t)(b, i_{t+1}, \ldots, i_{s-1}) = \mu' \mu'',
\]

say, where \( \mu' \) and \( \mu'' \) are two cycles, disjoint from each other and from all of the other \( \mu_j \) and the remaining one-element orbits. (Note: it is possible that one or both of \( \mu', \mu'' \) is a one-cycle. How could this happen?) It follows that the orbits of \( \rho \tau \) are the \( O_j \) for \( j \neq i \) together with the two new orbits \( O', O'' \) coming from \( \mu' \) and \( \mu'' \). Counting them, we see that there are \( N + 1 \) orbits in all, hence \( N(\rho \tau) = N(\rho) + 1 \).

**Case II.** In this case, we assume that \( a \) and \( b \) are in different orbits, say \( a \in O_i \) and \( b \in O_j \) with \( i \neq j \). We let \( \mu_i, \mu_j \) be the corresponding cycles.
Note that it is possible that either \( O_i \) or \( O_j \) is a one-element orbit, in which case we will simply define \( \mu_i \) or \( \mu_j \) to be the corresponding one-cycle (in particular, as an element of \( S_n \)), it would then be the identity. Then

\[
\rho \tau = \mu_1 \cdots \mu_M \tau = \mu_1 \cdots (\mu_i \mu_j \tau) \cdots \mu_M,
\]

where we have simply moved the \( \mu_i \) and \( \mu_j \) next to each other since all of the \( \mu_i \) commute. In this case, we claim that \( N(\rho \tau) = N(\rho) - 1 \). Write as before

\[
\mu_i = (i_1, \ldots, i_{t-1}, a, i_{t+1}, \ldots, i_k) \quad \text{and} \quad \mu_j = (\ell_1, \ldots, \ell_{s-1}, b, \ell_{s+1}, \ldots, \ell_{k_j}).
\]

A calculation similar to that in Case I shows that

\[
\rho \mu_j \tau = (i_1, \ldots, i_{t-1}, a, i_{t+1}, \ldots, i_k)(\ell_1, \ldots, \ell_{s-1}, b, \ell_{s+1}, \ldots, \ell_{k_j})(a, b)
\]

\[
= (a, \ell_{s+1}, \ldots, \ell_{k_j}, \ell_1, \ldots, \ell_{s-1}, b, i_{t+1}, \ldots, i_k, i_1, \ldots, i_{t-1}) = \mu,
\]
say. Thus the orbits of \( \rho \tau \) are the \( O_k \) for \( k \neq i, j \) together with \( O = \text{Supp} \mu \). Counting these up, we see that there are \( N - 1 \) orbits in all, hence \( N(\rho \tau) = N(\rho) - 1 \). \qed

**Definition 2.5.** If \( \sigma \in S_n \), we define \( \varepsilon(\sigma) = 1 \), if \( \sigma \) is a product of an even number of transpositions, and \( \varepsilon(\sigma) = -1 \), if \( \sigma \) is a product of an odd number of transpositions. Likewise, we define \( \sigma \) to be even if \( \sigma \) is a product of an even number of transpositions (i.e. \( \varepsilon(\sigma) = 1 \)) and odd if \( \sigma \) is a product of an odd number of transpositions (i.e. \( \varepsilon(\sigma) = -1 \)). The content of the previous theorem is then that \( \varepsilon \) is well-defined. Finally we define \( A_n \), the alternating group, to be the subset of \( S_n \) consisting of even permutations.

**Lemma 2.6.** Let \( \sigma_1 \) and \( \sigma_2 \) be elements of \( S_n \). If \( \sigma_1 \) and \( \sigma_2 \) are both even or both odd, then the product \( \sigma_1 \sigma_2 \) is even. If one of \( \sigma_1 \), \( \sigma_2 \) is even and the other is odd, then the product \( \sigma_1 \sigma_2 \) is odd. Equivalently, \( \varepsilon(\sigma_1 \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2) \).

**Proof.** Suppose that \( \sigma_1 \) can be written as a product \( \tau_1 \cdots \tau_k \) of \( k \) transpositions and that \( \sigma_2 \) can be written as a product \( \rho_1 \cdots \rho_\ell \) of \( \ell \) transpositions. Then \( \sigma_1 \sigma_2 = \tau_1 \cdots \tau_k \rho_1 \cdots \rho_\ell \) is a product of \( k + \ell \) transpositions. (Confusingly, for the **product** \( \sigma_1 \sigma_2 \) we take the **sum** of the numbers of transpositions.) From this, the proof of the lemma is clear: if \( k \) and \( \ell \) are both even or both odd, then \( k + \ell \) is even, and if one of \( k \), \( \ell \) is even and the other is odd, then \( k + \ell \) is odd. \qed

**Corollary 2.7.** \( A_n \) is a subgroup of \( S_n \).

**Proof.** If \( \sigma_1, \sigma_2 \in A_n \), then \( \sigma_1 \) and \( \sigma_2 \) are both even, hence \( \sigma_1 \sigma_2 \) is even, hence \( \sigma_1 \sigma_2 \in A_n \) and \( A_n \) is closed under taking products. The element 1 is
even, i.e. \(1 \in A_n\), either formally (1 is the product of 0 transpositions and 0 is even) or by writing \(1 = (a,b)(a,b)\), which is possible as long as \(n \geq 2\). Finally, if \(\sigma \in A_n\), then \(\sigma = \tau_1 \cdots \tau_k\) is the product of \(k\) transpositions, where \(k\) is even. Then

\[
\sigma^{-1} = (\tau_1 \cdots \tau_k)^{-1} = \tau_k^{-1} \cdots \tau_1^{-1} = \tau_k \cdots \tau_1,
\]

i.e. \(\sigma^{-1}\) is the product of the \(k\) transpositions \(\tau_1, \ldots, \tau_k\) in the opposite order. In particular, \(\sigma^{-1}\) is the product of an even number of transpositions, hence \(\sigma^{-1} \in A_n\). Thus \(A_n\) is a subgroup of \(S_n\).

**Example 2.8.** First note that, as we have already seen, a \(k\)-cycle is a product of \(k-1\) transpositions if \(k \geq 1\) (and a 1-cycle is a product of two transpositions). Thus a \(k\)-cycle is is **even** if \(k\) is odd and **odd** if \(k\) is even. Hence we can determine the parity of any product of cycles (whether or not they are disjoint). This gives the following description of the subgroups \(A_n\) for small \(n\):

- \(A_1 = S_1 = \{1\}\).
- \(A_2 = \{1\}\).
- \(A_3 = \{1, (1,2,3), (1,3,2)\}\).
- \(A_4\) consists of 1, the 8 3-cycles, and the 3 products of two disjoint 2-cycles. Hence \(#(A_4) = 12 = \frac{1}{2}#(S_4)\), and in hindsight we also see that \(#(A_2) = 1 = \frac{1}{2}#(S_2)\) and that \(#(A_3) = 3 = \frac{1}{2}#(S_3)\). This is not a coincidence:

**Proposition 2.9.** For \(n \geq 2\), \(#(A_n) = \frac{1}{2}#(S_n) = n!/2\).

**Proof.** Since every element of \(S_n\) is either even or odd, \(S_n\) is the disjoint union of \(A_n\), the set of even permutations, and \(S_n - A_n\), the set of odd permutations. We will find a bijection from \(A_n\) to \(S_n - A_n\). This will imply that \(#(A_n) = #(S_n - A_n)\), hence

\[
#(S_n) = #(A_n) + #(S_n - A_n) = 2#(A_n),
\]

hence \(#(A_n) = \frac{1}{2}#(S_n)\).

To find a bijection from \(A_n\) to \(S_n - A_n\), choose an odd permutation \(\rho\). For example, we can take \(\rho = (1,2)\). Note that this is only possible for \(n \geq 2\), since for \(n = 1\) \(S_1 = \{1\}\) and every permutation is trivially even. Then define a function \(f: A_n \rightarrow S_n - A_n\) by the rule:

\[
f(\sigma) = \rho \cdot \sigma,
\]

the product of \(\sigma\) with the fixed odd permutation \(\rho\). If \(\sigma \in A_n\), then \(\sigma\) is even and hence \(\rho \sigma\) is odd, so that \(f\) is really a function from \(A_n\) to \(S_n - A_n\).

To show that \(f\) is a bijection, it is enough to find an inverse function, i.e. a
function $g: S_n - A_n \to A_n$ such that $f \circ g$ and $g \circ f$ are both the identity (on their respective domains). Define $g: S_n - A_n \to A_n$ by the rule:

$$g(\sigma) = \rho^{-1}\sigma.$$ 

If $\rho$ is odd, then $\rho^{-1}$ is odd, hence, if $\sigma$ is also odd, then $\rho^{-1}\sigma$ is even. Thus $g$ is really a function from $S_n - A_n$ to $A_n$. Finally, if $\sigma \in A_n$, then $g \circ f(\sigma) = \rho^{-1}\rho\sigma = \sigma$, and if $\sigma \in S_n - A_n$, then $f \circ g(\sigma) = \rho\rho^{-1}\sigma = \sigma$. Thus $f \circ g = \text{Id}_{S_n - A_n}$ and $g \circ f = \text{Id}_{A_n}$, so that $g = f^{-1}$ and $f$ is a bijection. $\Box$

The group $A_n$ is a group of fundamental importance, for reasons which will gradually emerge this semester, and will play a major role in Modern Algebra II as well.