WORKSHEET ON SYMMETRIC GROUPS, MATH 504, FALL 2020

DUE FRIDAY, NOVEMBER 6

1. Introduction

This is an independent study homework where you'll have to develop the structure theory for symmetric groups by yourself. We recommend that you try proving all the facts on your own to gain familiarity and eventually proficiency with symmetric groups. If you are stuck consulting with literature is totally fine, of course. Just give yourself sufficient time to get stuck and unstuck.

This homework will also be an experiment on "peer review". We'll try to implement the standard scientific journal review procedure. Once submitted your assignment will be distributed back to your fellow students for review. Please consider this as an exercise in reviewing a journal submission - something you eventually might be doing. Check the writing assigned to you for correctness, readability and style, and leave respectful comments and suggestions for improvement. After the first round of reviews (the timeline will be announced on the assignment page on canvas), the papers go back to the authors who address the comments and send their revised assignments back to the reviewers to see if they are satisfied. If they are, the paper then goes back to Curtiss who will play the role of the editor. He'll evaluate the work of the reviewer and, hopefully, "accept the paper" in which case he'll record the final result and assign homework points. In rare circumstances the paper might go back to the reviewer with a request to do a more thorough job but hopefully it will never happen for us!

We'll discuss the process further in class. I expect the reviewing part to take up to a week. To let you focus on it, there will be no written homework that week, only presentation problems.

Your assignment is to supply all the proofs in the text below.

2. Generators of S_n

Definition 2.1. The symmetric group on n elements, denoted S_n is a group of self-bijections (or permutations) of the set $X = \{1, 2, \dots, n\}$. For the purposes of this worksheet, we multiply permutations from left to right. You could multiply from right to left as well - this will not change any of the main results but you'll need to adjust some of the formulas. Either way is fine as long as you (and I:)) are consistent.

Notation. Let $\sigma \in S_n$. Hence, $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ is a bijection. The commonly used notation for the corresponding permutation is the following:

$$\begin{pmatrix} 1 & 2 & \dots & \text{n} \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

Definition 2.2. A permutation $\sigma \in S_n$ is called a *cycle* if there exists a subset $\{x_1, \dots, x_k\} \subseteq \{1, 2, \dots, n\}$ such that $\sigma(x_i) = x_{i+1}$ and $\sigma(y) = y$ for any $y \neq x_i$. The standard notation for such a permutation is

$$(x_1, x_2, \dots, x_k).$$

Two cycles $(x_1, x_2, \dots, x_k)$ and $(y_1, y_2, \dots, y_\ell)$ are called *disjoint* if the sets $\{x_1, x_2, \dots, x_k\}$ and $\{y_1, y_2, \dots, y_\ell\}$ do not intersect.

Example 2.3. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix} = (234)$

Symmetric group has various sets of generators. For example:

Proposition 2.4. *The symmetric group S_n can be generated by two elements: a cycle of length 2 and a cycle of length n.*

Proof. Consider the two elements of S_n: $\alpha = (123\dots n)$ and $\beta = (12)$. I claim that α, β generate all of S_n. First, let us rephrase the definition of α: it is the map that sends $a \mapsto a + 1$, where the output is modded by n to ensure it stays in $\{1, 2, \dots, n\}$ (in fact, in the rest of the proof, the maps will implicitly have $(\mathsf{mod}\, n)$ in them, but I will not write them out). Thus, α^k (performing α k-times) is the map that sends $a \mapsto a+k(\mathsf{mod}\, n)$, and similarly α^{-k} is the map that sends $a \mapsto a-k$.

Now consider $\alpha^{k-1}\beta\alpha^{k-1}$ (doing α^{k-1}, then β, the α^{k-1}). α^{k-1} sends $k \mapsto 1$, $k + 1 \mapsto 2$, and so on. Then β keeps everything the same, except $2 \mapsto 1$ and $1 \mapsto 2$. That is to say, in conjunction $\alpha^{k-1}\beta$ sends $k \mapsto 2$, $k + 1 \mapsto 1$, and for $a \neq 1, 2$, $(k - 1) + a \mapsto a$. Then α^{k-1} sends $a \mapsto a + (k - 1)$ (so $1 \mapsto k, 2 \mapsto k+1$, and so on); and so we see that all together, $\alpha^{k-1}\beta\alpha^{k-1}$ takes $k \mapsto k+1$, $k+1 \mapsto k$, and keeps everything else the same. In other words, we've shown that $\alpha^{k-1}\beta\alpha^{k-1} = (k, k + 1)$.

Now observe that from these, we can now construct all the swaps of the form $(1k)$ — to start off, we have $(13) = (12)(23)(12)$, and $(14) = (13)(34)(13)$, and from here it's easy to see that $(1k) = (1, k - 1)(k - 1, k)(1, k - 1)$ (and so by induction all the $(1k)$ are constructed). With these, we can construct ALL swaps by $(k\ell) = (1k)(1\ell)(1k)$.

Finally, with all such swaps $(k\ell)$ we can construct any permutation in S_n; we show this by showing that given any permutation $\sigma \in S_n$, we can perform a series of swaps to get it to id, the identity permutation, i.e. showing for some swaps $\gamma_1, \gamma_2, \dots, \gamma_m$, $\sigma\gamma_1 \cdots \gamma_m = \text{id}$, which would of course imply that $\sigma = \gamma_m^{-1} \cdots \gamma_1^{-1} = \gamma_m \cdots \gamma_1$ (the inverse of a swap is itself):

Suppose that $\sigma(1) = x_1$; then let $\gamma_1 = (1, x_1)$ (if $x_1 = 1$ γ_1 is obviously id). Then $\sigma\gamma_1$ will send $1 \mapsto 1$. Then define $\gamma_2 = (2, x_2)$ where $x_2 = [\sigma\gamma_1](2)$ (x_2 can't be 1 because $\sigma\gamma_1$ is injective, and it already sent $1 \mapsto 1$); this guarantees that $\sigma\gamma_1\gamma_2$ sends $1 \mapsto 1, 2 \mapsto 2$. Continuing like this, we define $\gamma_k = (k, x_k)$ where $x_k = [\sigma\gamma_1 \dots \gamma_{k-1}](k)$ (where $x_k \notin \{1, \dots, k - 1\}$ because $\sigma\gamma_1 \dots \gamma_{k-1}$ is injective, and it already sent $1, \dots, k-1$ to $1, \dots, k$ resp.), thus guaranteeing that $\sigma\gamma_1 \dots \gamma_{k-1}$ sends $1, \dots, k$ to $1, \dots, k$ resp. So by $\sigma\gamma_1 \dots \gamma_{n-1}$, we already have that $1 \mapsto 1, \dots, n - 1 \mapsto n - 1$, which forces $n \mapsto n$. Some of the γ_i may be trivial (the id map), but regardless, we have shown that any permutation can be turned into the identity in m swaps ($m \leq n - 1$), so to summarize, we have shown that we can construct ANY permutation in S_n from the two building blocks α, β. □

Proposition 2.5. *Any permutation $\sigma \in S_n$ can be written as a composition of disjoint cycles.*

Proof. Define for each $i \in [n] := \{1, \dots, n\}$ $C_i := \{\sigma^m(i) : m \in \mathbb{N}\}$, where reminder $\sigma^m(i)$ is performing σ m-times on i. Obviously, because σ outputs only in $[n]$, each $C_i \subseteq [n]$. As the choice of letter suggests, the C_i are in fact cycles: let $n_i \in \mathbb{N}$ be the first natural number at which $\sigma^{n_i}(i)$ repeats something seen earlier; first, such a repeat must occur, because there are only n possible outputs (so in fact $n_i \leq n-1$), and second $\sigma^{n_i}(i)$ must be i itself, because if it were instead $\sigma^k(i)$ for some $k \in \{1, \dots, n_i - 1\}$, then we could perform σ^{-1} to get that $\sigma^{n_i-1}(i) = \sigma^{k-1}(i)$, contradicting that n_i was minimal.

Furthermore, if C_i contains some j, then we have that $C_i = C_j$ (implying that if $k \in C_i$ and C_j, then $C_i = C_k = C_j$; that is to say, all the C_i are either exactly equal or disjoint). To see this, observe that $j \in C_i \implies \sigma^k(i) = j$ (we can restrict $k \in \{0, \ldots, n_i - 1\}$, because it repeats after that), and so any $\sigma^m(j) = \sigma^m(\sigma^k(i)) = \sigma^{k+m}(i)$, and so $C_j \subseteq C_i$. For the other direction, we can write $\sigma^k(i) = j \implies i = \sigma^{-k}(j) \implies \sigma^{n_i}(i) = \sigma^{n_i}\sigma^{-k}(j) \implies i = \sigma^{n_i-k}(j)$, which implies that $i \in C_j$, and so we can use the above ($\subseteq$) direction to give that $C_i \subseteq C_j$.

Thus, selecting distinct C_i, calling them $\hat{C}_1, \ldots, \hat{C}_m$ (clearly $m \leq n$), we get disjoint cycles composing to σ; i.e. writing $\hat{C}_i$ as $(c_{i,1}, \ldots, c_{i,\hat{n}_i})$, we have $\sigma = (c_{1,1}, \ldots, c_{1,\hat{n}_m}) \cdots (c_{m,1}, \ldots, c_{m,\hat{n}_m})$, because each cycle describes what happens to all the elements in the cycle, and disjointness means that we can compose without interference between different cycles. □

Remark 2.6. Such decomposition is unique up to the order of the factors.

Example 2.7. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{pmatrix} = (15)(234)$

$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix} = (12)(3)(45) = (12)(45)$

A cycle of length 1, such as (3) in the example above, just indicates that the corresponding element is fixed under the permutation. These are often skipped when permutation is written as a product of cycles.

We now describe the conjugacy classes of S_n (= the orbits under the action by conjugation of S_n on itself).

Theorem 2.8. *Let $\sigma, \tau \in S_n$. Then σ and τ are conjugate if and only if their decompositions into disjoint cycles can be put into one-to-one correspondence such that the corresponding cycles are of the same length.*

In particular, the conjugacy class of a single cycle consists of all cycles of the same length.

Proof. If we have a cycle $\Gamma = (x_1, \ldots, x_m)$, and $\alpha \in S_n$, consider $\alpha\Gamma\alpha^{-1}$ (performing α, then Γ, then α^{-1}). For each $x_i\{x_1, \ldots, x_m\}$, let a_i be the element α takes to x_i ($a_i = \alpha^{-1}(x_i)$). For all elements $\notin \{a_1, \ldots, a_m\}$, after performing α on them, none of them are $\in \{x_1, \ldots, x_m\}$, and so Γ does nothing to them, and α^{-1} takes them back to where they were at the start; i.e. $\alpha\Gamma\alpha^{-1}$ does not affect anything outside $\{a_1, \ldots, a_m\}$. For $a_i \in \{a_1, \ldots, a_m\}$, α takes $a_i \mapsto x_i$, Γ maps $x_i \mapsto x_{i+1}$ (indices taken $\mathsf{mod}\, m$), and α^{-1} sends $x_{i+1} \to a_{i+1}$. That is to say, $\alpha\Gamma\alpha^{-1}$ maps $a_i \mapsto a_{i+1}$ (where "a_{m+1}" is "a_1"), so it is EXACTLY the cycle $(\alpha^{-1}(x_1), \ldots, \alpha^{-1}(x_m))$. Let us call this new cycle ${}_\alpha\Gamma$ (which has length m, the length of Γ).

This proves ($\implies$): from Proposition 2.5, we can write any $\sigma \in S_n$ as a composition of disjoint cycles $\Gamma_1 \cdots \Gamma_m$; then if τ is conjugate to σ, then for some $\alpha \in S_n$, $\tau = \alpha\sigma\alpha^{-1} = (\alpha\Gamma_1\alpha^{-1}) \cdots (\alpha\Gamma_m\alpha^{-1}) = {}_\alpha\Gamma_1 \cdots {}_\alpha\Gamma_m$, where $\Gamma_i \mapsto {}_\alpha\Gamma_i$ is a one-to-one correspondence of the cycles of σ and τ s.t. corresponding cycles are the same length.

Now for ($\impliedby$): if we have that $\sigma = \Gamma_1 \cdots \Gamma_m$, and $\tau = \Gamma'_1 \cdots \Gamma'_m$ (where $\Gamma_i = (x_{i,1}, \ldots, x_{i,n_i})$, and $\Gamma'_i = (x'_{i,1}, \ldots, x'_{i,n_i})$), where all the Γ_i are disjoint, and all the Γ'_i are disjoint. Let us now define a permutation α as follows: for every $x'_{i,j}$, define $\alpha(x'_{i,j}) = x_{i,j}$ (everywhere else just have α do nothing) — this is a bijection because all the $x'_{i,j}$ are distinct, and all the $x_{i,j}$ are distinct. Then, in the above notation, all the Γ'_i become ${}_\alpha\Gamma_i$, and we've already seen that if $\sigma = \Gamma_1 \cdots \Gamma_m$ and $\tau = {}_\alpha\Gamma_1 \cdots {}_\alpha\Gamma_m$, then $\tau = \alpha\sigma\alpha^{-1}$. □

Remark 2.9. The group S_n is non-commutative for $n \geq 3$. Nonetheless, disjoint cycles always commute.

Definition 2.10. A transposition is a cycle of length 2.

Proposition 2.11. *The symmetric group S_n is generated by transpositions.*

Proof. Proven in the last two paragraphs of Proposition 2.4. □

3. Alternating group

Note that the symmetric group S_n acts on polynomials on n variables. Namely, we define

$$(\sigma f)(x_1, \dots, x_n) = f(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$$

In short, $\sigma f = f \circ \sigma^{-1}$. For example, for $n = 3$, $\sigma = (12)$ a cycle of length 2,

$$\sigma(x_1^2 x_2 x_3^5) = x_2^2 x_1 x_3^5 = x_1 x_2^2 x_3^5.$$

For $\sigma = (123)$,

$$\sigma(x_1^2 x_2 x_3^5) = x_3^2 x_1 x_2^5 = x_1 x_2^5 x_3^2.$$

Let

$$f(x_1, \dots, x_n) = \prod_{i<j} (x_i - x_j).$$

Question. Do you know for which matrix $f(x_1, \dots, x_n)$ is a determinant?

Note that for any $\sigma \in S_n$, we have $\sigma f = \pm f$. Define a map

$$\mathrm{Sgn} : S_n \to \mathbb{Z}/2\mathbb{Z}$$

via $\mathrm{Sgn}(\sigma) = -1$ if $\sigma f = -f$ and $\mathrm{Sgn}(\sigma) = 1$ otherwise.

Proposition 3.1. Sgn *is a group homomorphism.*

Proof. We just need to verify that $\mathrm{Sgn}(\sigma\tau) = \mathrm{Sgn}(\sigma)\,\mathrm{Sgn}(\tau)$: from the definition,

$$\begin{aligned}
([\sigma\tau]f)(x_1, \dots, x_n) &= f(x_{[\sigma\tau]^{-1}(1)}, \dots, x_{[\sigma\tau]^{-1}(n)}) \\
&= f(x_{[\tau^{-1}\sigma^{-1}](1)}, \dots, x_{[\tau^{-1}\sigma^{-1}](1)](n)}) \\
&= f(x_{\sigma^{-1}(\tau^{-1}(1))}, \dots, x_{\sigma^{-1}(\tau^{-1}(n))})
\end{aligned}$$

Defining $y_i = x_{\tau^{-1}(i)}$ (a "renaming of the variables"), we have that $(\tau f)(x_1, \dots, x_n) = f(y_1, \dots, y_n)$, and $([\sigma\tau]f)(x_1, \dots, x_n) = (\sigma f)(y_1, \dots, y_n)$. Also notice that the definition of Sgn is equivalent to saying $\mathrm{Sgn}(\sigma) = \frac{\sigma f}{f}$. Then, we have

$$\begin{aligned}
\mathrm{Sgn}(\sigma\tau) &= \frac{([\sigma\tau]f)(x_1, \dots, x_n)}{f(x_1, \dots, x_n)} = \frac{([\sigma\tau]f)(x_1, \dots, x_n)}{(\tau f)(x_1, \dots, x_n)} \cdot \frac{(\tau f)(x_1, \dots, x_n)}{f(x_1, \dots, x_n)} \\
&= \frac{(\sigma f)(y_1, \dots, y_n)}{f(y_1, \dots, y_n)} \cdot \frac{(\tau f)(x_1, \dots, x_n)}{f(x_1, \dots, x_n)} = \mathrm{Sgn}(\sigma)\,\mathrm{Sgn}(\tau).
\end{aligned}$$

□

Definition 3.2. A permutation $\sigma \in S_n$ is called *even* if $\mathrm{Sgn}(\sigma) = 1$. Otherwise, it is called *odd*.

Corollary 3.3. *The subset of all even permutations is a normal subgroup of S_n.*

Proof. The subset of all even permutations is the kernel of Sgn, and we just showed that Sgn is a group homomorphism, and we know that kernels of homomorphisms are normal subgroups. □

Definition 3.4. The subgroup of even permutations is called an *alternating group A_n*.

As we shall see in the following theorem, the sign of a permutation can be determined from its decomposition into transpositions.

Theorem 3.5. (1) *If $\tau \in S_n$ is a transposition, then* $\operatorname{Sgn}(\tau) = -1$

(2) *A permutation σ is even if and only if it can be written as a product of even number of transpositions.*

Proof. For (1): suppose $\tau = (k, k+1)$. Then, consider all pairs (i,j), $i,j \in \{1,\dots,n\}$, $i<j$. If $i<k$ (so $j \geq k$), then even after doing τ, i is not changed (still $<k$), and j must remain $\geq k$, so all of those pairs stay the same. Similarly, if $j > k+1$, then after the swap j doesn't change (still $> k+1$), and i has to remain $\leq k+1$, so all of these pairs stay the same. The only pair left is $i=k, j=k+1$, and it does in fact swap $(x_k - x_{k+1})$ to $(x_{k+1}-x_k)$. Thus, the sign is flipped once, so $\operatorname{Sgn}((k,k+1)) = -1$. From paragraph 2 of the proof of Propostion 2.4, all transpositions of the form $(1k)$ could be written as $(1,k-1)(k-1,k)(1,k-1)$, so using induction on k we can show that $\operatorname{Sgn}((1k)) = \operatorname{Sgn}((1,k-1))\operatorname{Sgn}((k-1,k))\operatorname{Sgn}((1,k-1)) = -1\cdot -1 \cdot -1 = -1$ (using that Sgn is a homomorphism). From paragraph 3 of the same proof, ANY transposition $(k\ell) = (1k)(1\ell)(1k)$, and again using that Sgn is a homomorphism, $\operatorname{Sgn}((k\ell)) = -1 \cdot -1 \cdot -1 = -1$.

Suppose $\sigma = \tau_1 \cdots \tau_m$. Then, from (1) and the fact that Sgn is a homomorphism, we have that $\operatorname{Sgn}(\sigma) = \operatorname{Sgn}(\tau_1)\cdots\operatorname{Sgn}(\tau_m) = (-1)^m$, which is 1 (i.e. σ is even) if and only if m is even (i.e. σ can be written as a product of an even number of transpositions). □

We now determine generators of A_n.

Theorem 3.6. *The group A_n is generated by 3-cycles of the form $(12i)$, $3 \leq i \leq n$.*

Proof. For $\langle \{(12i)\}_{i=3}^n \rangle \geq A_n$: given any pair of transpositions $(ab)(cd)$ (only restrictions: $a \neq b, c \neq d$), if we can show that it can be written as some product of the above 3-cycles, then because A_n is generated by such pairs of transpositions by Theorem 3.5(2) (pairs to ensure there are an even number of transpositions), then the claim is proven. Recall from paragraph 3 from the proof of Proposition 4 that $(ab) = (1a)(1b)(1a)$, and so $(ab)(cd) = (1a)(1b)(1a)(1c)(1d)(1c) = [(1a)(1b)][(1a)(1c)][(1d)(1c)]$. Thus, it suffices to show that each $(1x)(1y)$ is a product of these 3-cycles: note that $(12i) = (12)(1i)$, and $(1i)(12) = (1i2) = (12i)(12i)$, and so $(1x)(1y) = (1x)(12)(12)(1y) = (1x2)(12y) = (12x)(12x)(12y)$.

The other direction of inclusion is trivial: all $(12i) = (12)(1i)$, so they have Sgn value 1, and so all products of such $(12i)$ must also have $\operatorname{Sgn} = 1$, i.e. they must be in A_n. □

4. Derived series for S_n

Theorem 4.1. *The symmetric group S_n is solvable for $n = 2, 3, 4$.*

Write down the explicit derived series in the proof.

Proof. For S_2: S_2 consists of two elements, $() = \text{id}$ or (12). It is clear that it is abelian, and so the commutator subgroup $[S_2, S_2] = e$, so the derived series is $\boxed{e \leq S_2}$.

We actually prove Theorem 4.2(1) first. For S_3: we know from Theorem 4.2(1) that $[S_3, S_3] = A_3$, which has $3!/2 = 3$ elements, and we know that any group of order p (p prime) is cyclic (by Lagrange), and so $A_3 = C_3$ is abelian. Thus, the derived series is $\boxed{e \leq A_3 \leq S_3}$

For S_4, Theorem 4.2(1) gives that $[S_4, S_4] = A_4$. Now we need to find $[A_4, A_4]$. From the last paragraph of the proof of Proposition 2.4, any permutation in S_4 can be written as a product of $\leq n - 1 = 4 - 1 = 3$ transpositions, and so all permutations of A_4 can be written as a product of 2 transpositions (or 0, i.e $() = \mathrm{id}$). There are a total of $4!/2 = 12$ elements in A_4, where one is $()$, and 3 more are of products of two disjoint cycles $(12)(34), (13)(24), (14)(23)$ (exactly 3, because once we fix the element that 1 goes to, everything else is fixed). That means that there are 8 more elements where the two cycles share one element; these can be written as 3-cycles. So suppose g, h are elements in this set of 8 elements. If h is a power $k = 1, 2$ of g, then obviously $[g, h] = ghg^{-1}h^{-1} = g^{k+1}g^{-1-k} = e$. So writing $g = (abc)$, we have 6 possibilities for h (elements that permute a, b, d or a, c, d or b, c, d), which we tackle individually below:

- $(abc)(abd)(cba)(dba)$ sends $a \mapsto b \mapsto d \mapsto d \mapsto b$ (which I will henceforth shorthand to $abddb$. Similarly, we have $bccba$, $cabad$, and $ddacc$, and so this is $(ab)(cd)$.
- $(abc)(acd)(cba)(dca)$ maps $abbad$, $bcddc$, $cacbb$, and $ddaca$, so this is $(ad)(bc)$.
- $(abc)(bcd)(cba)(dcb)$ maps $abcbd$, $bcddc$, $caacb$, and $ddbaa$, so this is $(ad)(bc)$.
- $(abc)(adb)(cba)(bda)$ maps $abacc$, $bccbd$, $cadda$, and $ddbab$, so this is $(ac)(bd)$.
- $(abc)(adc)(cba)(cda)$ maps $abbac$, $bcacd$, $cadda$, and $ddcbb$, so this is $(ad)(bc)$.
- $(abc)(bdc)(cba)(cdb)$ maps $abddb$, $bcbaa$, $caacd$, and $ddcbc$, so this is $(ab)(cd)$.

Thus we see that all of these products are in the Klein 4-group $K_4 := \{(), (12)(34), (13)(24), (14)(23)\}$ (and so the commutators of the 3 "disjoint 2-cycle" elements are obviously in K_4) too. Finally, if we have a 3-cycle (cba), and some element in K_4. say $(xy)(zw)$, the commutator will be in K_4 too because out of the 6 representations above, 2 of them correspond to the particular element $(xy)(zw)$, so we can write the commutator as

$$\begin{aligned} & (cba)(abc)(\cdots)(cba)(\cdots)^{-1}(abc)(\cdots)(abc)(\cdots)^{-1}(cba) \\ = & \quad (\cdots)(cba)(\cdots)^{-1}(abc)(\cdots)(abc)(\cdots)^{-1}(cba) \\ = & \quad [(\cdots)(cba)(\cdots)^{-1}(abc)][(\cdots)(abc)(\cdots)^{-1}(cba)], \end{aligned}$$

which is just a product of two elements of the type described in the 6 bullet points above (and so still in K_4). The Klein 4-group is abelian, and so the commutator is e, and so the derived series of S_4 is $\boxed{e \leq K_4 \leq A_4 \leq S_4}$ □

Theorem 4.2. *(1)* $[S_n, S_n] = A_n$
(2) For $n \geq 5$, $[A_n, A_n] = A_n$

Proof. The commutator subgroup $[S_n, S_n]$ is the subgroup generated by all commutators $[g, h] = ghg^{-1}h^{-1}$. Thus, $(12i)(12i) = (12)(1i)(12)(1i) = (12)(1i)(12)^{-1}(1i)^{-1} \in [S_n, S_n]$, and so $(12i) = [(12i)(12i)(12i)](12i) = [(12i)(12i)][(12i)(12i)] \in [S_n, S_n]$. Thus, from Theorem 3.6, we know that $A_n \leq [S_n, S_n]$. For the other direction, $\mathrm{Sgn}(ghg^{-1}h^{-1}) = \mathrm{Sgn}(g)\,\mathrm{Sgn}(h)\,\mathrm{Sgn}(g^{-1})\,\mathrm{Sgn}(h^{-1}) = \mathrm{Sgn}(g)\,\mathrm{Sgn}(g^{-1})\,\mathrm{Sgn}(h)\,\mathrm{Sgn}(h^{-1}) = \mathrm{Sgn}(e)\,\mathrm{Sgn}(e) = 1$, so all elements of $[S_n, S_n]$ must have Sgn value 1.

For (2): $[A_n, A_n] \leq A_n$ follows exactly the same way as above $[S_n, S_n] \leq A_n$. For $\geq$, we again show that all $(12i) \in [A_n, A_n]$ $(i \geq 3)$. Because $n \geq 5$, we can find $x, y \in [n]$ s.t. $\{1, 2, i, x, y\}$ are all distinct. Then, we can write $(12i)(12i) = (12)(1i)(12)(1i) = (12)(xy)(1i)(xy)(xy)(12)(xy)(1i)$ (which is true because (xy) is disjoint from any of the other cycles involving $1, 2, i$, so we can move them around willy-nilly, and here we have 4 of them which multiply to just id). This long product is of the form $ghg^{-1}h^{-1}$, where $g = (12)(xy)$ and $h = (1i)(xy)$, and because we have 2 (an even number) of transpositions, $g, h \in A_n$, and so $(12i)(12i) \in [A_n, A_n]$. From this, we can conclude like we did in (1) that $(12i) \in [A_n, A_n]$, and so $A_n \leq [A_n, A_n]$. □

The following lemma might be useful (prove it if you use it):

Lemma 4.3. *Let i, j, k, ℓ, m be distinct integers. Then*

(1) $(ij)(k\ell) = [(ijk), (ij\ell)]$,
(2) $(ijk) = [(ik), (ij)]$,
(3) $(ijk) = [(ik\ell), (ijm)]$.

Theorem 4.4. *For $n \geq 5$, the group A_n is simple.*

Proof. We want to show that there is no non-trivial proper normal subgroup N of A_n. Suppose there is a non-trivial normal subgroup N; we show that it must be exactly A_n. Because A_n is generated by 3-cycles $(12i)$, it suffices to show that N must contain such 3-cycles. But it turns out that the normality condition of N, and the fact that $n \geq 5$ together make it so that we only need to show that N contains ONE 3-cycle σ. This is because $n \geq 5$ implies that every 3-cycle is conjugate via an $\alpha \in A_n$, and normality means that $N = \alpha N \alpha^{-1}$ for all $\alpha \in A_n$, i.e. N contains all conjugates of σ via $\alpha \in A_n$ (which are exactly all the 3-cycles).

We now prove that $n \geq 5 \implies$ all 3-cycles are conjugate via some element in A_n: if we have an arbitrary 3-cycle τ, we know from Theorem 2.8 that any other 3-cycle is a conjugate of τ via $\alpha \in S_n$. That is to say, picking some concrete 3-cycle like (123), we know there must be some $\alpha \in S_n$ s.t. $\alpha\tau\alpha^{-1} = (123)$. If $\alpha \in A_n$, set $\alpha' = \alpha$. If not, then set $\alpha' = (45)\alpha$ (where $\mathrm{Sgn}(\alpha') = -1 \cdot -1 = 1 \implies \alpha' \in A_n$), and consider $\alpha'\tau(\alpha')^{-1} = [(45)\alpha]\tau[(45)\alpha]^{-1} = (45)[\alpha\tau\alpha^{-1}](45) = (45)(123)(45) = (123)$. Then either way, we will have that τ can in fact be conjugated to (123) by some $\alpha' \in A_n$. Thus, for any 3-cycles τ_1, τ_2, there are $\alpha_1, \alpha_2 \in A_n$ s.t. $\alpha_1\tau_1\alpha_1^{-1} = (123) = \alpha_2\tau_2\alpha_2^{-1} \implies [\alpha_2^{-1}\alpha_1]\tau_1[\alpha_1^{-1}\alpha_2] = \tau_2$, and so the claim is proven.

Now, we have to show that N always must contain one 3-cycle. Let σ be any non-trivial element of N. We know from Proposition 2.5 that $\sigma = \Gamma_1 \cdots \Gamma_m$ for disjoint cycles Γ_i. Luckily, we can prove this by a relatively nice use of casework: our 3 main situations are if (1) at least one Γ_i has length ≥ 4; (2) all Γ_i have length ≤ 3 and some actually have length 3; and (3) all Γ_i have length ≤ 2.

Case 1: by relabelling (we can reorder the Γ_i since they are disjoint cycles), we can say $\Gamma_1 = (x_1, \ldots, x_{n_1})$ has length ≥ 4. Furthermore, we can relabel the x_i (using the glyphs "1", "2", and so on) so that we can write Γ_1 to be $(123 \ldots n_1)$ (so Γ_i for $i \geq 2$ will not involve any of $1, 2, \ldots, n_1$). Consider conjugation by the element (123) (reminder that by normality $(123)\sigma(321) \in N$):

$$\begin{aligned}(123)\sigma(321) &= (123)\Gamma_1 \cdots \Gamma_m(321) = (123)\Gamma_1(321)\Gamma_2 \cdots \Gamma_m \\ &= (123)\Gamma_1(321)\Gamma_1^{-1} \cdot \Gamma_1\Gamma_2 \cdots \Gamma_m = (123)(123 \ldots n_1)(321)(n_1 \ldots 321)\sigma \\ &= (23n_1)\sigma\end{aligned}$$

and so we get that $(23n_1) = [(123)\sigma(321)]\sigma^{-1} \in N$.

Case 2a: all the Γ_i have length ≤ 3, and ≥ 2 of them have length $= 3$. Again by relabelling, let these 2 particular 3-cycles be $\Gamma_1 = (123)$ and $\Gamma_2 = (456)$, and consider conjugation by the element (124) — yes unfortunately (123) does not work (again by normality $\implies (124)\sigma(421) \in N$):

$$\begin{aligned}(124)\sigma(421) &= (124)\Gamma_1 \cdots \Gamma_m(421) = (124)\Gamma_1\Gamma_2(421)\Gamma_3 \cdots \Gamma_m \\ &= (124)\Gamma_1\Gamma_2(421)\Gamma_2^{-1}\Gamma_1^{-1} \cdot \Gamma_1\Gamma_2 \cdots \Gamma_m = (124)(123)(456)(421)(654)(321)\sigma \\ &= (12436)\sigma\end{aligned}$$

and so we get that $(12436) = [(124)\sigma(421)]\sigma^{-1} \in N$. Use Case 1 on this 5-cycle to get a 3-cycle.

Case 2b: exactly one Γ_i has length 3, and the rest have length 2 (length 2 means it is its own inverse). So again relabelling, we have $\Gamma_1 = (123)$, and then $\sigma^2 = \Gamma_1 \cdots \Gamma_m \cdot \Gamma_1 \cdots \Gamma_m = \Gamma_1 \cdots \Gamma_m \cdot \Gamma_m \cdots \Gamma_1 = \Gamma_1^2 = (132)$.

Case 3: all the Γ_i have length 2 (and there is an even number of them). Even number and non-trivial guarantees us that there are two, so again relabelling we have $\Gamma_1 = (12)$ and $\Gamma_2 = (34)$. Then, like above, we have that

$$\begin{aligned}(123)\sigma(321) &= (123)\Gamma_1 \cdots \Gamma_m(321) = (123)\Gamma_1\Gamma_2(321)\Gamma_3 \cdots \Gamma_m \\ &= (123)\Gamma_1\Gamma_2(321)\Gamma_2^{-1}\Gamma_1^{-1} \cdot \Gamma_1\Gamma_2 \cdots \Gamma_m = (123)(12)(34)(321)(34)(12)\sigma \\ &= (14)(23)\sigma \implies (14)(23) \in N\end{aligned}$$

Then, $(14)(23)\big[(145) \cdot (14)(23) \cdot (541)\big] \in N$, where

$$(14)(23)\big[(135) \cdot (14)(23) \cdot (531)\big] = (14)(145)(14)(541) = (145)$$

and so here $(145) \in N$. Thus in all cases, N has a 3-cycle and we are done. □

5. Sylow subgroups of S_{p^n}

In this section you'll give an alternative proof of the first Sylow theorem. So you are NOT allowed to assume any of them!
Let $\nu(n)$ denote the maximal power of p dividing $(p^n)!$; that is, $p^{\nu(n)} | (p^n)!$ but $p^{\nu(n)+1} \not| (p^n)!$.

Lemma 5.1. $\nu(n) = 1 + p + \ldots + p^{n-1}$.

Proof. There's 1 p^n, p many p^{n-1}'s, p^2 many p^{n-2}'s, and so on until p^{n-1} many p's. Adding all this up, we have a total of $1 + p + p^2 + \ldots + p^{n-1}$ many p's in the prime factorization of $(p^n)!$, which is to say that the maximal power $\nu(n)$ of p dividing $(p^n)!$ is exactly $1 + p + \ldots + p^{n-1}$. □

Proposition 5.2. *The symmetric group S_{p^n} has a Sylow p-subgroup.*

Proof. Hint: by induction. For the induction step $n - 1 \mapsto n$, subdivide S_{p^n} into p equal parts. Consider the permutation σ of order p defined as a product of p^{n-1} disjoint cycles as follows:

$$\sigma = (1, p^{n-1}+1, \ldots, (p-1)p^{n-1}+1) \ldots (j, p^{n-1}+j, \ldots, (p-1)p^{n-1}+j) \ldots (p^{n-1}, 2p^{n-1}, \ldots, p^n)$$

Now using a Sylow p-subgroup of $S_{p^{n-1}}$ and the permutation σ, construct a Sylow p-subgroup for S_{p^n}.

Base case ($n = 0$): $S_{p^0} = S_1 = \{1\}$ does in fact have a Sylow p-subgroup $\{1\}$ (of order p^0). We now assume that the claim holds for $n - 1$, and now we want to prove it for S_{p^n}. As the hint suggests, we split the SET S_{p^n} into p equal SETS $T_1, \ldots, T_p$ where $T_i = \{(i-1)p^{n-1}+1, \ldots, ip^{n-1}\}$. Let $\mathcal{T}_i = \{\tau \in S_{p^n} : \tau \text{ only permutes elements in } T_i\}$. $\mathcal{T}_i$ has $|T_i|! = (p^{n-1})!$ elements, and it is clear that $\mathcal{T}_i \simeq S_{p^{n-1}}$. By the induction hypothesis, we now have R_i ($i \in \{1, \ldots, p\}$) s.t. R_i is a Sylow p-subgroup in $\mathcal{T}_i$ with order $p^{\nu(n-1)}$. Now consider $M = \prod_{i=1}^p R_i = \{\tau_1 \cdots \tau_p : \tau_i \in \mathcal{T}_i\}$; because the T_i are disjoint, τ_i and τ_j for $i \neq j$ are disjoint and so they commute with each other. Moreover, disjointness means that $|M| = (p^{\nu(n-1)})^p = p^{\nu(n)-1}$.

M is a subgroup: closure is because $(\tau_1\tau_2 \cdots \tau_p)(\tau_1'\tau_2 \cdots \tau_p') = (\tau_1\tau_1') \cdots (\tau_p\tau_p')$ (using the reordering property of $\tau_i\tau_j$, $i \neq j$); identity is because $() = \text{id} \in R_i$ for all $i \in \{1, \ldots, p\}$ because the R_i are subgroups; associativity is inherited, and inverse because $(\tau_1\tau_2 \cdots \tau_p)^{-1} = \tau_1^{-1} \cdots \tau_p^{-1}$ (reordering property is crucial here), where $\tau_i^{-1} \in R_i$ because R_i is a subgroup. Also, to give some intuition to M, M contains permutations that send T_i to a scrambled T_i, i.e. think of each T_i as a box, and think of M as going through all p boxes and stirring the contents.

Unfornately, M has order $p^{\nu(n)-1}$, so we are off by a factor of p. However, there is a quick fix: for each permutation α in M that sends T_i to a scrambled T_i, consider the p permutations (one for each $k \in \{0, \dots, p-1\}$) that send T_i to a scrambled T_{i+k} (indices taken $\operatorname{mod} p$). Here's where σ comes in — intuively we can say that it sends T_i to T_{i+1}, and so the "quick fix" in more mathematical notion is just $\sigma^k \alpha$, where σ^k sends T_i to T_{i+k}, and α sends T_{i+k} to a scrambled T_{i+k}.

Now we do this formally: consider the set $Q = \{\sigma^k M : k \in \{0, \dots, p-1\}\}$. All elements are distinct because $\sigma^{k_1}\alpha_1 = \sigma^{k_2}\alpha_2 \implies \sigma^{k_1-k_2} = \alpha_2\alpha_1^{-1}$, but the RHS is in M, so k_1 must $= k_2 \implies \sigma^{k_1} = \sigma^{k_2}$ and $\alpha_1 = \alpha_2$ (because σ is not in M, as σ moves elements between "boxes", and elements of M do not). This shows that $|Q| = p \cdot |M| = p^{\nu(n)-1+1} = p^{\nu(n)}$.

Lastly Q is a subgroup: closure because we have that $\sigma^{k-1}\mathcal{T}_1\sigma^{k-1} = \mathcal{T}_k \implies \sigma^{k-1}M\sigma^{k-1} = M$, i.e. for any $k \in \{0, \dots, p-1\}$, $\sigma^k M = M\sigma^k$, and so $\sigma^{k_1}\alpha_1\sigma^{k_2}\alpha_2 = \sigma^{k_1}\sigma^{k_2}\alpha_1'\alpha_2$; identity because $\sigma^0() = \text{id}$; associativity is inherited; and inverse because $(\sigma^k\alpha)^{-1} = \alpha^{-1}\sigma^{-k} = \sigma^{-k}(\alpha^{-1})'$ (by $\sigma^{k-1}M = \sigma^{k-1}M$). Thus, Q is a subgroup of order $p^{\nu(n)}$, i.e. it is a Sylow p-subgroup. □

Definition 5.3. Let G be a group, and H, K be subgroups of G. For an element $x \in G$, the set

$$HxK := \{hxk \,|\, h \in H, k \in K\}$$

is called a *double coset* of H, K in G.

The next three statements constitute $\boxed{\text{Problem 10}}$.

Lemma 5.4. *Suppose H, K are finite subgroups of G. Then for any $x \in G$, $|HxK| = \frac{|H||K|}{|H \cap xKx^{-1}|}$.*

Proof. Proposition 3.13 in Dummit & Foote gives that for any subgroups H, K in G, $|HK| = \frac{|H||K|}{|H \cap K|}$. Notice also that $HxKx^{-1}$ and HxK are in bijection (consider the map $g \mapsto gx^{-1} : HxK \to HxKx^{-1}$ and its inverse $g \mapsto gx : HxKx^{-1} \to HxK$), as well as xKx^{-1} and K (consider the map $g \mapsto xgx^{-1} : K \to xKx^{-1}$ and its inverse $g \mapsto x^{-1}gx : xKx^{-1} \to K$). Thus, we have that

$$|HxK| = |HxKx^{-1}| = \frac{|H||xKx^{-1}|}{|H \cap xKx^{-1}|} = \frac{|H||K|}{|H \cap xKx^{-1}|}$$

□

Proposition 5.5. *Let $H < G$ be finite groups, and suppose that G has a Sylow subgroup Q. Then H has a Sylow subgroup P. Moreover, $P = H \cap xQx^{-1}$ for some $x \in G$.*

Proof. Hint: Consider double cosets HxQ, and let p^n be the maximal power of p dividing $|H|$ (so that the expected order of P is p^n). Using the formula for the size of double cosets in the Lemma above and the fact that G is a union of disjoint double cosets, show (by contradiction) that at least one intersection $H \cap xQx^{-1}$ must have the maximal possible size p^n.

Denote $|G| = p^k m$ (where k is maximal, i.e. $\gcd(p, m) = 1$), and similarly $|H| = p^n\ell$. Then, $|Q| = p^k$, and because H, xQx^{-1}, $H \cap xQx^{-1}$ is a subgroup of xQx^{-1}, and so by Lagrange it must have order p^a for some $a \le \min\{k, n\} = n$. We want to show that $a = n$ for some x. Suppose not. Then, all $a < n$, and by Lemma 5.4, we have that all $|HxQ| = \frac{p^n\ell p^k}{p^a} = \ell p^{k+n-a}$ are divisible by p^{k+1}. Then, because there is some set $X \subseteq G$ s.t. HxQ are all disjoint, and $G = \bigcup_{x \in X} HxQ$, we have that $|G| = \sum_{x \in X} |HxQ|$. But since all the $|HxQ|$ are divisible by p^{k+1}, the sum on right is too, and therefore $|G|$ must be too. But this is impossible, because k was the maximal power of p in $|G|$. Thus, we must have some $H \cap xQx^{-1}$ (a subgroup) with order p^n, and so this is our Sylow p-subgroup of H. □

Now, the first Sylow theorem is an easy consequence of what you already proved.

Theorem 5.6. *Any finite group whose order is divisible by p has a Sylow p-subgroup.*

Proof. By Cayley's theorem, our finite group G of order n is isomorphic to some subgroup of S_n, which in turn is isomorphic to the subgroup of S_{p^m} consisting of permutations only permuting the elements $\{1, \ldots, n\}$ (we only require $n \leq p^m$). We know by Proposition 5.2 that S_{p^m} has a Sylow p-subgroup, and by Proposition 5.5, we have that $G < S_{p^m}$ also has a Sylow p-subgroup. □