M3P10: GROUP THEORY

LECTURES BY DR. JOHN BRITNELL; NOTES BY ALEKSANDER HORAWA

These are notes from the course M3P10: Group Theory taught by Dr. John Britnell, in Fall 2015 at Imperial College London. They were LaTeX'd by Aleksander Horawa.

This version is from January 6, 2017. Please check if a new version is available at my website https://sites.google.com/site/aleksanderhorawa/. If you find a mistake, please let me know at aleksander.horawa13@ic.ac.uk.

Website: http://wwwf.imperial.ac.uk/~jbritnel/Teaching/index.html

Contents

Introduction 1
1. Quotient Groups 4
2. Group Actions 9
3. Sylow's Theorems 15
4. Automorphism Groups and Semidirect Products 18
5. Composition Series 22
6. The Lower Central Series and nilpotent groups 29
7. More on actions 33
Examples of Sylow subgroups 40
Appendix A. The alternating group A_n is simple for $n \geq 5$ 45

Introduction

We first review the basic notions of group theory.

Definition. A *group* is a set G equipped with a binary operation $*\colon G \times G \to G$ such that:

(*associativity*) $(x * y) * z = x * (y * z)$ for all $x, y, z \in G$,
(*identity*) there exists $e \in G$, an *identity element*, such that $x * e = e * x = x$ for all $x \in G$,
(*inverses*) for all $x \in G$, there exists $y \in G$, an *inverse of* x, such that $x * y = y * x = e$.

The identity element e is unique and the inverse of each element is unique. We usually use multiplicative notation for groups, i.e. xy for $x * y$, x^{-1} for the inverse of x, and 1 for e.

We have *right cancellation*: $xz = yz$ implies that $x = y$, and *left cancellation*: $xy = xz$ implies that $y = z$.

The group G is *abelian* if $x * y = y * x$ for all $x, y \in G$. We often write abelian groups additively: $x + y$ for $x * y$, $-x$ for the inverse of x, and 0 for e.

A *subgroup* H of G is a non-empty subset which is closed under $*$ and taking inverses. We then write $H \leq G$. Every group G has subgroups G itself and $\{e\}$, the trivial subgroup. Other subgroups of G are called *non-trivial proper subgroups*.

We write x^k for $\underbrace{xx\ldots x}_{k \text{ times}}$ (or kx for $\underbrace{x + x + \cdots + x}_{k \text{ times}}$ if we are using additive notation). We write $\langle x \rangle$ for $\{x^k \;:\; k \in \mathbb{Z}\}$, the *cyclic subgroup generated by* x. More generally, if $x_1, \ldots, x_k \in G$, we define $\langle x_1, \ldots, x_k \rangle$ to be the subgroup generated by $x_1, \ldots, x_k$, the smallest subgroup of G which contains $x_1, \ldots x_k$. More formally,

$$\langle x_1, \ldots, x_k \rangle = \bigcap H$$

where the intersection is over all subgroups H of G containing $x_1, \ldots, x_k$. Alternatively, take any *word* in $x_1, \ldots, x_k, x_1^{-1}, \ldots, x_k^{-1}$, e.g. $x_1^2 x_2^{-3} x_1^{-1} x_2 x_2^{-1}$. This represents some group element. It is not hard to show that the subset of elements of G which we can represent in this way is the subgroup $\langle x_1, \ldots, x_k \rangle$.

If $X = \{x_1, \ldots, x_k\}$, we can write $\langle X \rangle$ for $\langle x_1, \ldots, x_k \rangle$. This also works if X is infinite.

Remarks.

(1) If $H \leq G$, then $\langle H \rangle = H$.
(2) By convention, $\langle \emptyset \rangle = \{e\}$. (This is clear from the definition as an intersection.)
(3) If $G = \langle x_1, \ldots, x_k \rangle$, we say that $\{x_1, \ldots, x_k\}$ is a *generating set*. We will say that G is k*-generated* if it has a generating set of order k. So 0-generated is equivalent to being trivial, 1-generated is equivalent to being cyclic. The 2-generated groups are a massive family.

Theorem (Lagrange's Theorem)**.** *If G is finite and $H \leq G$, then $|H|$ divides $|G|$.*

The proof uses the idea of cosets. A *left coset* is $gH = \{gh : h \in H\}$. We write $|G : H|$ for the *index* of H in G (the number of cosets), and we have that $|G| = |H|\,|G : H|$.

A subgroup $H \leq G$ is *normal* (and we write $H \trianglelefteq G$) if one of the following equivalent conditions holds:

(1) Every left coset is a right coset.
(2) Every right coset is a left coset.
(3) $Hg = gH$ for all $g \in G$.
(4) $H = gHg^{-1}$ for all $g \in G$.

If $H \trianglelefteq G$, then the set of cosets of H in G inherits a group structure from G:

$$(xH)(yH) = (xyH).$$

This is the *quotient group* G/H.

A homomorphism from a group G to a group H is a function $\theta\colon G \to H$ such that $\theta(g_1g_2) = \theta(g_1)\theta(g_2)$ for all $g_1, g_2 \in G$. Then the *image* of θ is $\mathrm{Im}(\theta) = \{\theta(g) : g \in G\} \subseteq H$ and the *kernel* of θ is $\mathrm{Ker}(\theta) = \{g \in G : \theta(g) = e\} \trianglelefteq G$. If $\mathrm{Im}\,\theta = H$ and $\mathrm{Ker}\,\theta = \{e\}$, then θ is an *isomorphism.*

Theorem (First Isomorphism Theorem)**.** *If $\theta\colon G \to H$ is a surjective homomorphism with kernel K, then $G/K \cong H$ with the isomorphism given by $\tilde{\theta}\colon G/K \to H$ given by $\tilde{\theta}(gK) = \theta(g)$.*

The map $G \to G/N$ given by $g \mapsto gN$ is called the *canonical map.* It is a surjective homomorphism with image G/N and kernel N.

If A and B are groups, then the *direct product* $A \times B$ is the set of pairs $\{(a,b) : a \in A, b \in B\}$ with the operation $(a_1,b_1)(a_2,b_2) = (a_1a_2, b_1b_2)$.

Facts.

- $|A \times B| = |A||B|$
- $A \times \{e_B\}$ *is a normal subgroup of $A \times B$, isomorphic to A*
- $\{e_A\} \times B$ *is a normal subgroup of $A \times B$, isomorphic to B*

More generally, if we have groups $\{A_i : i \in I\}$, we can form the direct product

$$\prod_{i \in I} A_i.$$

(If the indexing set is infinite, there are two possible products, but we will not go into this—the course is focused on finite groups, so products will be finite.)

Theorem (Characterization of finite abelian groups)**.** *Any finite abelian group is a direct product of cyclic groups. Moreover, for any finite abelian group A, there exists a unique sequence $q_1, \ldots, q_k \in \mathbb{N}$ such that q_{i+1} divides q_i and*

$$A \cong \prod_i C_{q_i}.$$

Examples (Groups)**.** *Cyclic groups*: C_n (or $\mathbb{Z}_n$), C_∞ (or $\mathbb{Z}$).

Dihedral groups: A group is *dihedral* if it is generated by elements a and b such that $b^2 = e$ and $a^-1 = bab$. For any even order $2n$, there is a unique dihedral group D_{2n} (the group of symmetries of an n-gon). For infinite order, there is a unique infinite dihedral group D_∞ ($a\colon \mathbb{Z} \to \mathbb{Z}$, $a(n) = n+1$ and $b\colon \mathbb{Z} \to \mathbb{Z}$, $b(n) = -n$).

Symmetric groups: S_n is the group of permutations of $\{1, \ldots, n\}$; for any set X, $\mathrm{Sym}(X)$ is the group of permutations of X. A permutation of a finite set has a signature $+$ or $-$, i.e. there is a homomorphism $\mathrm{sgn}\colon S_n \to \{1, -1\}$. If g is a transposition, then $\mathrm{sgn}(g) = -1$.

Alternating groups: $A_n = \mathrm{Ker}(\mathrm{sgn})$. An element of A_n is called *even.* Note that a permutation is even if it has an even number of cycles of even length.

Vector spaces are groups under $+$.

General linear groups: If F is a field, then $\mathrm{GL}_n(F)$ is the set of invertible $n \times n$ matrices with entries from F. If F is a finite field with p^r elements, we write $\mathrm{GL}_n(p^r) = \mathrm{GL}_n(F)$. We have a homomorphism $\det\colon \mathrm{GL}_n(F) \to F^\times$.

Special linear groups: $\mathrm{SL}_n(F) = \mathrm{Ker}(\det)$.

1. Quotient Groups

We will look at subgroups of G/K and relate them to subgroups of G.

Suppose $\theta\colon G \to H$ is a homomorphism. For a subset $S \subseteq G$, we will write

$$\theta(S) = \{\theta(s) : s \in S\} \subseteq H,$$

and for a subset $T \subseteq H$, we will write

$$\theta^{-1}(T) = \{g \in G : \theta(g) \in T\}.$$

For $S, T \subseteq G$, write $ST = \{st : s \in S,\ t \in T\}$.

Proposition 1. *Let $\theta\colon G \to H$ is a surjective*[1] *homomorphism with kernel K. Then:*

(1) $\theta(L) \leq H$ *for all* $L \leq G$,
(2) $K \leq \theta^{-1}(X) \leq G$ *for all* $X \leq H$,
(3) *if* $K \leq L \leq G$, *then* $K \trianglelefteq L$ *and* $L/K \cong \theta(L)$,
(4) $\theta(\theta^{-1}(X)) = X$ *for all* $X \leq H$,
(5) $\theta^{-1}(\theta(L)) = KL \leq G$ *for all* $L \leq G$; *in particular, if* $K \leq L$, *then* $\theta^{-1}(\theta(L)) = L$.

Proof. (1) Let $\theta_{|L}$ be the restriction of θ to L. Then $\theta_{|L}\colon L \to H$ is a homomorphism with image $\theta(L)$, and hence $\theta(L) \leq H$.

(2) If $k \in K$, then $\theta(k) = e_H \in X$, so $k \in \theta^{-1}(X)$. Hence $K \subseteq \theta^{-1}(X)$. We check that it is a subgroup. If $g_1, g_2 \in \theta^{-1}(X)$, then $\theta(g_1) \in X$ and $\theta(g_2) \in X$, so $\theta(g_1 g_2) = \theta(g_1)\theta(g_2) \in X$, so $g_1 g_2 \in \theta^{-1}(X)$. If $g \in \theta^{-1}(X)$, then $\theta(g^{-1}) = \theta(g)^{-1} \in X$. Hence $K \leq \theta^{-1}(X) \leq H$.

(3) If $K \trianglelefteq G$, then $gK = Kg$ for all $g \in G$. In particular, $gK = Kg$ for all $g \in L$, so if $K \leq L$, then $K \trianglelefteq L$. To get $L/K \cong \theta(L)$ we apply the First Isomorphism Theorem to $\theta_{|L}$.

(4) Let $x \in X$. By definition, $\theta^{-1}(x) = \{g \in G : \theta(g) = x\}$ and hence $\theta(\theta^{-1}(x)) \subseteq \{x\}$. Hence $\theta(\theta^{-1}(X)) \subseteq X$. Since θ is surjective, $\theta^{-1}(x)$ is non-empty for all $x \in X$, so $x \in \theta(\theta^{-1}(X))$. Hence $\theta(\theta^{-1}(X)) = X$.[2]

(5) Suppose $g \in \theta^{-1}(\theta(L))$. Then $\theta(g) \in \theta(L)$, so for some $l \in L$ we have that $\theta(g) = \theta(l)$. Now $\theta(gl^{-1}) = \theta(g)\theta(l)^{-1} = e_H$ and hence $gl^{-1} \in K$. Thus $g = (gl^{-1})l \in KL$. Hence $\theta^{-1}(\theta(L)) \subseteq KL$.
Conversely, if $k \in K$, $l \in L$, then $\theta(kl) = \theta(k)\theta(l) = e_H\theta(l) \in \theta(L)$. Hence $KL \subseteq \theta^{-1}(\theta(L))$, and we have equality. It follows from (1) and (2) that KL is a subgroup.

For the *in particular* clause, note that if $K \leq L$, then $L = \{e_G\}L \subseteq KL \subseteq L$, so $KL = L$. $\square$

[1] If θ was not surjective, we could simply replace H by $\mathrm{Im}(\theta)$ to get a surjective homomorphism. Therefore, we are not losing any generality by assuming surjectivity.

[2] We really have to assume surjectivity for this argument to work.

Example. Let $G = C_{12}$. We can represent the subgroups of G in an *order diagram*[3] as follows

$$\begin{tikzcd}[column sep=small]
 & & C_{12} \arrow[dl, dash] \arrow[dr, dash] & \\
 & C_6 \arrow[dl, dash] \arrow[dr, dash] & & C_4 \arrow[dl, dash] \\
C_3 \arrow[dr, dash] & & C_2 \arrow[dl, dash] & \\
 & 1 & &
\end{tikzcd}$$

The quotient Q of C_{12} by C_3 is isomorphic to C_4. The subgroup diagram for Q is

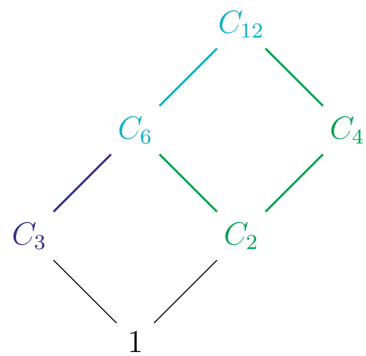
$$\begin{tikzcd}
Q \cong C_4 \arrow[d, dash] \\
C_2 \arrow[d, dash] \\
1
\end{tikzcd}$$

Now, take the quotient R of C_{12} by C_2, isomorphic to C_6. It has subgroups

$$\begin{tikzcd}[column sep=small]
 & R \cong C_6 \arrow[dl, dash] \arrow[dr, dash] & \\
C_3 \arrow[dr, dash] & & C_2 \arrow[dl, dash] \\
 & 1 &
\end{tikzcd}$$

In the following diagram, blue represents the subgroups of C_{12} containing C_3 and green represents the subgroups of C_{12} containing C_2 (with turquise representing the parts of the diagram contained in both).

[3]The vertices are groups and the edges represent the subgroup relation, i.e. we write G above H and draw an edge between G and H to mean that $G \geq H$.



In both cases, the diagram for the quotient is the same as the part of the diagram for C_{12} with the subgroups containing the kernel. This idea is formalized in the following theorem.

Theorem 2 (Correspondence Theorem)**.** *Let $\theta\colon G \to H$ be a surjective homomorphism with kernel K. Write $\mathrm{sub}_K(G)$ for the set of subgroups of G containing K and $\mathrm{sub}(H)$ for the set of subgroups of H. The map*

$$\hat{\theta}\colon \mathrm{sub}_K(G) \to \mathrm{sub}(H)$$

defined by $\hat{\theta}(L) = \theta(L)$ is a bijection, with the following properites:

1. *if $L, M \in \mathrm{sub}_K(G)$ then $L \leq M$ if and only if $\hat{\theta}(L) \leq \hat{\theta}(M)$,*
2. *if $L \in \mathrm{sub}_K(G)$ then $L \trianglelefteq G$ if and only if $\hat{\theta}(L) \trianglelefteq H$.*

Proof. Certainly, $\hat{\theta}$ is indeed a map $\mathrm{sub}_K(G) \to \mathrm{sub}(H)$ by Proposition 1 (1).

For injectivity, suppose $\hat{\theta}(L) = \theta(\hat{M})$. Then $\theta(L) = \theta(M)$, so $\theta^{-1}(\theta(L)) = \theta^{-1}(\theta(M))$. But $K \leq L$ and $K \leq M$, so $L = \theta^{-1}(\theta(L)) = \theta^{-1}(\theta(M)) = M$ by Proposition 1 (5).

For surjectivity, let $X \leq H$. Then $\theta^{-1}(X) \in \mathrm{sub}_K(G)$ by Proposition 1 (2), and hence

$$\hat{\theta}(\theta^{-1}(X)) = \theta(\theta^{-1}(X)) = X$$

by Proposition 1 (4).

For (1), suppose $L, M \in \mathrm{sub}_K(G)$ with $L \leq M$. It is clear that $\theta(L) \subseteq \theta(M)$ and both are subgroups of H, so $\hat{\theta}(L) \leq \hat{\theta}(M)$. Suppose conversely that $\hat{\theta}(L) \leq \hat{\theta}(M)$. Then clearly $\theta^{-1}(\theta(L)) \subseteq \theta^{-1}(\theta(M))$. But $\theta^{-1}(\theta(L)) = L$ and $\theta^{-1}(\theta(M)) = M$, so $L \leq M$.

For (2), let $L \in \mathrm{sub}_K(G)$. First, suppose that $L \trianglelefteq G$. Then $gL = Lg$ for all $g \in G$. Now take any $h \in H$ and consider $h\hat{\theta}(L)$. By surjectivity of θ, there exists $g \in G$ such that $\theta(g) = h$. Now,

$$h\hat{\theta}(L) = \theta(g)\theta(L) = \theta(gL) = \theta(Lg) = \theta(L)\theta(g) = \hat{\theta}(L)h.$$

Hence $\hat{\theta}(L) \trianglelefteq H$.
Conversely, let $X = \theta(L)$ and suppose that $X \trianglelefteq H$. Let ϱ be the cannonical map $\varrho\colon H \to H/X$. Now, the kernel of the composition homomorphism $\varrho \circ \theta\colon G \to H/X$ is $\theta^{-1}(X)$, and so $\theta^{-1}(X) \trianglelefteq G$. Hence $\theta^{-1}(X) = L$, so $L \trianglelefteq G$. □

Example. Take S_4 has a normal subgroup $V_4 = \{e, (12)(34), (13)(24), (14)(23)\}$. What is S_4/V_4? It has order 6, and it cannot be cyclic, since S_4 has no elements of order 6. Therefore, we must have $S_4/V_4 \cong S_3$. The subgroups of S_3 are

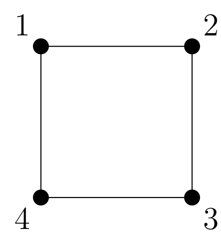
$$\begin{tikzcd}[column sep=small]
 & & S_3 \arrow[dll, dash] \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
A_3 \arrow[drr, dash] & \langle(13)\rangle \arrow[dr, dash] & \langle(12)\rangle \arrow[d, dash] & \langle(23)\rangle \arrow[dl, dash] \\
 & & 1 &
\end{tikzcd}$$

So the correspondence theorem tells us that the part of the subgroup diagram for S_4 above V_4 looks the same.

$$\begin{tikzcd}[column sep=small]
 & S_4 \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] \arrow[drr, dash] & & \\
A_4 \arrow[dr, dash] & G_1 \arrow[d, dash] & G_2 \arrow[dl, dash] & G_3 \arrow[dll, dash] \\
 & V_4 & &
\end{tikzcd}$$

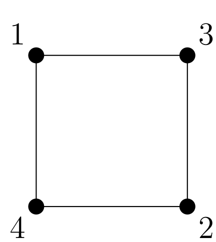
with G_1, G_2, G_3 subgroups of order 8. What are they? The are actually dihedral groups.

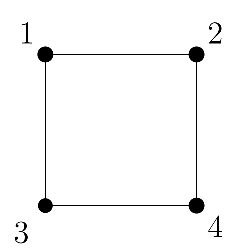
The dihedral group permutes the vertices 1, 2, 3, 4 of the square.



We can hence write the elements as permutations. This gives rotations $e, (1234), (13)(24), (1432)$ and reflections $(14)(23)$, $(12)(34)$, $(13)(24)$.

Relabelling the vertices gives different copies of D_8 in S_4:





Theorem 3 (Second Isomorphism Theorem)**.** *Let $K, L \trianglelefteq G$ with $K \leq L$. Then*

$$\frac{G}{L} \cong \frac{G/K}{L/K}.$$

Proof. The idea of the proof is to apply the First Isomorphism Theorem to the map $\theta\colon \frac{G}{K} \to \frac{G}{L}$, defined by
$$\theta(gK) = gL.$$
We first check it is well-defined. Suppose that $g_1K = g_2K$. Then $g_1g_2^{-1} \in K$, so $g_1g_2^{-1} \in L$ since $K \leq L$. Thus $g_1L = g_2L$, as requested.

Moreover, θ is a homomorphism:
$$\theta((g_1K)(g_2K)) = \theta(g_1g_2K) = g_1g_2L = (g_1L)(g_2L) = \theta(g_1K)\theta(g_2K).$$

Finally, θ is surjective: for $gL \in G/L$, we have $gL = \theta(gK)$.

To find the kernel of θm note that $\theta(gK) = e_{G/L} = L$ if and only if $gL = L$, i.e. $g \in L$, which is equivalent to $gK \in L/K$. Hence $\operatorname{Ker}\theta = L/K$.

By the First Isomorphism Theorem, we obtain
$$\frac{G/K}{L/K} \cong \frac{G}{L},$$
as requested. □

Proposition 4. *Let A, B be subgroups of G. Then:*

(1) *If A, B are finite, then $|AB| = \frac{|A||B|}{|A\cap B|}$.*
(2) *The set AB is a subgroup of G if and only if $AB = BA$. In particular, if $A \trianglelefteq G$, then AB is a subgroup of G.*

Proof. Homework 1, Question 1. □

Proposition 5. *Let $K \trianglelefteq G$ and $L \leq G$. Then*

(1) *$KL \leq G$,*
(2) *if $\theta\colon G \to H$ is a homomorphism with kernel K, then $\theta(L) = \theta(KL)$.*

Proof. Note that (1) follows from Proposition 4 (2).

For (2), we have
$$\theta(KL) = \{\theta(kl) : k \in K, l \in L\} = \{\theta(k)\theta(l) : k \in K, l \in L\} = \{\theta(L) : l \in L\} = \theta(L),$$
since $K = \operatorname{Ker}(\theta)$ so $\theta(k) = e$. □

Theorem 6 (Third Isomorphism Theorem)**.** *Let $K \trianglelefteq G$ and $L \leq G$. Then $K \cap L \trianglelefteq L$ and*
$$\frac{KL}{K} \cong \frac{L}{K\cap L}.$$

Proof. Let $\theta\colon G \to G/K$ be the canonical map and $\theta_{|L}$ be its restriction to L. Clearly, $\operatorname{Ker}(\theta_{|L}) = K \cap L$. We see that $\operatorname{Im}(\theta_{|L}) = \{lK : l \in L\} = LK/K = KL/K$ and hence the First Isomorphism Theorem yields
$$\frac{L}{K\cap L} \cong \frac{KL}{K},$$
as requested. □

We can represent the Third Isomorphism Theorem 6 on a subgroup diagram as follows, with the blue lines corresponding to the quotients that are equal.

$$\begin{tikzcd}[arrows=dash]
 & G \arrow[d, dotted] \arrow[ddl, dotted] \arrow[ddr, dotted] & \\
 & KL \arrow[dl] \arrow[dr, blue] & \\
L \arrow[dr, blue] & & K \arrow[dl] \\
 & K \cap L &
\end{tikzcd}$$

Note that KL is the smallest subgroup of G which contains K and L, and $K \cap L$ is the largest subgroups contained in K and L.

Remark. Compare: for integers m, n we have that

$$\begin{tikzcd}[arrows=dash]
 & \operatorname{lcm}(m,n) \arrow[dl] \arrow[dr, blue] & \\
m \arrow[dr, blue] & & n \arrow[dl] \\
 & \gcd(m,n) &
\end{tikzcd}$$

with the lines corresponding to divisibility, and indeed

$$\frac{\operatorname{lcm}(m,n)}{n} = \frac{m}{\gcd(m,n)}.$$

2. Group Actions

Definition. Let G be a group and Ω a set. A *left action* of G on Ω is a map $\psi\colon G \times \Omega \to \Omega$ such that

(1) $\psi(e, x) = x$ for all $x \in \Omega$,
(2) $\psi(gh, x) = \psi(g, \psi(h, x))$ for all $g, h \in G$, $x \in \Omega$.

We usually write gx for $\psi(g, x)$, except when there is a good reason not to. By (2), we have that $(gh)x = g(hx)$, so we can just write ghx for this element of Ω.

Definition. Let G be a group and Ω a set. A *right action* of G on Ω is a map $\psi\colon \Omega \times G \to \Omega$ such that

(1) $\psi(x, e) = x$ for all $x \in \Omega$,
(2) $\psi(x, gh) = \psi(\psi(x, g), h)$ for all $g, h \in G$, $x \in \Omega$.

Note that this is not quite the same, since in the product gh, the elements g and h are applied to x in a different order. However, if $\psi\colon G \times \Omega \to \Omega$ is a left-action, then we can define the associated right action $\psi^{\mathrm{op}}\colon \Omega \times G \to \Omega$ by

$$\psi^{\mathrm{op}}(x, g) = \psi(g^{-1}, x).$$

It is clear that this actually defines a right action.

The elements of Ω are often called *points*.

Examples. The dihedral group D_{2n} acts *naturally* on the vertices of a regular n-gon (or on edges or interior points, etc).

The symmetric group S_n acts *naturally* on the set $[n] := \{1, \ldots, n\}$. Is also acts on the pairs $(i, j) \in [n]^2$ by

$$\psi(g, (i, j)) = (gi, gj).$$

The general linear group $\mathrm{GL}_n(F)$ acts *naturally* on the vector space F^n. For the left action, we notationally take F^n to consist of column vectors. There is also a right action, for which we take F^n to consist of row vectors.

Definition. Let G act on Ω and $x \in \Omega$. The *orbit* of x is the set

$$\mathrm{Orb}_G(x) = \{gx : g \in G\} \subseteq \Omega.$$

The *stabilizer* of x is the subgroup

$$\mathrm{Stab}_G(x) = \{g \in G : gx = x\} \leq G.$$

Proposition 7. *The stabilizer, as defined above, is indeed a subgroup of G.*

Proof. First, $\varphi(e, x) = x$ by (1), so $e \in \mathrm{Stab}_G(x)$. Next, for $g, h \in \mathrm{Stab}_G(x)$, we have

$$\begin{aligned} \psi(gh, x) &= \psi(g, \psi(h, x)) && \text{by axiom (2)} \\ &= \psi(g, x) && \text{since } h \in \mathrm{Stab}_G(x) \\ &= x && \text{since } g \in \mathrm{Stab}_G(x) \end{aligned}$$

so $gh \in \mathrm{Stab}_G(x)$. Finally, for $g \in \mathrm{Stab}_G(x)$, we have

$$\begin{aligned} x &= \psi(e, x) && \text{by axiom (1)} \\ &= \psi(g^{-1}g, x) \\ &= \psi(g^{-1}, \psi(g, x)) && \text{by axiom (2)} \\ &= \psi(g^{-1}, x) && \text{since } g \in \mathrm{Stab}_G(x) \end{aligned}$$

so $g^{-1} \in \mathrm{Stab}_G(x)$. Hence $\mathrm{Stab}_G(x) \leq G$. □

Examples.

(1) Take D_{2n} acting on vertices of an n-gon and let x be any vertex. Then

$$\mathrm{Orb}_{D_{2n}}(x) = \{\text{all vertices of the } n\text{-gon}\}$$

since we can rotate x to any other vertex. Moreover,

$$\mathrm{Stab}_{D_{2n}}(x) = \{I, T_x\}, \text{ where } T_x \text{ is the reflection through the axis passing through x.}$$

(2) Take S_n acting on $[n]$ and $x \in [n]$. Then
$$\operatorname{Orb}_{S_n}(x) = [n],$$
since if $y \neq x$, then $(xy) \in S_n$. Moreover,
$$\operatorname{Stab}_{S_n}(x) \cong \operatorname{Sym}([n] \setminus \{x\}) \cong S_{n-1}.$$

Theorem 8 (Orbit-Stabilizer Theorem)**.** *Let G be a finite group acting on a set Ω and $x \in \Omega$. Then*
$$|G| = |\operatorname{Orb}_G(x)| \cdot |\operatorname{Stab}_G(x)|.$$

Proof. Write $H = \operatorname{Stab}_G(x)$. For $g_1, g_2 \in G$, we have $g_1 x = g_2 x$ if and only if $g_2 g_1^{-1} x = x$ if and only if $g_2^{-1} g_1 \in H$, i.e. $g_1 H = g_2 H$. Hence the elements of $\operatorname{Orb}_G(x)$ are in '1-1' correspondence with the left cosets of H. So
$$|\operatorname{Orb}_G(x)| = |G|/|H| = |G|/|\operatorname{Stab}_G(x)|,$$
as requested. □

Remark. It follows immediately from the Orbit-Stabilizer Theorem 8 that $|\operatorname{Orb}_G(x)|$ divides $|G|$.

Definition. The action of G on Ω is *transitive* if $\operatorname{Orb}_G(x) = \Omega$ for any $x \in \Omega$.

Proposition 9. *Let G be a group acting on Ω. Define a relation $\sim$ by $x \sim y$ if $y \in \operatorname{Orb}_G(x)$. Then $\sim$ is an equivalence relation.*

Proof. Reflexivity: $ex = x$ by axiom (1), so $x \sim x$.

Symmetry: if $x \sim y$ then $y = gx$ for some $g \in G$, and hence $x = g^{-1} y \in \operatorname{Orb}_G(y)$, so $y \sim x$.

Transitivity: if $x \sim y$, $y \sim z$ then $y = gx$, $z = hy$ for some $g, h \in G$, and hence $z = hgx \in \operatorname{Orb}_G(x)$, so $z \sim x$. □

Consequence of Proposition 9: the orbits of the action give a *partition* of Ω (into equivalence classes of $\sim$).

Proposition 10. *Let G be a group acting on a set Ω. For $g \in G$, let $\varphi_g \colon \Omega \to \Omega$ given by $\varphi_g(x) = gx$. Then φ_g is a permutation of Ω, and moreover the map $\varphi \colon G \to \operatorname{Sym}(\Omega)$ given by $g \mapsto \varphi_g$ is a homomorphism.*[4]

Proof. Certainly φ_g is a function $\Omega \to \Omega$. Since $\varphi_{g^{-1}}$ is an inverse of φ_g, it follows that φ_g is a permutation. Thus φ defines a map $G \to \operatorname{Sym}(\Omega)$, as claimed.

Note that $\varphi_{gh}(x) = gh(x) = \varphi_g(hx) = \varphi_g(\varphi_h(x))$ for any $x \in \Omega$, and hence $\varphi(gh) = \varphi(g) \circ \varphi(h)$, so φ is a homomorphism. □

Definition. Let G act on Ω.

(1) The *kernel* of the action is the kernel of the homomorphism φ from Propostion 10, i.e. the set $\{g \in G : gx = x \text{ for all } x \in \Omega\}$. The kernel is a normal subgroup of G.

[4] We could actually define an action of G on Ω as a homomorphism $G \to \operatorname{Sym}(\Omega)$. The proposition guarantees that an action yields such a homomorphism, and obviously any such a homomorphism yields an action, so the definitions are indeed equivalent.

(2) The action is *faithful* if the kernel is trivial, i.e. the homomorphism φ is injective.

Some important actions.

Action 1. The action of G on itself by *left translation.* Let $\Omega = G$ and define the action $\psi\colon G \times \Omega \to \Omega$ by
$$\psi(g, x) = g * x = gx,$$
where $*$ is the group operation. We check this is an action:

(1) $\psi(e, x) = ex = x$,
(2) $\psi(gh, x) = (gh)(x) = g(hx) = \psi(g, \psi(h, x))$ by associativity.

This action is often called the *left regular action*[5].

Suppose $x, y \in G$ and put $g = yx^{-1}$. Then $gx = y$ and so $y \in \mathrm{Orb}_G(x)$. Hence the action is transitive. The kernel of the action is $\{e\}$, so the action is faithful.

Theorem 11 (Cayley's Theorem)**.** *Let G be a finite group. Then G is isomorphic to a subgroup of S_n for some n.*

Proof. Let $n = |G|$. Then $S_n \cong \mathrm{Sym}(G)$ and by Proposition 10, there is a homomorphism
$$\varphi\colon G \to \mathrm{Sym}(G)$$
corresponding to the left regular action of G with a trivial kernel. Therefore, by the First Isomorphism Theorem,
$$G \cong \mathrm{Im}\ \varphi \leq \mathrm{Sym}(G) \cong S_n,$$
as requested. $\square$

Action 2. The action of G on left cosets of a subgroup by *left translation*[6]. Let H be a subgroup of G and $\Omega = G/H$, the set of left cosets of H in G. Define the action $\psi\colon G \times \Omega \to \Omega$ by
$$\psi(g, xH) = gxH.$$
It is easy to check that is indeed an action. (Similar to Action 1.)

Let $xH, yH \in \Omega$. Letting $g = yx^{-1}$, we obtain $g(xH) = yH$, so the action is transitive. In fact, we will soon prove (Theorem 12) that, conversely, any transitive action is *equivalent* (which we will soon define) to an action of this type. This is why studying these actions is important.

For $xH \in \Omega$, we have that $gxH = xH$ is equivalent to $gx \in xH$, i.e. $g \in xHx^{-1}$. Therefore, $\mathrm{Stab}_G(xH) = xHx^{-1}$. (Notice that this implies that xHx^{-1} is a subgroup.)

Action 3. The action of G on itself by *conjugation.* Let $\Omega = G$ and define an action $\psi\colon G \times \Omega \to \Omega$ by
$$\psi(g, x) = gxg^{-1}.$$
This is a case where the notation gx for $\psi(g, x)$ is impossible. We instead write ${}^g x$ for gxg^{-1}. We call ${}^g x$ the *conjugate of x by g.*[7]

[5]There is also an analogous *right regular action.*
[6]There is also an analogous *right* action on the *right* cosets by *right translation.*
[7]There is also an analogous right action given by $(x, g) \mapsto x^g = g^{-1}xg$. Note that $x^g = {}^{g^{-1}}x$.

We check this is an action:

(1) $exe^{-1} = x$,
(2) $\psi(gh, x) = (gh)x(gh)^{-1} = ghxh^{-1}g^{-1} = \psi(g, hxh^{-1}) = \psi(g, \psi(h, x))$.

The orbits of the conjugation action are called *conjugacy classes*. We write ${}^{G}x$ for the conjugacy class containing x. (Other notations: $\mathrm{Con}_G(x)$, $\mathrm{Class}_G(x)$.)

The stabilizer of x is the subgroup

$$\mathrm{Stab}_G(x) = \{g \in G : gxg^{-1} = x\} = \{g \in G : gx = xg\},$$

so the set of all elements of G which commute with x. This is the *centralizer* of x in G, written $\mathrm{Cent}_G(x)$. For this action, the Orbit-Stabilizer Theorem 8 becomes

$$|{}^{G}x||\mathrm{Cent}_G(x)| = |G|.$$

In particular, the conjugacy class sizes divide $|G|$.

The action is never transitive if $|G| > 1$, since the identity e lies in an orbit of size 1.

The kernel of the action consists of those $g \in G$ which commute with everything in G. This is the *centre* of G, written $Z(G)$. Notice that $Z(G) = G$ whenever G is abelian.

Action 4. The action of G on its subgroups by conjugation. Let Ω be the set of subgroups of G and define an action $\psi\colon G \times \Omega \to \Omega$ by

$$\psi(g, H) = gHg^{-1} = {}^{g}H.$$

(We saw for Action 2 that $gHg^{-1} \leq G$.) It is easy to check that this is indeed a left action.[8] (Similar to Action 3.) The orbits are called *conjugacy classes* (of subgroups).

The stabilizer of H is the subgroup

$$\{g \in G : gHg^{-1} = H\} = \{g \in G : gH = Hg\}.$$

This is the *normalizer* of H, written $N_G(H)$.

Remark. Clearly, $N_G(H) \geq H$ and, in fact, $N_G(H)$ is the largest subgroup of G in which H is normal. Therefore, $H \trianglelefteq G$ if and only if $N_G(H) = G$ if and only if $\mathrm{Orb}_G(H) = \{H\}$.

Definition. Let G act on sets Ω_1 and Ω_2. We say that the two actions are *equivalent* if there exists a bijection $f\colon \Omega_1 \to \Omega_2$ such that $f(gx) = gf(x)$ for all $x \in \Omega_1$, $g \in G$.

Theorem 12 (Orbit-Stabilizer Theorem Revisited)**.** *Let G act transitively on a set Ω. Let $x \in \Omega$ and let $H = \mathrm{Stab}_G(x)$. Then the action of G on Ω is equivalent to the action of G on the cosets of H (Action 2).*

Proof. In the proof of the original Orbit Stabilizer Theorem 8, we found a '1-1' correspondence between the cosets of H and the orbit of x given by $gH \mapsto gx$. We show that the map $f\colon G/H \to \Omega$ given by $f(gH) = gx$ is an equivalence of actions.

[8]There is also an analogous right action given by $(H, g) \mapsto H^g = g^{-1}Hg$. Note that $H^g = {}^{g^{-1}}H$.

Certainly, f is a bijection, since $\mathrm{Orb}_G(x) = \Omega$. So we just need to show that $gf(hH) = f(ghH)$ for all $g, h \in G$. We see that

$$\begin{aligned} gf(hH) &= g(hx) \\ &= (gh)x \qquad \text{by action axiom (2)} \\ &= f(ghH) \end{aligned}$$

as requested. □

Thus every transitive action is equivalent to a coset action.

Automorphisms.

Definition. An *automorphism* of a group G is an isomorphism from G to itself. We write $\mathrm{Aut}(G)$ for the set of automotphisms.

Proposition 13. *For a group G, $\mathrm{Aut}(G)$ is a group under composition.*

Proof. Every automorphism is a permutation, so we want to show that $\mathrm{Aut}(G) \leq \mathrm{Sym}(G)$.

The identity permutation 1 is certainly an automorphism. The inverse of an isomorphism is an isomorphism, so $\mathrm{Aut}(G)$ is closed under inverses. The composition of two isomorphisms is an isomorphism, so $\mathrm{Aut}(G)$ is closed under composition. □

Example. Let $G = \mathbb{Z}_3 = \{0, 1, 2\}$. Any automorphism must fix 0. So the possibilities are id and θ such that $1 \longleftrightarrow 2$. It is easy to check that θ is an automorphism, indeed:

$$\begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \longmapsto \begin{array}{c|ccc} & 0 & 2 & 1 \\ \hline 0 & 0 & 2 & 1 \\ 2 & 2 & 1 & 0 \\ 1 & 1 & 0 & 2 \end{array}$$

Thus $\mathrm{Aut}(G) = \{\mathrm{id}, \theta\} \cong C_2 \cong \mathbb{Z}_3^\times$.

Proposition 14. *Let G be a group and $g \in G$. Let $\psi_g \colon G \to G$ be the conjugation map $x \mapsto {}^g x = gxg^{-1}$. Then ψ_g is an automorphism. For the conjugation action (Action 3), the corresponding homomorphism $G \to \mathrm{Sym}(G)$ is in fact a map $G \to \mathrm{Aut}(G)$.*

Proof. The second claim follows immediately from the first, so it will suffice to show that $\psi_g \in \mathrm{Aut}(G)$. Since ψ_g is a permutation, we just need to show it is a homomorphism. We see that $\psi_g(xy) = gxyg^{-1} = gxg^{-1}gyg^{-1} = \psi_g(x)\psi_g(y)$, as required. □

The set of conjugation maps on G is a subgroup of $\mathrm{Aut}(G)$, since it is the image of the homomorphism $g \mapsto \psi_g$. In general, this image is not the whole of $\mathrm{Aut}(G)$ (see, example above). We call it the group of *inner automorphisms*, $\mathrm{Inn}(G)$.

We know that $\mathrm{Ker}\, \psi$ is the centre, $Z(G)$ (from discussion of Action 3). Therefore

$$\mathrm{Inn}(G) \cong G/Z(G)$$

by the First Isomorphism Theorem.

Proposition 15.

(1) *Let H be a subgroup of G and $g \in G$. Then ${}^gH \cong H$. (So each orbit of Action 4 consists of isomorphic subgroups.)*
(2) *Let G act on Ω and $x \in \Omega$ have stabilizer H. If $y = gx$, then $\mathrm{Stab}_G(y) = {}^gH$.*

Proof. (1) Observe that ${}^gH = \mathrm{Im}(\psi_g|_H)$. The result follows easily from the First Isomorphism Theorem.

(2) We have $hy = y$ if and only if $hgx = gx$ if and only if $g^{-1}hgx = x$. Thus $h \in \mathrm{Stab}_G(y)$ if and only if $h \in {}^gH$. □

3. Sylow's Theorems

This chapter is about *converses* to Lagrange's Theorem. Suppose m divides $|G|$. Is there a subgroup of G of order m? In general, no. The group A_4, of order 12, has no subgroup of order 6. However, in some special cases, we can get a converse.

Definition. Let p be a prime number.

(1) A *p-group* is a group whose order is p^a for some $a \in \mathbb{N} \cup \{0\}$.
(2) If G is a group, then a *p-subgroup* of G is a subgroup which is a p-group.
(3) A *p-element* of a group G is an element whose order is p^a for some $a \in \mathbb{N} \cup \{0\}$.

Note that $\{e\}$ is a p-group for any prime p, and e is a p-element.

Theorem 16 (Cauchy's Theorem). *Let G be a finite group whose order is divisible by a prime p. Then G has an element of order p.*

Proof. Let $\Omega \subseteq \underbrace{G \times \cdots \times G}_{p \text{ times}}$ be the set of p-tuples $(g_1, \ldots, g_p)$ such that $g_1 \ldots g_p = e$. How big is Ω? We can choose $g_1, \ldots, g_{p-1}$ in any way we like; this forces $g_p = (g_1 \ldots g_{p-1})^{-1}$, so we have no choice for g_p. So $|\Omega| = |G|^{p-1}$.

Let C_p act on Ω by *rotations*. Explicitly, $C_p = \langle t \rangle$ and let

$$t(g_1, \ldots, g_p) = (g_2, g_3, \ldots, g_p, g_1).$$

(This gives a well-defined action on C_p.) Every orbit of C_p on Ω has size 1 or p. Say there are A orbits of size 1, B of size p. Then

$$|\Omega| = A + pB.$$

Since $p||G|$, we have $p||\Omega|$, and hence $p|A$.

Now, an orbit of size 1 in Ω consists of a tuple $(g, g, \ldots, g)$ with all entries the same. Such a tuple is in Ω only if $g^p = e$. So A is the number of solutions to $g^p = e$ in G. Now, clearly e is a solution, so $A \geq 1$, and since $p|A$, we must have $A \geq p$. Hence there exists $g \neq e$ such that $g^p = e$, an element of order p. □

Remark. If G has an element g of order p, then G has a subgroup $\langle g \rangle$ of order p.

Theorem 17 (First Sylow Theorem). *Let G be a finite group and let p^a be the largest power of p dividing $|G|$. Then G has a subgroup of order p^a.*

Such a subgroup is called a *Sylow p-subgroup* of G.

Proof. The proof goes by induction on $|G|$. (The base case is the trivial group $\{e\}$.)

Inductive hypothesis. For any group G of order less than n, G has a Sylow p-subgroup.

Inductive step. Let $|G| = n$.

We will use the *class equation*:

$$|G| = |Z(G)| + \sum_{i=1}^{k} |C_i|$$

where $C_1, \ldots, C_k$ are the conjugacy classes of non-central elements of G. (Note that if $z \in Z(G)$, then ${}^g z = z$ for all g, so ${}^G z = \{z\}$. Also, if ${}^G h = \{h\}$, then ${}^g h = h$ for all g, so $h \in Z(G)$. So $|Z(G)|$ is the number of conjugacy classes of size 1 in G.)

We may assume that $p \bigm| |G|$; otherwise, $\{e\}$ is a Sylow p-subgroup.

Case (i). $p \bigm| |Z(G)|$. Then $Z(G)$ contains an element z of order p. Since ${}^g z = z$ for all $g \in G$, we see that $\langle z \rangle \trianglelefteq G$. Now, $G/\langle z \rangle$ has order $n/p < n$, so $G/\langle z \rangle$ has a Sylow subgroup X of order p^{a-1} by the inductive hypothesis. The Correspondence Theorem 2 tells us that if $\theta \colon G \to G/\langle z \rangle$ is the canonical map, then $\theta^{-1}(X)$ is a subgroup of order p^a in G.

Case (ii). $p \nmid |Z(G)|$. Then $p \nmid \sum_{i=1}^{k} |C_i|$ (by the class equation). So for some i, we have $p \nmid |C_i|$. Take $x \in C_i$. Then $|G| = |\mathrm{Cent}_G(x)| \cdot |C_i|$ and so p^a divides $|\mathrm{Cent}_G(x)|$. Since $x \notin Z(G)$, we see that $\mathrm{Cent}_G(x) < G$. So $|\mathrm{Cent}_G(x)| < n$. Thus by the inductive hypothesis, $\mathrm{Cent}_G(x)$ has a Sylow p-subgroup P of order p^a. Finally, $P \leq G$ shows that G has a Sylow p-subgroup. □

Write $\mathrm{Syl}_p(G)$ for the set of Sylow p-subgroups of G and $n_p = n_p(G)$ for the number of Sylow p-subgroups.

Lemma 18. *Let G be a finite group and let P be a Sylow p-subgroup of G. Let Q be any p-subgroup of G. Then either $Q \leq P$ or $Q \not\leq N_G(P)$. In other words, if $Q \not\leq P$, then there exists $q \in Q$ such that ${}^q P \neq P$.*

Proof. We show that if $Q \leq N_G(P)$, then $Q \leq P$. If $Q \leq N_G(P)$, then ${}^q P = P$ for any $q \in Q$, so $qP = Pq$ for all $q \in Q$. In particular, this implies that $QP = PQ$. So PQ is a subgroup of G, with order $|P||Q|/|P \cap Q|$. But $|P||Q|/|P \cap Q|$ is a power of p, and it divides $|G|$, since PQ is a subgroup. But $|P|$ is the largest power of p dividing $|G|$, and so we must have $|Q|/|P \cap Q| = 1$. Thus $Q \leq P$. □

Theorem 19 (Second Sylow Theorem)**.** *Let G be a finite group and p be prime. Then $n_P(G) \equiv 1 \bmod p$.*

Theorem 20 (Third Sylow Theorem)**.** *Let G be a finite group, p be a prime, and Q be a p-subgroup of G. Then Q is contained in a Sylow p-subgroup.*

Theorem 21 (Fourth Sylow Theorem)**.** *Let G be a finite group and p be a prime. The Sylow p-subgroups of G form a single conjugacy class of subgroups of G.*

Remark. The Fourth Sylow Theorem 21 implies that $n_p(G)$ divides $|G|$ via the Orbit-Stabilizer Theorem 8.

Proof of Theorem 19. Notice that if $H \leq G$ and $g \in G$, then $|{}^gH| = |H|$. In particular, if H is a Sylow p-subgroup, then so is the conjugate. Thus G acts by conjugation on $\mathrm{Syl}_p(G)$.

Let P be a Sylow p-subgroup. (We know one exists by Theorem 17.) Look at the action of P on $\mathrm{Syl}_p(G)$. Clearly, ${}^PP = P$, so $\{P\}$ is an orbit of size 1. Suppose that $\{Q\}$ is an orbit of size 1. Then ${}^PQ = Q$, so $P \leq N_G(Q)$. But Q is a p-subgroup, so by Lemma 18, we have $P \leq Q$. But $|P| = |Q|$, so $P = Q$. So P has exactly one orbit of size 1 on $\mathrm{Syl}_P(G)$.

Every other orbit has size dividing $|P|$ (by Orbit-Stabilizer Theorem 8), so a power of p greater than 1. So p divides the size of every orbit except $\{P\}$, and so $|\mathrm{Syl}_p(G)| \equiv 1 \bmod p$, as required. □

Proof of Theorem 20. Let Q be a p-subgroup of G. Let Q act on $\mathrm{Syl}_p(G)$. Then every orbit has size p^a for some $a \in \mathbb{N} \cup \{0\}$. Since $|\mathrm{Syl}_p(G)| \equiv 1 \bmod p$ by the Second Sylow Theorem 19, not every orbit has size divisible by p, and so there is an orbit $\{P\}$. Then $Q \leq N_G(P)$, and so $Q \leq P$ by Lemma 18. □

Proof of Theorem 21. Let $P \in \mathrm{Syl}_p(G)$ and Ω be the conjugacy class of P in G. Then G acts on Ω by conjugation.

Let P act on Ω. Then P has one orbit $\{P\}$ of size 1, and the others have size divisible by p. So $|\Omega| \equiv 1 \bmod p$. (By the same argument as in the proof of the Second Sylow Theorem 19.)

Now, let $Q \in \mathrm{Syl}_p(G)$, and let Q act on Ω. The orbits all have p-power length. Since $|\Omega| \equiv 1 \bmod p$, there must be an orbit $\{R\}$ of length 1. Now $Q \leq N_G(R)$, and so $Q \leq R$ by Lemma 18. But $|Q| = |R|$, since they are both Sylow p-subgroups, so $Q = R$. Therefore, $Q \in \Omega$, showing that $\mathrm{Syl}_p(G) = \Omega = {}^GP$, as required. □

Example (Groups of order 15)**.** Let G has order 15. We use *Sylow arithmetic* to show that $G \cong C_{15}$. We have:

$$n_5 \equiv 1 \bmod 5 \quad \text{and it divides } 15.$$

Thus $n_5 = 1$, G has a unique Sylow 5-subgroup $N \trianglelefteq G$. Also:

$$n_3 \equiv 1 \bmod 3 \quad \text{and it divides } 15.$$

Thus $n_3 = 1$ as well. So there is a unique Sylow 3-subgroup $M \trianglelefteq G$.

Note that G has exactly 4 elements of order 5, 2 elements of order 3, 1 element of order 1. There are 8 remaining elements, which must have order 15. Thus $G \cong C_{15}$.

Proposition 22. *Let p be a prime and G be a non-trivial p-group. Then $Z(G)$ is non-trivial. (In particular, since $p \bigm| |Z(G)|$, by Cauchy's Theorem 16, G has a central element g of order p, and so a normal subgroup $\langle g \rangle \cong C_p$.)*

This proposition allows us to use induction arguments for p-groups by considering quotients by the normal subgroup $\langle g \rangle \cong C_p$.

Proof. We recall the Class Equation:

$$|G| = |Z(G)| + \sum_{i=1}^{k} |C_i|$$

where $C_1, \ldots, C_K$ are the non-central conjugacy classes. Since each $|C_i|$ is a p-power greater than 1, we have that p divides $|C_i|$ for all i. Moreover, $p \mid |G|$, so $p \mid |Z(G)|$, and hence $|Z(G)| \neq 1$. □

Remark. A stronger form of Proposition 22 is the following: if G is a p-group and if N is a non-trivial normal subgroup of G, then $N \cap Z(G)$ is non-trivial.

Proposition 23. *Let G be a finite group, and suppose that p^b divides $|G|$. Then G has a subgroup of order p^b.*

Proof. Let p^a be the largest power of p dividing $|G|$, so $(b \leq a)$. Then G has a subgroup P of order p^a by First Sylow Theorem 17. So if P has a subgroup of order p^b, then so does G. So it is enough to prove the proposition for p-groups. So assume that $|G| = p^a$. We can also assume that $b > 0$, since otherwise $\{e\}$ is a subgroup of order p^b.

We work by induction on a. If $a = 1$, the statement is trivial.

Inductive hypothesis. A group of order p^a has a subgroup of order p^b for all $b \leq a$.

Inductive step. Suppose $|G| = p^{a+1}$. Let K be a normal subgroup of G of order p (which exists by Proposition 22). Then G/K is a group of order p^a. If $b \leq a+1$, then $b - 1 \leq a$, and so G/K has a subgroup X of order p^{b-1} by the inductive hypothesis. Under the Subgroup Correspondence (Theorem 2), X corresponds to a subgroup L of G of order $|K||X| = pp^{b-1} = p^b$. Therefore, G has a subgroup of order p^b for all $b \leq a+1$. □

4. Automorphism Groups and Semidirect Products

Recall that an *automorphism* of G is an isomorphism $G \to G$, and that the automorphisms of G form a group $\mathrm{Aut}(G)$ under composition.

Examples.

(1) Let $G = C_n \cong \mathbb{Z}_n$. Note that $\mathbb{Z}_n = \langle 1 \rangle$, and that any homomorphism from $\mathbb{Z}_n$ to a group H is determined by where it sends 1. So an automorphism of $\mathbb{Z}_n$ is of the form $\varphi_t \colon 1 \mapsto t$ for $t \in \mathbb{Z}_n$. In fact, this gives all the homomorphisms $\mathbb{Z}_n \to \mathbb{Z}_n$ (endomorphisms of $\mathbb{Z}_n$).

We need to identify which φ_t are invertible. Note that $\varphi_t(x) = tx$, so φ_t is *multiplication by* t. So φ_t is invertible whenever t has a multiplicative inverse, i.e. $t \in \mathbb{Z}_n^*$. Note that $\varphi_t \circ \varphi_s = \varphi_{ts}$, so the map $\mathrm{Aut}(\mathbb{Z}_n) \to \mathbb{Z}_n^*$ given by $\varphi_t \mapsto t$ is an isomorphism.

Fact. $\mathbb{Z}_p^* \cong C_{p-1}$ whenever p is a prime. (This is proved using elementary number theory.) Thus $\mathrm{Aut}(C_p) \cong C_{p-1}$.

(2) Let $G = C_2 \times C_2$. Think of G as $\{e, a, b, c\}$ with multiplication given by

$$a^2 = b^2 = c^2 = e, ab = ba = c, ac = ca = b, bc = cb = a.$$

Thinking of an automorphism φ as a permutation on G, we see that $\varphi(e) = e$ (for any automorphism). So we may as well consider φ as an element of $\mathrm{Sym}(\{a, b, c\})$. Thus

Aut(G) is isomorphic to a subgroup of S_3. Now, any permutation of a, b, c preserves the multiplication equations above, so is a homomorphism. Hence $\operatorname{Aut}(G) \cong S_3$.

(3) Let $G = S_3$. We have

Conjugacy classes	Order of elements
$\{\mathrm{id}\}$	1
$\{(123), (132)\}$	3
$\{(12), (13), (23)\}$	2

Since the disjoint conjugacy classes have elements of different orders, any automorphism of S_3 must preserve them. So, in particular, $\operatorname{Aut}(S_3)$ acts on $\Omega = \{(12), (13), (23)\}$. Now, $\operatorname{Inn}(S_3) \cong S_3/Z(S_3) = S_3/\{\mathrm{id}\} \cong S_3$, and it is easy to see that $\operatorname{Inn}(S_3)$ acts as $\operatorname{Sym}(\Omega)$ on Ω. Now, suppose that the homomorphism

$$\operatorname{Aut}(S_3) \to \operatorname{Sym}(\Omega)$$

has kernel K and $\alpha \in K$. So $\alpha(12) = (12)$, $\alpha(13) = (13)$, $\alpha(23) = (23)$. Then

$$\alpha(123) = \alpha(13)(12) = \alpha(13)\alpha(12) = (13)(12) = (123),$$

$$\alpha(132) = \alpha(12)(13) = \alpha(12)\alpha(13) = (12)(13) = (132),$$

so α is the identity map on S_3. Hence $\operatorname{Aut}(S_3)$ is isomorphic to a subgroup of $\operatorname{Sym}(\Omega) \cong S_3$, but $|\operatorname{Aut}(S_3)| \geq |\operatorname{Inn}(S_3)| = 6$, so $\operatorname{Aut}(S_3) \cong S_3$.

Definition. Let G be a group and let H and K be subgroups. We say that H and K are *complementary* subgroups of G if

(1) $G = HK$,
(2) $H \cap K = \{e\}$.

Recall that $|HK| = |H||K|/|H \cap K|$ by Proposition 4. So if H and K are complementary in G, then $|G| = |H||K|$, and furthermore, every element of G has a unique representation as hk for $h \in H$, $k \in K$. This gives us a sort of *decomposition* of G into subgroups—but the multiplication is hard to understand.

We look at the case where H is normal in G. Let us try multiplying two elements of HK, h_1k_1 and h_2k_2. The product is $h_1k_1h_2k_2$, and we know that this is h_3k_3 for some $h_3 \in H$, $k_3 \in K$. So $k_1h_2 = h_1^{-1}h_3k_3k_2^{-1}$. We want to know what $h_1^{-1}h_3$ and $k_3k_2^{-1}$ are. If H is normal in G, then $k_1H = Hk_1$, so we must have $k_3k_2^{-1} = k_1$ (since the representation of every element is unique). Hence we have *halved* the problem.

All we need to know is $h_1^{-1}h_3 = k_1h_2k_1^{-1} = {}^{k_1}h_2$. So to understand multiplication in HK, we need to understand how the conjugation maps by elements of K act on H.

In general, we can define the following product of groups.

Definition. Let N, K be groups, and let $\varphi\colon K \to \operatorname{Aut}(N)$ be a homomorphism. The *semidirect product* of N by K via φ is the set of pairs $N \times K$ with multiplication given by

$$(n_1, k_1)(n_2, k_2) = (n_1\varphi_{k_1}(n_2), k_1k_2),$$

where φ_{k_1} is the image of k_1 under φ.

In the discussion above, we looked at the situation where $N \trianglelefteq G$, $K \leq G$ were complementary subgroups. We saw that

$$n_1k_1n_2k_2 = n_1k_1n_2k_1^{-1}k_1k_2 = n_1({}^{k_1}n_1)k_1k_2.$$

Now, $n \mapsto {}^{k_1}n$ is an automorphism of N. Writing φ_{k_1} for this automorphism, we have

$$(n_1k_1)(n_2k_2) = n_1\varphi_{k_1}(n_2)k_1k_2$$

which explains the definition above.

Proposition 24. *Let N, K be groups and $\varphi\colon K \to \mathrm{Aut}(N)$. Then*

1. *The semidirect product of N by K via φ is a group, written $N \rtimes_\varphi K$.*
2. *The set $\{(n, e_K) : n \in N\}$ is a normal subgroup isomorphic to N. The set $\{(e_N, k) : k \in K\}$ is a subgroup isomorphic to K. These two subgroups are complementary.*
3. *If G is a group with complementary subgroups N and K with N normal, and if $\varphi_k(n) = {}^kn$ for all $k \in K$, $n \in N$, then*
$$G \cong N \rtimes_\varphi K.$$

Proof. (1) **Associativity:**

$$\begin{aligned}((n_1,k_1)(n_2,k_2))(n_3,k_3) &= (n_1\varphi_{k_1}(n_2), k_1k_2)(n_3,k_3)\\ &= (n_1\varphi_{k_1}(n_2)\varphi_{k_1k_2}(n_3), k_1k_2k_3)\\ &= (n_1\varphi_{k_1}(n_2\varphi_{k_2}(n_3)), k_1k_2k_3) \qquad \text{since } \varphi, \varphi_{k_1} \text{ are homomorphisms}\\ &= (n_1,k_1)(n_2\varphi_{k_2}(n_3), k_2k_3)\\ &= (n_1,k_2)((n_2,k_2)(n_3,k_3))\end{aligned}$$

Identity: (e_N, e_K), since φ_{e_K} is the identity automorphism.

Inverses: (n,k) has the inverse $(\varphi_{k^{-1}}(n^{-1}), k^{-1})$.

Therefore, $N \rtimes_\varphi K$ is a group.

(2) It is easy to check that the map $n \mapsto (n, e_K)$ is an injective homomorphism $N \to G$. Similarly, the map $k \mapsto (e_N, k)$ is also an injective homomorphism $K \to G$. The images of these maps are the sets identified in the proposition. Clearly, $\{(n, e_K)\} \cap \{(e_N, k)\} = \{(e_N, e_K)\}$ and any element (n,k) can be written as $(n, e_K)(e_N, k)$, so the two subgroups are complementary.

(3) This follows from the discussion above. □

Example (Groups of order 21)**.** Let $|G| = 21$. First, use Sylow's Theorems. For 7,

$$n_7 \equiv 1 \bmod 7 \text{ and divides } 21,$$

so $n_7 = 1$, and G has a normal Sylow 7-subgroup N. (For 3, $n_3 \equiv 1 \bmod 3$ divides 21, so n_3 can be 1 or 7. This is not very helpful.) Let K be a Sylow 3-subgroup.

Now $K \cap N = \{e\}$ (considering orders) and

$$|NK| = \frac{|N||K|}{|N \cap K|} = \frac{7 \times 3}{1} = 21,$$

so $NK = G$. So N and K are complementary subgroups, with N normal.

So $G \cong N \rtimes_\varphi K$ for some $\varphi\colon K \to \mathrm{Aut}(N)$ (with φ determined by the conjugacy maps of K on N). Now

$$\mathrm{Aut}(N) \cong \mathrm{Aut}(C_7) \cong C_6.$$

We see that $\operatorname{Im}\varphi$ can have order 1 or 3.

Case 1. $\operatorname{Im}\varphi = \{\mathrm{id}\}$. In this case, our multiplication is $(n_1, k_2)(n_2, k_2) = (n_1n_2, k_1k_2)$. Thus

$$G \cong C_7 \times C_3 \cong C_{21}.$$

Case 2. $\operatorname{Im}\varphi = C_3$. Let α be an automorphism of N of order 3. Let k be the element of K such that $\varphi_k = \alpha$. Let $\langle n\rangle = N$. Then our multiplication is

$$(n^i, k^j)(n^u, k^v) = (n^i\alpha^j(n^u), k^jk^v).$$

Take $\alpha(n) = n^2$. Then we have

$$(n^i, k^j)(n^u, k^v) = (n^{i+2^ju}, k^{j+v}).$$

This gives the non-cyclic group of order 21.

(So there are exactly two groups of order 21.)

Proposition 25. *Let p and q be distinct primes, with $p < q$. If $q \equiv 1 \bmod p$, then there are exactly two groups of order pq, up to isomorphism. Otherwise, there is only one (the cyclic group).*

Proof. Let $|G| = pq$. We have $n_q \equiv 1 \bmod q$ divides pq. Since $p, q, pq \not\equiv 1 \bmod q$, we have $n_q = 1$, so G has a normal Sylow q-subgroup, N. Also, $n_p \equiv 1 \bmod p$ divides pq.

If $q \not\equiv 1 \bmod p$, then we must have $n_p = 1$. In this case, G has a normal Sylow p-subgroup M. Thus G has

$$\begin{gathered} 1 \text{ element of order } 1 \\ p-1 \text{ elements of order } p \\ q-1 \text{ elements of order } q \end{gathered}$$

and $p + q - 1 < pq$, so G has an element of order pq. Hence $G \cong C_{pq}$.

Now, suppose $q \equiv 1 \bmod p$. Let K be a Sylow p-subgroup. Then N and K are complementary subgroups, so

$$G \cong N \rtimes_\varphi K$$

for some $\varphi\colon K \to \mathrm{Aut}(N)$. Now, $\mathrm{Aut}(N) \cong C_{q-1}$, so $\operatorname{Im}\varphi$ is either trivial or else isomorphic to C_p. If $\operatorname{Im}\varphi$ is trivial, then $G \cong C_q \times C_p \cong C_{pq}$. Otherwise, let a be an element of order p in $\mathbb{Z}_q^*$. Then $n \mapsto n^a$ defines an automorphism θ_a of N of order p, where $N = \langle n\rangle$. Let $K = \langle k\rangle$, where $\varphi(k) = \theta_a$. Then we have a representation of G as

$$\langle n, k \mid n^q = 1,\ k^p = 1, {}^kn = n^a\rangle.$$

(This is a *generator-relation presentation*[9].) So there is only one group of order pq other than the cyclic group. □

[9]The group is generated by the elements on the left and the only relations in the group can be derived from the relations on the right.

Remark. When G is a group with complementary subgroups N, K, with N normal, then $G \cong N \rtimes_\varphi K$ with φ given by conjugation maps. In this case, we often omit the φ and just write $G \cong N \rtimes K$. Then this notation only really means that G is a group with complementary subgroups N and K, where N is normal.

Examples.

- $S_n \cong A_n \rtimes C_2$. (Take $C_2 \cong \langle (12) \rangle$. Clearly, $\langle (12) \rangle \cap A_n = \{e\}$ and $S_n = A_n \cup \langle (12) \rangle A_n$, so they are complementary subgroups with A_n normal.)
- $S_4 \cong V_4 \rtimes S_3$. (Consider $S_3 = \mathrm{Stab}_{S_4}(4) \leq S_4$, and recall that
$$V_4 = \{e, (12)(34), (13)(24), (14)(23)\}.$$
These are complementary and V_4 is normal.)
- $\mathrm{GL}_n(F) = \mathrm{SL}_n(F) \rtimes H$, where $H \cong F^\times$. (Take
$$H = \left\{ \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & \lambda \end{pmatrix} : \lambda \in F^\times \right\}.$$
Then $\mathrm{SL}_n(F)$ and H are complementary subgroups of $\mathrm{GL}_n(F)$ with $\mathrm{SL}_n(F)$ normal.)

5. Composition Series

Definition. A group G is *simple* if it has no normal subgroups except $\{e\}$ and G.

If G is **not** simple, then there exists $N \lhd G$, $N \neq \{e\}$ and we can study G by looking at the smaller groups N and G/N.

Definition. A *composition series* for G is a chain of subgroups $\{G_i : 0 \leq i \leq k\}$, with
$$G = G_0 > G_1 > G_2 > \cdots > G_k = \{e\},$$
such that $G_{i+1} \lhd G_i$ and $\frac{G_i}{G_{i+1}}$ is simple for all i. These quotients $\frac{G_i}{G_{i+1}}$ are the *composition factors* of the series. The number k is the *length* of the series[10].

Example. The group S_4 has a composition series (with composition factors written below):
$$\begin{array}{ccccccccc} S_4 & > & A_4 & > & V_4 & > & \langle (12)(34) \rangle & > & \{e\} \\ & C_2 & & C_3 & & C_2 & & C_2 & \end{array}$$
(Note that this series is not unique—we could have chosen $\langle (13)(24) \rangle$ or $\langle (14)(23) \rangle$ instead of $\langle (12)(34) \rangle$.)

Proposition 26. *Let G be a finite group and $N \trianglelefteq G$. Then G has a composition series including N. In particular, taking $N = G$, every finite group has a composition series.*

Remark. The proposition does not hold for infinite groups. For example, take $G = \mathbb{Z}$. The subgroups of $\mathbb{Z}$ are $n\mathbb{Z}$ for $n \in \mathbb{N} \cup \{0\}$, but, if $n \neq 0$, then $n\mathbb{Z} \cong \mathbb{Z}$, so it is not a simple group. After k terms of a '*composition series*', the group G_k must still be isomorphic to $\mathbb{Z}$. So no finite series exists.

[10] Note that there are $k + 1$ groups in the series, and the length is in fact the number of inclusions.

Proof. We work by induction. Suppose the statement is true for groups of order $< n$. Suppose $|G| = n$. If G is simple, then $G = G_0 > G_1 = \{e\}$ is a composition series. So assume that G has a non-trivial, proper, normal subgroup N. Then $|N|$ and $|G/N|$ are both less than n, and so these groups have composition series:

$$N = N_0 > N_1 > \cdots > N_k = \{e\},$$

$$G/N = Q = Q_0 > Q_1 > \cdots > Q_l = \{e\}.$$

By the Correspondence Theorem 2, there exist subgroups $L_i = \theta^{-1}(Q_i) < G$ (where θ is the canonical map $G \to G/N$). We have

$$G = L_0 > L_1 > \cdots > L_l = \ker(\theta) = N.$$

We also have $L_{i+1} \trianglelefteq L_i$ and $\frac{L_i}{L_{i+1}} \cong \frac{Q_i}{Q_{i+1}}$ by the Second Isomorphism Theorem 3, which is simple. So we have

$$G = L_0 > L_1 > \cdots > L_l = N = N_0 > N_1 > \cdots > N_k = \{e\},$$

a composition series for G, including N. □

A group can have multiple composition series. We saw that S_4 has three.

Example. The group C_{12} has a few composition series (with composition factors written below):

$$\begin{array}{ccccccc} C_{12} & > & C_6 & > & C_3 & > & C_1 \\ & C_2 & & C_2 & & C_3 & \end{array}$$

$$\begin{array}{ccccccc} C_{12} & > & C_6 & > & C_2 & > & C_1 \\ & C_2 & & C_3 & & C_2 & \end{array}$$

$$\begin{array}{ccccccc} C_{12} & > & C_4 & > & C_2 & > & C_1 \\ & C_3 & & C_2 & & C_2 & \end{array}$$

We see that the composition factors are the same for the three series, but their order varies.

Theorem 27 (Jordan-Hölder Theorem)**.** *Any two composition series for a finite group G have the same length, and the same composition factors (with the same multiplicities, but not necessarily the same order).*

Proof. We work by induction on $|G|$. Suppose the theorem is valid for groups of size less than n. Suppose that $|G| = n$, and that G has composition series

$$\mathcal{A} : G = G_0 > G_1 > \cdots > G_k = \{e\},$$

$$\mathcal{B} : G = H_0 > H_1 > \cdots > H_l = \{e\}.$$

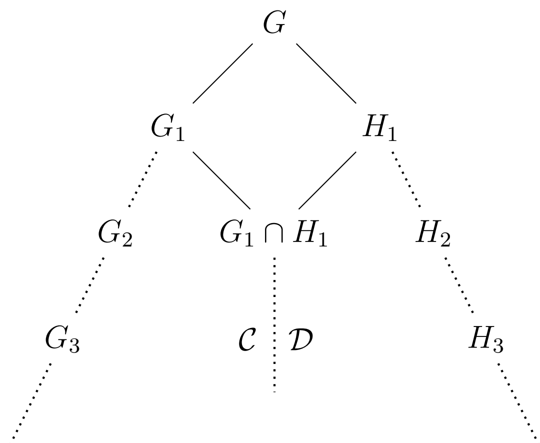
Case 1. $G_1 = H_1$. Since $|G_1| < n$, any two series for G_1 have the same length and factors. Now,

$$\mathcal{A}_1 : G_1 > G_2 > \cdots > G_k,$$

$$\mathcal{B}_1 : H_1 > H_2 > \cdots > H_l$$

are both composition series for G_1, so $k = l$ and $\mathcal{A}_1$, $\mathcal{B}_1$ have the same factors. Since $G_0/G_1 = H_0/H_1$, the series $\mathcal{A}$, $\mathcal{B}$ have the same factors.

Case 2. $G_1 \neq H_1$. Note that G/G_1 is simple, and so G has no normal subgroups strictly between G_1 and G. Similarly with H_1. But $G_1H_1 \trianglelefteq G$, and $G_1 \leq G_1H_1 \leq G$. Since $H_1 \not\leq G_1$, we must have $G_1H_1 = G$. Also, note that $G_1 \cap H_1 \trianglelefteq G_1$. Therefore, G_1 has a composition series containing $\mathcal{C}$ which includes $G_1 \cap H_1$ by Proposition 26. Since $|G_1| < n$, any composition series for G_1 has the same length and factors as $\mathcal{C}$. So $\mathcal{A}_1$ (as defined in Case 1) has the same factors as $\mathcal{C}$. Similarly, H_1 has a composition series $\mathcal{D}$ including $G_1 \cap H_1$, and $\mathcal{B}_1$ (as defined in Case 1) has the same factors as $\mathcal{D}$. We may assume that $\mathcal{C}$ and $\mathcal{D}$ agree below $G_1 \cap H_1$.



Note that by the Second Isomorphism Theorem 3:

$$\frac{G_1}{G_1 \cap H_1} \cong \frac{G_1H_1}{H_1} \cong \frac{G}{H_1},$$
$$\frac{H_1}{G_1 \cap H_1} \cong \frac{G_1H_1}{G_1} \cong \frac{G}{G_1},$$

which are both simple.

Therefore, the composition factors of $\mathcal{A}$ are those of $G_1 \cap H_1$ together with $\frac{G}{G_1}$ and $\frac{G_1}{G_1 \cap H_1} \cong \frac{G}{H_1}$. Similarly, the composition factors for $\mathcal{B}$ are those of $G_1 \cap H_1$ together with $\frac{G}{H_1}$ and $\frac{H_1}{G_1 \cap H_1} \cong \frac{G}{G_1}$. Hence $\mathcal{A}$ and $\mathcal{B}$ have the same length and composition factors. $\square$

Henceforth, we can refer to composition factors of a group, not just a series.

Note that different groups can have the same composition factors. Therefore, the composition factors do not determine a group, but we can use them to distinguish between them (if the composition factors are different).

Example. The groups C_4 and $C_2 \times C_2$ both have composition factors C_2, C_2. Similarly, C_6 and S_3 both have composition factors C_2, C_3.

Composition factors are simple groups. The finite simple groups are classified into 14 infinite families and 26 *sporadic* examples. The easiest family to understand is $\{C_p : p \text{ prime}\}$. These are the only abelian simple groups.

A group whose composition factors are all abelian is said to be *soluble* (*solvable*). We give a more general definition which works for infinite groups too.

Definition. A group G is *soluble* (*solvable*) if there exist subgroups

$$G = G_0 > G_1 > \cdots > G_k = \{e\}$$

such that $G_{i+1} \lhd G_i$ for all i, and $\frac{G_i}{G_{i+1}}$ is abelian.

Remark. If $H \lhd G$ and G is finite, then the composition factors of G are those of N together with those of G/N.[11] Now, if G has a series as in the definition above, and if G has a composition series, then then every composition factor of G is a composition factor of some abelian group $\frac{G_i}{G_{i+1}}$. So the composition factors of G are abelian.

Examples.

(1) Any abelian group is soluble.
(2) A dihedral group D_{2n} has a normal subgroup isomorphic to C_n with index 2, so $D_{2n} > C_n > \{e\}$ is a series as required by the definition. So D_{2n} is soluble.
(3) For S_4, we have a series $S_4 > A_4 > V_4 > \{e\}$, as required by the definition. So S_4 is soluble.
(4) For $n \geq 5$, S_n is **not** soluble, since it has A_n as a composition factor, which is non-abelian and simple (for the proof that A_n is simple for $n \geq 5$, see Appendix A).

Remark. If $N \trianglelefteq G$, then G is soluble if and only if both N and G/N are soluble.

Theorem 28. *If a finite group G is soluble and $H \leq G$, then H is soluble.*

Proof. Take a composition series for G:

$$G = G_0 > G_1 > \cdots > G_k = \{e\}.$$

Define $H_i = H \cap G_i$ for $i = 0, 1, \ldots, k$. Then

$$H = H_0 \geq H_1 \geq \cdots \geq H_k = \{e\}.$$

Now, $H_{i+1} = H \cap G_{i+1} \trianglelefteq H \cap G_i = H_i$, and by the Third Isomorphism Theorem 6:

$$\frac{H_i}{H_{i+1}} = \frac{H \cap G_i}{H \cap G_{i+1}} \cong \frac{G_{i+1}(H \cap G_i)}{G_{i+1}} = \frac{G_{i+1}H_i}{G_{i+1}} \leq \frac{G_i}{G_{i+1}},$$

since $H \cap G_i \cap G_{i+1} = H \cap G_{i+1}$. But $\frac{G_i}{G_{i+1}}$ is a composition factor of G so it is abelian, and the subgroup $\frac{H_i}{H_{i+1}}$ is also abelian. Now, simply deleting all H_i such that $H_i = H_{i+1}$ from the series, we obtain a composition series for H with abelian composition factors. Thus H is soluble. □

Definition. Let G be a group and let $x, y \in G$. The *commutator* $[x, y]$ of x and y is the element $xyx^{-1}y^{-1}$.

Notice that $[x, y] = e$ if and only if x and y commute (since $xyx^{-1}y^{-1} = e$ if and only if $yx = xy$).

Proposition 29. *Let G be a group and let $N \trianglelefteq G$. Then $\frac{G}{N}$ is abelian if and only if N contains every commutator in G.*

[11] We have seen this in inductive step of the proof of Proposition 26.

Proof. We have xN and yN commute if and only if $[xN, yN] = eN$. Now

$$[xN, yN] = xNyNx^{-1}Ny^{-1}N = xyx^{-1}y^{-1}N = [x, y]N$$

and this is equal to eN if and only if $[x, y] \in N$. So $\frac{G}{N}$ is abelian if and only if $[x, y] \in N$ for all $x, y \in G$. $\square$

Definition. Let G be a group, $X, Y \leq G$. We define the *commutator* $[X, Y]$ of X and Y to be the subgroup of G generated by all commutators $[x, y]$ for $x \in X$, $y \in Y$; symbolically:

$$[X, Y] = \langle [x, y] : x \in X, y \in Y \rangle$$

In particular, $[G, G]$ is the *derived group* of G, often written G'.

Warning. Not every element of $[X, Y]$ needs to be a commutator. Every element is a product of commutators and their inverses.

Remark.

(1) We have that $[x, y]^{-1} = [y, x]$ for all $x, y \in G$.
(2) From (1), $[X, Y] = [Y, X]$ for all $X, Y \subseteq G$.
(3) If $N \trianglelefteq G$, then $[X, N] \leq N$ for all $X \leq G$. Indeed, $xnx^{-1}n = {}^{x}nn^{-1} \in N$, so all the commutators $[x, n]$ are contained in N, and these generate $[X, N]$.

Proposition 30. *Suppose that $X, Y \trianglelefteq G$. Then $[X, Y] \trianglelefteq G$.*

Proof. Every element of $[X, Y]$ is a product of commutators and their inverses. So a general element looks like

$$z = [x_1, y_1]^{\varepsilon_1}[x_2, y_2]^{\varepsilon_2} \dots [x_k, y_k]^{\varepsilon_k}$$

where $x_i \in X$, $y_i \in Y$ and $\varepsilon_i \in \{\pm 1\}$. For $g \in G$, we have

$$\begin{aligned} {}^{g}z &= g[x_1, y_1]^{\varepsilon_1}g^{-1}g[x_2, y_2]^{\varepsilon_2}g^{-1} \dots g[x_k, y_k]^{\varepsilon_k}g^{-1} \\ &= [{}^{g}x_1, {}^{g}y_1]^{\varepsilon_1}[{}^{g}x_2, {}^{g}y_2]^{\varepsilon_2} \dots [{}^{g}x_k, {}^{g}y_k]^{\varepsilon_k} && \text{since } g[x, y]g^{-1} = gxyx^{-1}y^{-1}g^{-1} = [{}^{g}x, {}^{g}y] \\ &\in [X, Y] && \text{since } {}^{g}x_i \in X \text{ and } {}^{g}y_i \in Y \text{ for all } i \end{aligned}$$

which shows that $[X, Y] \trianglelefteq G$. $\square$

Recall that $G' = [G, G]$, the derived group. By Proposition 30, we see that $G' \trianglelefteq G$. Now, Proposition 29 tells us that G' is contained in every normal subgroup $N \trianglelefteq G$ such that G/N is abelian. So G' is the **smallest** normal subgroup with this property (that G/G' is abelian).

Examples.

(1) If A is abelian, then $A/\{e\}$ is abelian. So $A' = \{e\}$.
(2) We know that S_4' is a normal subgroup of S_4, so one of $\{e\}$, V_4, A_4, S_4. We have $\frac{S_4}{V_4} \cong S_3$, which is not abelian. But $\frac{S_4}{A_4} \cong C_2$, which is abelian. Hence $S_4' = A_4$.
(3) The normal subgroups of A_4 are $\{e\}$, V_4, A_4. Now, $\frac{A_4}{\{e\}}$ is not abelian, but $\frac{A_4}{V_4} \cong C_3$, so $A_4' \cong V_4$.
(4) In general, S_n ($n \geq 5$) has normal subgroups $\{e\}, A_n, S_n$ and no others (for the proof that A_n is simple for $n \geq 5$, see Appendix A). Clearly, $\frac{S_n}{\{e\}}$ is not abelian, but $\frac{S_n}{A_n} \cong C_2$ is abelian. Thus $S_n' = A_n$.
Furthermore, A_n has no normal subgroups except $\{e\}, A_n$. Since $\frac{A_n}{\{e\}}$ is not abelian, we have $A_n' = A_n$.

(5) Let $G = D_{2n}$ and $H = \langle h \rangle$ be the rotation subgroup. Since $G/H \cong C_2$, we have $G' \leq H$. Now, let x be a reflection. We have $[h, x] = hxh^{-1}x^{-1} = h\,{}^{x}(h^{-1}) = h^2$, so $\langle h^2 \rangle \leq G'$.

If n is odd, then $\langle h \rangle = \langle h^2 \rangle$, so $G' = H \cong C_n$.

If n is even, then $\left|\frac{G}{\langle h^2 \rangle}\right| = 4$, so $\frac{G}{\langle h^2 \rangle}$ is abelian. Hence $G' = \langle h^2 \rangle \cong C_{n/2}$ in this case.

Definition. Let $H \leq G$. We say that H is a *characteristic* subgroup if $\alpha(H) = H$ for every $\alpha \in \operatorname{Aut}(G)$. We write $H \operatorname{char} G$.

(This is a strengthening of the normality condition.)

Examples. We have:

- $A_n \operatorname{char} S_n$ for all n. (The only subgroup of index 2.)
- $V_4 \operatorname{char} S_4$. (The only normal subgroup of order 4).
- $\langle (12)(34) \rangle \trianglelefteq V_4$, but $\langle (12)(34) \rangle$ is not characteristic, since V_4 has an automorphism α such that $\alpha(12)(34) = (13)(24)$.

Proposition 31. *If $X, Y \operatorname{char} G$, then $[X, Y] \operatorname{char} G$.*

Proof. Essentially the same as the proof of Proposition 30. □

If follows from Proposition 31 that $G' \operatorname{char} G$.

Proposition 32. *Let G be a group.*

(1) *If $N \trianglelefteq G$ and $X \operatorname{char} N$, then $X \trianglelefteq G$.*
(2) *If $N \operatorname{char} G$ and $X \operatorname{char} N$, then $X \operatorname{char} G$.*

Proof. Let $\alpha \in \operatorname{Aut}(G)$ such that $\alpha(N) = N$. Then $\alpha_{|N}$ is an automorphism of N. Now, if $X \operatorname{char} N$, then $\alpha_{|N}(X) = X$, so $\alpha(X) = X$.

(For (1), take $\alpha \in \operatorname{Inn}(G)$. For (2), take $\alpha \in \operatorname{Aut}(G)$.) □

Warning. We can have $N \operatorname{char} G$ and $X \trianglelefteq N$ without having $X \trianglelefteq G$. For example, take $G = S_4$, $N = V_4$, $X = \langle (12)(34) \rangle$.

Definition. The *derived series* of a group G is the sequence of subgroups $\left(G^{(i)}\right)_{i \geq 0}$ defined by $G^{(0)} = G$, $G^{(i+1)} = [G^{(i)}, G^{(i)}]$.

Proposition 33. *We have that $G^{(i)} \operatorname{char} G$ for all i.*

Proof. The proof goes by induction on i. Certainly, $G^{(0)} \operatorname{char} G$. Suppose that $G^{(i)} \operatorname{char} G$. Then $[G^{(i)}, G^{(i)}] \operatorname{char} G$ by Proposition 31. Thus $G^{(i+1)} \operatorname{char} G$. □

Examples.

(1) Let G be abelian. Then the derived series of G is

$$G \geq \{e\} \geq \{e\} \geq \cdots$$

(2) The group S_4 has the derived series

$$S_4 \geq A_4 \geq V_4 \geq \{e\} \geq \cdots$$

The derived series for A_4 is obtained by cutting this series at A_4 instead.

$$A_4 \geq V_4 \geq \{e\} \geq \cdots$$

(3) The group D_{2n} has the derived series

$$D_{2n} \geq H \geq \{e\} \geq \cdots$$

where, as we have seen before:

$$H = D'_{2n} = \begin{cases} C_n & \text{if } n \text{ odd} \\ C_{n/2} & \text{if } n \text{ even} \end{cases}$$

(4) For $n \geq 5$, S_n has the derived series

$$S_5 \geq A_n \geq A_n \geq \cdots$$

Proposition 34. *A group G is soluble if and only if $\{e\}$ appears in its derived series.*

Proof. Recall that G is soluble if there exists

$$G \geq G_0 \geq \cdots \geq G_k = \{e\}$$

such that $G_{i+1} \trianglelefteq G_i$ and $\frac{G_i}{G_{i+1}}$ is abelian for all i. Suppose $\{e\} = G^{(k)}$ for some k. Then

$$G = G^{(0)} \geq G^{(1)} \geq \cdots \geq G^{(k)} = \{e\}$$

with $G^{(i+1)} \trianglelefteq G^{(i)}$, and $G^{(i)}/G^{(i+1)}$ is abelian, since $G^{(i+1)}$ is the derived group of $G^{(i)}$. Hence G is soluble.

For the converse, suppose that G is soluble, and $G = G_0 \geq G_1 \geq \cdots \geq G_k = \{e\}$ with $G_{i+1} \trianglelefteq G_i$ and $\frac{G_i}{G_{i+1}}$ abelian. We claim that

$$G^{(i)} \leq G_i$$

for all i. The proof goes by induction on i. Certainly, $G^{(0)} = G \leq G = G_0$. Now, suppose that $G^{(i)} \leq G_i$. We know that $\frac{G_i}{G_{i+1}}$ is abelian, so $G'_i \leq G_{i+1}$. But

$$G^{(i+1)} = [G^{(i)}, G^{(i)}] \leq [G_i, G_i] = G'_i \leq G_{i+1}$$

and the induction is complete.

We have shown that $G^{(i)} \leq G_i$ for all i, and, in particular, $G^{(k)} \leq G_k = \{e\}$. □

This means that for soluble groups, the derived series have finite length. Thus, to prove facts about soluble groups, we could use induction on the length of the series.

6. The Lower Central Series and nilpotent groups

Definition. Let G be a group. The *lower central series* (LCS) for G is the sequence of subgroups $(\gamma_i(G))_{i\geq 1}$ defined by $\gamma_1(G) = G$ and $\gamma_{i+1}(G) = [\gamma_i(G), G]$.

There are also *upper central series* and all the corresponding results hold for them by analogy, but they will not be discussed in this course. They were covered in the 2012 course, so the interested reader is encouraged to refer to lecture notes from that year.

Remark.

(1) Note that the LCS starts at 1, not at 0.
(2) It is clear that $\gamma_1(G) = G = G^{(0)}$ and $\gamma_2(G) = G' = G^{(1)}$. Beyond these terms, the LCS and the derived series are generally different.

Examples.

(1) If G is abelian, then the LCS is
$$G \geq \{e\} \geq \{e\} \geq \cdots$$
(2) Let $G = D_{2n}$. Then $\gamma_1(G) = G$, and $\gamma_2(G) = G'$, which is a group of rotations. Let x be any rotation, y any reflection. Then $[x, y] = xyx^{-1}y^{-1} = x^2$, so $[\langle x\rangle, G] = \langle x^2\rangle$. Clearly,
$$\langle x^2\rangle = \begin{cases} \langle x\rangle & \text{if } \mathrm{ord}(x) \text{ odd} \\ \text{index 2 subgroup of } \langle x\rangle & \text{if } \mathrm{ord}(x) \text{ even} \end{cases}$$
For example, we get the LCS:
$$D_{24} \geq C_6 \geq C_3 \geq C_3 \geq \cdots$$
$$D_{16} \geq C_4 \geq C_2 \geq C_1 = \{e\} \geq \{e\} \geq \cdots$$

Definition. A group G is *nilpotent* if $\{e\}$ appears in its LCS. If $\gamma_{c+1}(G)$ is the first term of the LCS equal to $\{e\}$, then we say that G has *nilpotency class* c.

Examples.

(1) As we have seen above, D_{24} is not nilpotent, but D_{16} is nilpotent of class 3.
(2) The trivial group $\{e\}$ is the unique group of nilpotency class 0.
(3) A group is nilpotent of class 1 if and only if it is a non-trivial abelian group.

Proposition 35. *Every nilpotent group is soluble.*

Proof. We show inductively that $G^{(i)} \leq \gamma_{i+1}(G)$ for all i. The base case $i = 0$ is trivial. Suppose that $G^{(i)} \leq \gamma_{i+1}(G)$. We have
$$G^{(i+1)} = [G^{(i)}, G^{(i)}] \leq [\gamma_{i+1}(G), G],$$
since $G^{(i)} \leq \gamma_{i+1}(G)$ by the inductive hypothesis, and $G^{(i)} \leq G$. But $[\gamma_{i+1}(G), G] = \gamma_{i+2}(G)$, so $G^{(i+1)} \leq \gamma_{i+2}(G)$, completing the induction. Now, if G is nilpotent, then $\gamma_{c+1}(G) = \{e\}$ for some c, but now $G^{(c)} = \{e\}$, so G is soluble. $\square$

Remark. There exist groups which are soluble but not nilpotent—we have seen one, D_{24}. There are smaller examples, such as S_3 which has LCS $S_3 \geq A_3 \geq A_3 \geq \cdots$.

Proposition 36. *Let $N \trianglelefteq G$. Then $[N,G]$ is the smallest normal subgroup H of G, contained in N, such that $\frac{N}{H} \leq Z\left(\frac{G}{H}\right)$.*

Proof. We know that $[N,G] \trianglelefteq G$ by Proposition 30, and certainly $[N,G] \leq N$, since N is normal. For any $x \in N$, we have $xH \in Z\left(\frac{G}{H}\right)$ if and only if $[xH, gH] = e_{G/H}$ for all $g \in G$, which is equivalent to $[x,g] \in H$. Therefore, $\frac{N}{H} \leq Z\left(\frac{G}{H}\right)$ if and only if $[x,g] \in H$ for all $x \in N$, $g \in G$, which is equivalent to $[N,G] \leq H$. So clearly $[N,G]$ is the smallest subgroup H with this property. $\square$

While Proposition 36 is technical, it is very useful incomputing the LCS of some groups.

Examples.

(1) The LCS for S_4 begins $\gamma_1(S_4) = S_4$, $\gamma_2(S_4) = S_4' = A_4$. Now, $\gamma_3(S_4)$ is normal in S_4, and contained in A_4, so one of $\{\text{id}\}, V_4, A_4$. But $\frac{A_4}{\{\text{id}\}} \nleq Z\left(\frac{S_4}{\{\text{id}\}}\right)$ and $\frac{A_4}{V_4} \nleq Z\left(\frac{S_4}{V_4}\right)$, because $\frac{S_4}{V_4} \cong S_3$ (since $Z(S_4)$ and $Z(S_3)$ are trivial). So $\gamma_3(S_4) = A_4$. (Hence $\gamma_{i+1}(S_4) = A_4$ for $i > 0$.)

(2) We have $\gamma_1(A_4) = A_4$, $\gamma_2(A_4) = A_4' = V_4$. Now, $\gamma_3(A_4)$ is normal in A_4, and a subgroup of V_4, so one of $\{\text{id}\}$ or V_4. But $\frac{V_4}{\{\text{id}\}} \nleq Z\left(\frac{A_4}{\{\text{id}\}}\right)$, so $\gamma_3(A_4) = V_4$.

Remark. Note that, unlike the derived series, we cannot get the LCS for $\gamma_i(G)$ get the LCS for $\gamma_i(G)$ just by truncating the LCS for G.

Proposition 37. *We have that $\gamma_i(G) \operatorname{char} G$ for all i.*

Proof. Similar to the proof of Proposition 33. $\square$

Proposition 38. *For any $i, j \in \mathbb{N}$, we have*

$$[\gamma_i(G), \gamma_j(G)] \leq \gamma_{i+j}(G).$$

Note that a generator of $\gamma_i(G)$ looks like $[\dots[[[g_1, g_2], g_3], g_4]\dots, g_i]$.

The proof of this proposition is omitted. It was proved in the course in 2012, so the interested reader is encouraged to refer to lecture notes from that year for the proof.

Proposition 38 is one reason why we begin the LCS at 1 instead of 0.

Proposition 39. *Let $\varphi\colon G \to H$ be a surjective homomorphism. Then $\gamma_i(H) = \varphi(\gamma_i(G))$.*

An analogous result also holds for the derived series; however, it will not be used anywhere in this course, which is why it was not stated and proved before.

Proof. The proof goes by induction on i. The base case $i = 1$ is clear. Suppose that $\varphi(\gamma_i(G)) = \gamma_i(H)$ for some i. We have $\gamma_{i+1}(H) = [\gamma_i(H), H]$. So a general element of $\gamma_{i+1}(H)$ is $y = [x_1, h_1]^{\varepsilon_1} \dots [x_t, h_t]^{\varepsilon_t}$ with $x_j \in \gamma_i(H)$, $h_j \in H$, $\varepsilon_j = \pm 1$. For each j, let $z_j \in \gamma_i(G)$ be such that $\varphi(z_j) = x_j$, and g_j such that $\varphi(g_j) = h_j$. Now

$$y = \varphi\left([z_1, g_1]^{\varepsilon_1} \dots [z_t, g_t]^{\varepsilon_t}\right) = \varphi(w)$$

for some $w \in [\gamma_i(G), G] = \gamma_{i+1}(G)$. So $\gamma_{i+1}(H) \leq \varphi(\gamma_{i+1}(G))$. The reverse containment is similar. $\square$

Theorem 40. *Every p-group is nilpotent.*

Proof. The proof goes by induction on the order of the group. The base case $|P| = 1$ is trivial. For the inductive step, suppose that all groups of order p^a are nilpotent. Suppose $|P| = p^{a+1}$. Then P is a non-trivial p-group, so it has a normal central subgroup N of order p by Proposition 22. Now, $\left|\frac{P}{N}\right| = p^a$, so $\frac{P}{N}$ is nilpotent by the inductive hypothesis. Thus $\gamma_c\left(\frac{P}{N}\right) = \{e\}$ for some c. Now, if $\theta\colon P \to \frac{P}{N}$ is the canonical map, then by Proposition 39, we have

$$\gamma_{c+1}\left(\frac{P}{N}\right) = \theta(\gamma_{c+1}(P)).$$

So $\gamma_{c+1}(P) \leq \operatorname{Ker}\theta = N$. So either $\gamma_{c+1}(P) = \{e\}$, or else $\gamma_{c+1}(P) = N$. But in the latter case, $\gamma_{c+2}(P) = [N, P]$, which is $\{e\}$, since $N \leq Z(P)$. □

Example. We have seen that in general dihedral groups are not nilpotent. But if $n = 2^a$, then D_{2n} is a 2-group, so in this case D_{2n} is nilpotent. (We saw that D_{16} is nilpotent earlier.)

Proposition 41. *Let G be nilpotent of class c. Then*

1. *Every subgroup of G is nilpotent of class at most c.*
2. *If $N \trianglelefteq G$, then $\frac{G}{N}$ is nilpotent of class at most c.*
3. *If H is nilpotent of class d, then $G \times H$ is nilpotent of class* $\max(c, d)$.

Proof. To show (1), we show by an easy induction that if $L \leq G$, then $\gamma_i(L) \leq \gamma_i(G)$ for all i. In particular, $\gamma_{c+1}(L) = \{e\}$.

For (2), we consider the canonical map $\theta\colon G \to \frac{G}{N}$ and use Proposition 39 to obtain

$$\gamma_{c+1}\left(\frac{G}{N}\right) = \theta(\gamma_{c+1}(G)) = \theta(\{e\}) = \{e_{G/N}\}.$$

For (3), notice that $[(g_1, h_1), (g_2, h_2)] = ([g_1, g_2], [h_1, h_2])$ for all $g_1, g_2 \in G$, $h_1, h_2 \in H$. Now, an easy induction shows that

$$\gamma_i(G \times H) = \gamma_i(G) \times \gamma_i(H).$$

If $m = \max(c, d)$, then

$$\gamma_{m+1}(G \times H) = \gamma_{m+1}(G) \times \gamma_{m+1}(H) = \{e_G\} \times \{e_H\} = \{e_{G\times H}\},$$

which completes the proof. □

Example. Let $A = \langle a \rangle \cong C_{12}$ and $B = \langle b \rangle \cong C_2$. Let $G = A \rtimes_\varphi B$, where

$$\varphi_b(a) = a^7.$$

Is G nilpotent? Write $A = A_3A_4$ where $A_3 = \langle a^4 \rangle \cong C_3$ and $A_4 = \langle a^3 \rangle \cong C_4$. Identify A with $\{(a, e) : a \in A\}$ and B with $\{(e, b) : b \in B\}$. So $G = AB = A_3A_4B$. Notice that $A_4 \operatorname{char} A \trianglelefteq G$, so $A_4 \trianglelefteq G$. Hence A_4B is a subgroup of order 8. Hence A_4B is nilpotent. Moreover, A_3 is also nilpotent, and A_3 and A_4B are complements in G.

So certainly $G \cong A_3 \rtimes A_4B$. We show that actually $G \cong A_3 \times A_4B$. We have $A_3 = \langle a^3 \rangle$, so it is enough to show that $A_4B \leq \operatorname{Cent}_G(a^4)$. Certainly, $A_4 \leq A \leq \operatorname{Cent}_G(a^4)$, and we have that $\varphi_b(a^4) = (a^4)^7 = a^{28} = a^4$, so the $B \leq \operatorname{Cent}_G(a^4)$. Hence $A_4B \leq \operatorname{Cent}_G(a^4)$.

Therefore, $G \cong A_3 \times A_4 B$, so it is nilpotent by Proposition 41 (3).

Proposition 42 (Frattini argument). *Let G be a finite group. For $K \trianglelefteq G$ and $P \in \mathrm{Syl}_p(K)$, we have that $G = KN_G(P)$.*

Proof. Let $g \in G$. Then ${}^gP \leq {}^gK = K$. Hence ${}^gP \in \mathrm{Syl}_p(K)$. But the Sylow p-subgroup of K are all conjugate in K, so ${}^gP = {}^kP$ for some $k \in K$. Now, ${}^{k^{-1}g}P = P$, so $k^{-1}g \in N_G(P)$. Thus we have $g = k(k^{-1}g) \in KN_G(P)$. □

Definition. A subgroup M of G is *maximal* if $M \neq G$, and there exists no subgroup of H with $M < H < G$.

Fact. *If G is finite then every proper subgroup of G is contained in at least one maximal subgroup.*

(If H is contained in no maximal subgroup, then there exist an infinite ascending chain $H = H_0 < H_1 < \cdots$, and so G must be infinite.)

Theorem 43. *Let G be a finite group. Then the following are equivalent:*

(1) *G is nilpotent,*
(2) *$H < N_G(H)$ (strict inequality) for every proper subgroup H,*
(3) *$M \lhd G$ for every maximal subgroup M,*
(4) *$P \trianglelefteq G$ for every Sylow subgroup P,*
(5) *G is isomorphic to a direct product of p-groups,*
(6) *any two elements of G with coprime orders commute.*

Normally, we would prove the consecutive implications in a cycle; however, it is hard to structure the proof that way. Instead, the proof will be structured as follows:

$$\begin{tikzcd}
(1) \arrow[r] & (2) \arrow[r] & (3) \arrow[r] & (4) \arrow[r] & (5) \arrow[llll, bend right] \arrow[dl] \\
& & & (6) \arrow[u] &
\end{tikzcd}$$

Proof. (1) implies (2). Let G be nilpotent of class c and let $H < G$. Then $H \ngeq \gamma_1(G)$, but $H \geq \gamma_{c+1}(G)$. So there must exist j such that $\gamma_j(G) \nleq H$, but $\gamma_{j+1}(G) \leq H$.

Now, $\gamma_j(G) \trianglelefteq G$, and so $H\gamma_j(G) \leq G$. Since $\frac{\gamma_j(G)}{\gamma_{j+1}(G)} \leq Z\left(\frac{G}{\gamma_{j+1}(G)}\right)$,

$$\frac{H\gamma_j(G)}{\gamma_{j+1}(G)} \leq \frac{H}{\gamma_{j+1}(G)} \frac{\gamma_j(G)}{\gamma_{j+1}(G)} \leq \frac{H}{\gamma_{j+1}(G)} Z\left(\frac{G}{\gamma_{j+1}(G)}\right).$$

and so

$$\frac{H}{\gamma_{j+1}(G)} \trianglelefteq \frac{H\gamma_j(G)}{\gamma_{j+1}(G)}.$$

Hence H is normal in $H\gamma_j(G)$ by the Subgroup Correspondence Theorem 2 (2). So $N_G(H) \geq H\gamma_j(G) \nleq H$ (since $\gamma_j(G) \nleq H$). Hence $N_G(H) > H$.

(2) implies (3). Suppose $H < N_G(H)$ for all $H < G$. Let M be maximal. Then $M < N_G(M) \leq G$. So by maximality of M, we have $N_G(M) = G$.

(3) implies (4). Let $P \in \mathrm{Syl}_p(G)$. Suppose (for a contradiction) that P is not normal. Then $N_G(P) < G$, and so $N_G(P) \leq M$ for some maximal M. Now, $M \lhd G$ by (3), so, by the Frattini argument 42, we have $G = MN_G(P) \leq MM = M$, a contradiction. So $P \trianglelefteq G$.

(4) implies (5). Suppose that the Sylow subgroups of G are normal. Let $p_1, \dots, p_k$ be the prime divisors of $|G|$, and let $P_1, \dots, P_k$ be the corresponding Sylow subgroups.

We show by induction that for $1 \leq j \leq k$ we have $P_1P_2 \dots P_j$ is a normal subgroup of G isomorphic to $P_1 \times P_2 \times \cdots \times P_j$. The base case, $j = 1$, is obvious. Suppose true for j. Then

$$P_1 \dots P_j P_{j+1} = (P_1 P_2 \dots P_j)P_{j+1}.$$

This is a subgroup since $P_{j+1} \trianglelefteq G$. Now, $P_1 \dots P_j$ is normal, and P_{j+1} is normal, so $(P_1 \dots P_j)P_{j+1}$ is normal, and isomorphic to $(P_1 \dots P_j) \times P_{j+1}$, which is isomorphic to $P_1 \times \cdots \times P_{j+1}$ by the inductive hypothesis.

Therefore, we have $|P_1 \dots P_k| = |G|$, so $P_1 \times \cdots \times P_k \cong P_1 \dots P_k = G$, as required.

(5) implies (1). We have that every p-group is nilpotent and any direct product of nilpotent groups is nilpotent (Theorem 40 and Proposition 41 (3)). So if G is isomorphic to a direct product of p-groups then G is nilpotent.

(5) implies (6). Since $G \cong P_1 \times \cdots \times P_k$ (where P_i is a p_i-group, for distinct primes p_i), we can write an element g of G as $g = (g_1, \dots, g_k)$ where $g_i \in P_i$ for all i. Let $h = (h_1, \dots, h_k)$ be another element. Now, $\mathrm{ord}(g) = \mathrm{ord}(g_1) \dots \mathrm{ord}(g_k)$ and $\mathrm{ord}(h) = \mathrm{ord}(h_1) \dots \mathrm{ord}(h_k)$. So g and h have coprime orders if and only if, for all i, we have at least one of g_i or h_i equal to e. But in this case, it is clear that g and h commute.

(6) implies (4). Let $|G| = n = p_1^{a_1} \dots p_k^{a_k}$. Let P_i be a Sylow p_i-subgroup. Certainly, $P_i \leq N_G(P_i)$. But if $j \neq i$, then every element of P_j (a Sylow p_j-subgroup) commutes with every element of P_i (by (6)). So $P_j \leq N_G(P_i)$ for all j. Now, $p_j^{a_j}$ divides $|N_G(P_i)|$ for all j, and so $|N_G(P_i)| = |G|$. Thus $N_G(P_i) = G$, and $P_i \trianglelefteq G$. □

7. More on actions

Let G act on a set Ω and $k \leq |G|$. Define

$$\Omega^{(k)} = \{(x_1, \dots, x_k) \mid x_i \in \Omega, x_i \neq x_j \text{ if } i \neq j\},$$

the set of k-tuples of distinct elements of Ω.[12]

There is an action of G on $\Omega^{(k)}$ given by

$$g(x_1, \dots, x_k) = (gx_1, \dots, gx_k).$$

(Notice that if $x_i \neq x_j$ then $gx_i \neq gx_j$.)

[12] The notation $\Omega^{(k)}$ is not standard and will not be found in literature. There is no *standard* notation for this set.

Definition. We say that G acts *k-transitively* on Ω if the action of G on $\Omega^{(k)}$ is transitive. In other words, G is k-transitive if any k-tuple of distinct points can be mapped to any other by an element of G.

Note that the assumption that the k-tuples are distinct is really necessary; otherwise, no actions would be k-transitive, because we could not send $(x, x, \ldots, x)$ to $(x, y, \ldots, y)$ if $x \neq y$. Note also that if $k > |G|$, then the action of G on $\Omega^{(k)}$ can never be transitive, hence the assumption $k \leq |G|$.

Remark. It is clear that if G is k-transitive, then it is l-transitive for all $l \leq k$.

Examples.

(1) The group S_n is n-transitive on $\{1, 2, \ldots, n\}$. If $x = (x_1, \ldots, x_n)$ is a tuple of distinct points, then $\{x_1, \ldots, x_n\} = \{1, \ldots, n\}$, so $\sigma_x : i \mapsto x_i$ is a permutation. Hence $\sigma_x \in S_n$ and $x = \sigma_x(1, 2, \ldots, n)$, and so $\mathrm{Orb}_{S_n}(1, \ldots, n) = \Omega^{(n)}$.

(2) The group A_n is only $(n-2)$-transitive on $\{1, \ldots, n\}$. There is no element of A_n which maps

$$(1, 2, \ldots, n-2, n-1) \mapsto (1, 2, \ldots, n-2, n),$$

since the only element of S_n which does this is $(n-1, n)$. So A_n is not $(n-1)$-transitive.

To show that A_n is $(n-2)$-transitive, let $x = (x_1, \ldots, x_{n-2})$ and $y = (y_1, \ldots, y_{n-2})$ be tuples of distinct points. Let x_{n-1} and x_n be the two points not in x, and y_{n-1} and y_n be two points not in y. Since S_n is n-transitive, there exists $\sigma, \tau \in S_n$ such that

$$\sigma(x_1, \ldots, x_n) = (y_1, \ldots, y_{n-2}, y_{n-1}, y_n),$$
$$\tau(x_1, \ldots, x_n) = (y_1, \ldots, y_{n-2}, y_n, y_{n-1}).$$

Now, $\tau = (y_{n-1}y_n)\sigma$, so σ and τ have different signatures, so one of them (say σ) is in A_n. Then $\sigma x = y$, so x and y are in the same orbit of A_n.

(3) If $n > 3$, then D_{2n} is only 1-transitive on the vertices of an n-gon. Given a vertex u, we can find vertices v, w such that u and v are adjacent but u and w are not. Now, no element of D_{2n} maps (u, v) to (u, w), because the action of D_{2n} preserves the adjacency relation.

Not many finite groups are highly transitive. In fact, the only finite groups with 4-transitive actions are S_n for $n \geq 4$, A_n for $n \geq 6$, and four other groups. These are known as the Mathieu Groups, M_{11}, M_{12}, M_{23}, M_{24}, which are subgroups of S_{11}, S_{12}, S_{23}, S_{24} respectively. They are simple groups (in the classification of finite simple groups, they probably the easiest examples of sporadic groups). Moreover, M_{12} and M_{24} are 5-transitive.

Remark. If G acts on Ω and $H = \mathrm{Stab}_G(x)$ for $x \in \Omega$, then H acts on $\Omega \setminus \{x\}$ by restriction.

Proposition 44. *Let G act transitively on Ω, and $H = \mathrm{Stab}_G(x)$ for $x \in \Omega$. Let $k \in \mathbb{N}$. Then G is k-transitive on Ω if and only if H is $(k-1)$-transitive on $\Omega \setminus \{x\}$.*

Note that using this proposition, we can get an inductive prove that S_n is n-transitive on $\{1, \ldots, n\}$; indeed, the stabilizer of a point is S_{n-1}.

Proof. We first prove the 'only if' implication. Suppose that G is k-transitive on Ω. Let $y = (y_1, \dots, y_{n-1})$, $z = (z_1, \dots, z_{n-1})$ be tuples of $(k-1)$ distinct elements of $\Omega \setminus \{x\}$. Then $y' = (y_1, \dots, y_{n-1}, x)$ and $z' = (z_1, \dots, z_{n-1}, x)$ are elements of $\Omega^{(k)}$. Since G is transitive on $\Omega^{(k)}$, there is some $g \in G$ such that $gy' = z'$. So we have $gy = z$ and $gx = x$, and hence $g \in H$. Therefore, y and z are in the same orbit of H. So H is $(k-1)$-transitive on $\Omega \setminus \{x\}$.

For the 'if' implication, suppose that H is $(k-1)$-transitive on $\Omega \setminus \{x\}$. Let $y, z \in \Omega^{(k)}$, $y = (y_1, \dots, y_k)$, $z = (z_1, \dots, z_k)$. Since G is transitive, there exist $f, g \in G$ such that $fy_k = x$, $gz_k = x$. So

$$fy = (fy_1, \dots, fy_{k-1}, x),\ gz = (gz_1, \dots, gz_{k-1}, x).$$

Now, $fy, gz \in \Omega^{(k)}$, so $y' = (fy_1, \dots, fy_{k-1})$ and $z' = (gz_1, \dots, gz_{k-1})$ are $(k-1)$-tuples are distinct points from $\Omega \setminus \{x\}$. Hence there exists $h \in H$ such that $hy' = z'$. Also, $hx = x$, so $hfy = gz$. Now, $g^{-1}hfy = z$, and so y and z are in the same orbit of G. Hence G is k-transitive on Ω. ☐

Recall that a *partition* on a set Ω is a division of Ω into non-overlapping subsets $\{X_i \subseteq \Omega : i \in I\}$, where I is some indexing set, such that $X_i \cap X_j = \emptyset$ when $i \neq j$ and $\bigcup_{i \in I} X_i = \Omega$.

Partitions correspond to equivalence relations. We write $x \sim y$ if and only if x and y lie in the same part X_i.

Definition.

(1) We say that a partition of Ω is *trivial* if it has only one part, or if every part has size 1.
(2) Let G act on Ω. We say that G *preserves* the equivalence relation $\sim$ on Ω (or the corresponding partition) if $gx_1 \sim gx_2$ if and only if $x_1 \sim x_2$.

Clearly, any group acting on Ω preserves the trivial partitions.

Definition. Let G act transitively on Ω, where $|\Omega| > 1$. The action of G is *primitive* if it preserves no non-trivial equivalence relations (or partitions) on Ω.

If $\sim$ is a non-trivial relation preserved by G then the action is *imprimitive* and $\sim$ is a *system of imprimitivity* for G. In this case, the equivalence classes are *blocks* for the action.

Warning. Both primitive and imprimitive actions are transitive by definition.

Examples.

(1) Let $g = (12 \dots n) \in S_n$ and consider $G = \langle g \rangle$ acting on $\{1, 2, \dots, n\}$. The action of G is primitive if and only if n is prime.

We can assume that $n > 1$. Suppose that d is a divisor of n. Then define $\sim$ by $i \sim j$ if and only if $i \equiv j \mod d$. It is clear that $\sim$ is preserved by G and $\sim$ is non-trivial provided $1 < d < n$. Hence if n has such a divisor, i.e. n is not prime, then G is imprimitive.

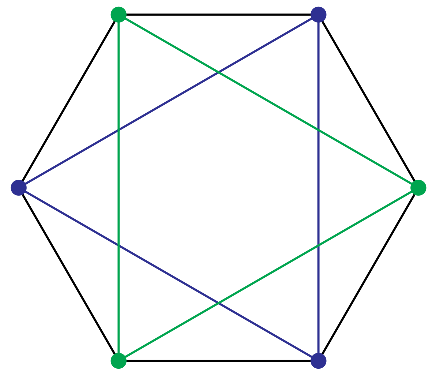
(The proof of the converse implication comes later—see Corollary 47.)

(2) The group S_n is primitive on $\{1, \dots, n\}$ for $n > 1$ and A_n is primitive on $\{1, \dots, n\}$ for $n > 2$.

Suppose we have a non-trivial relation $\sim$ on $\{1, \ldots, n\}$. Then we can find distinct i, j, k such that $i \sim j$ but $i \not\sim k$. Since S_n is 2-transitive, and A_n is 2-transitive if $n \geq 4$, they have an element g such that $g(i,j) = (i,k)$. Hence $\sim$ is not preserved under the action. (The case A_3 is as in (1) above.)

(3) The group D_{2n} acts primitively on the vertices of an n-gon if and only if n is prime.

If n is not prime, let d be a proper divisor of n. Assuming, $d > 2$, we can embed $\frac{n}{d}$ copies of d-gons in our n-gon, partitioning the vertices. In the case, $n = 6$, $d = 2$, we get 2 copies of triangles:



These form a system of imprimitivity for D_{2n}. (If $d = 2$, embed long diagonals instead.)

(The proof of the converse implication comes later—see Corollary 47.)

Proposition 45. *Let $\mathcal{B} = \{B_i : i \in I\}$ be a system of imprimitivity for the action of G on Ω. (So $\mathcal{B}$ is a partition of Ω.) For $B \in \mathcal{B}$, define $gB = \{gx : x \in B\}$, for any $g \in G$. Then $gB \in \mathcal{B}$ and the map $(g, B) \mapsto gB$ defines a transitive action of G on $\mathcal{B}$.*

Proof. Let $\sim$ be the equivalence relation corresponding to $\mathcal{B}$. Let $x \in gB$. Then $g^{-1}x \in B$. We have $x \sim y$ if and only if $g^{-1}x \sim g^{-1}y$, or, equivalently, $g^{-1}y \in B$, i.e. $y \in gB$. Hence gB is the equivalence class of x under $\sim$, so $gB \in \mathcal{B}$.

It is clear that $eB = B$ and $(g_1 g_2)B = g_1(g_2 B)$, so $(g, B) \mapsto gB$ defines an action. To show that the action on $\mathcal{B}$ is transitive, take any $B' \in \mathcal{B}$ and let $y \in B'$. Let $z \in B$. Since G is transitive on Ω, there exists $g \in G$ such that $gz = y$. So $y \in gB$, and so $gB = B'$. $\square$

Proposition 46. *Let $\mathcal{B}$ be a system of imprimitivity for the action of G on Ω. Then all of the blocks of $\mathcal{B}$ have the same size.*

Proof. Let B_1 and B_2 be blocks of $\mathcal{B}$. Then there exists $g \in G$ such that $B_2 = gB_1$ by Proposition 45. But $|gB_1| = |B_1|$, and hence $|B_2| = |B_1|$. $\square$

Corollary 47. *Let Ω be a set such that $|\Omega|$ is prime. Then any transitive group action on Ω is primitive.*

Proof. Suppose we have k blocks in a system of imprimitivity. They all have the same size, m. So $|\Omega| = km$, but $k, m > 1$, so $|\Omega|$ is composite. $\square$

This establishes the claim from before that the actions of C_p and D_{2p} on the vertices of a regular p-gon are primitive for p prime.

Remark. Primitivity implies transitivity (as part of the definition). It is a strictly stronger property, since, for example, C_4 acting on $\{1,2,3,4\}$ is transitive but not primitive.

We show next that primitivity is strictly weaker than 2-transitivity.

Proposition 48. *Let G act 2-transitively on Ω. Then the action is primitive.*

The proof will be analogous to the proof that the action of S_n on $\{1,\dots,n\}$ for $n>1$ is primitive.

Proof. Let $\sim$ be a non-trivial equivalence relation on Ω. There exists distinct $x,y,z\in\Omega$ such that $x\sim y$, $x\not\sim z$. But since G is 2-transitive, there exists $g\in G$ such that $g(x,y)=(x,z)$. Therefore, $\sim$ is not preserved by G, and hence G acts primitively on Ω. □

So 2-transitivity is at least *weakly stronger* than primitivity. But C_3 acting on $\{1,2,3\}$ is primitive but not 2-transitive, so in fact 2-transitivity is strictly stronger.

Even primitivity is a strong condition on a group action. It can be shown that for "*almost all*" $n\in\mathbb{N}$, the only primitive subgroups of S_n are S_n and A_n. (In particular, the other families of primitive groups that we have seen, such as C_p and D_{2p}, do not give a contribution, asymptotically. This is because the density of the primes in $\mathbb{N}$ tends to 0.)

Proposition 49. *Let G act transitively on Ω, and let $x\in\Omega$. Let $H=\mathrm{Stab}_G(x)$. Then the action is primitive if and only if H is maximal in G.*

Therefore, by Theorem 12, studying primitive actions is equivalent to studying groups and their maximal subgroups.

Proof. For the 'if' implication, we show that if G is imprimitive then H is not maximal. Suppose that $\mathcal{B}$ is a system of imprimitivity for the action of G. Let B_x be the block of $\mathcal{B}$ containing x. Recall that G acts transitively on $\mathcal{B}$ by Proposition 45.

Let L be the stabilizer of B_x. For $h\in H$, we have $hx=x$, so $x\in B_x\cap hB_x$. But $hB_x\in\mathcal{B}$, and so $hB_x=B_x$. So $h\in L$, and hence $H\le L$. Since $\mathcal{B}$ is associated with a non-trivial equivalence relation, we have:

(1) There exists $y\neq x$ such that $y\in B_x$. Now, suppose $gx=y$ (since G is transitive on Ω). Then $y\in B_x\cap gB_x$, and so $gB_x=B_x$, and so $g\in L$. But $g\notin H$, so $H\neq L$.
(2) There exists $z\notin B_x$. If $gx=z$ (and such g exists by transitivity of G on Ω) then $gB_x\neq B_x$, so $L\neq G$.

Hence $H<L<G$, so H is not maximal.

For the 'only if' implication, recall that by Theorem 12, the action of G on Ω is equivalent to its coset action on cosets of H. We show that if H is not maximal, then G preserves a non-trivial equivalence relation on the cosets of H.

Suppose that $H<L<G$. Define relation $\sim$ by $g_1H\sim g_2H$ if and only if $g_1L=g_2L$. (Note that $g_1H=g_2H$ if and only if $g_2^{-1}g_1\in H$, so $g_2^{-1}g_1\in L$, and hence $g_1L=g_2L$. Therefore, this relation is well-defined, and clearly it is an equivalence relation.) Since $H<L$, there

exists $l \in L \setminus H$. Now, $H \neq lH$, but $H \sim lH$ since $eL = lL$. So $\sim$ has an equivalence class of size greater than 1. Also, $L < G$, so there exists $g \in G \setminus L$. Now, $H \not\sim gH$, since $eL \neq gL$. Hence $\sim$ is non-trivial.

Finally, for $x, y, g \in G$, we have $gxH \sim gyH$ if and only if $gxL = gyL$ if and only if $y^{-1}x = (gy)^{-1}gx \in L$, which is equivalent to $xL = yL$, i.e. $xH \sim yH$. Hence the action of G preserves $\sim$, and so it is imprimitive. $\square$

Example. Let C_n act on $\{1, \ldots, n\}$. This action is transitive and the point stabilizers are trivial. Now $\{e\}$ is maximal in G if and only if G has no subgroups except $\{e\}$, which holds if and only if $G \cong C_p$ for some prime p.

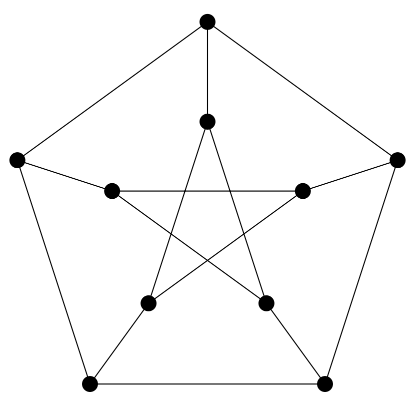
Therefore, C_n acts primitively if and only if n is prime. We have seen a proof of this before, but the new theory gives an alternative proof.

Example. What are the primitive actions of S_5? Look at maximal subgroups of S_5. These are (up to conjugacy):

(1) A_5
(2) S_4 (E.g. $\mathrm{Stab}_{S_5}(5)$)
(3) $C_5 \rtimes C_4$, where C_4 is acting as the full automorphism group of C_5. (E.g. take the normalizer in S_5 of $\langle(12345)\rangle$. Note that $(13524) = (12345)^2$ and (2354) conjugates (12345) to (13524), so (2354) is in this normalizer.)
(4) $S_3 \times S_2$ (E.g. $\mathrm{Sym}(\{1, 2, 3\}) \times \mathrm{Sym}(\{4, 5\})$.)

So S_5 has primitive actions of degrees (number of points of Ω) equal to 2, 5, 6, 10. Note that the action of 2 points is not faithful, as it has kernel A_5. The other three are faithful. So S_5 embeds as a primitive subgroup of S_6 and S_{10}.

The action on 10 points can be seen as the action of S_5 on 2-subsets of $\{1, 2, \ldots, 5\}$, i.e. $\{i, j\}$ for $i \neq j$, with the action given by $g\{i, j\} = \{gi, gj\}$. Alternatively, it can be seen as the automorphisms[13] of the Petersen graph:



Proposition 50. *If G is nilpotent and G acts primitively on Ω then $|\Omega|$ is prime, and the image of G in $\mathrm{Sym}(\Omega)$ is cyclic.*

Proof. Every maximal subgroup of a nilpotent group is normal by Theorem 43. Since G is transitive, its point stabilizers are all conjugate, so in fact G has only one point stabilizer, K,

[13] An automorphism is a permutation of the vertices which maps adjacent vertices to adjacent vertices.

which is the kernel of the action. Now, subgroups of G/K correspond to subgroups of G containing K by the Subgroup Correspondence Theorem 2. But K is maximal by Proposition 49, so G/K has no non-trivial proper subgroups. Hence $G/K \cong C_p$ for some prime p. Therefore, $|\Omega| = p$, and since G/K is isomorphic to the image of G in $\mathrm{Sym}(\Omega)$, this image is cyclic. □

Proposition 51. *Let G act faithfully and primitively on a set Ω. Let $N \neq \{e\}$ be a normal subgroup of G. Then N acts transitively on Ω.*

(If we omit the assumption of faithfulness, we could take N to be the kernel of the action, whence N acts on Ω trivially. Therefore, we could only conclude that N acts trivially or transitively on Ω.)

Proof. The orbits of the action of N on Ω form a partition of Ω. Let $\sim$ be the equivalence relation associated to this partition. (So $y \sim x$ if and only if $y = nx$ for some $n \in N$.) Since G is faithful and $N \neq \{e\}$, the orbits of N are not all of size 1.

We show that $\sim$ is preserved by G. We have $gy \sim gx$ if and only if $gy = ngx$ for some $n \in N$ which is equivalent to $y = g^{-1}ngx$ for some $n \in N$, but $g^{-1}ng \in N$, since N is normal. We have hence shown that $gy \sim gx$ if and only if $y \sim x$.

But G is primitive, so $\sim$ must be trivial. Since the parts do not all have size 1, there must be only one part. Hence N has only one orbit on Ω, so N is transitive. □

Example. We claim that any subgroup of S_7 of order 168 is simple.

Let $G \leq S_7$ have order $168 = 7 \times 3 \times 2^3$. Then G has an element of order 7 (by Cauchy's Theorem 16), which must be a 7-cycle. Hence G is transitive on $\{1, \ldots, 7\}$, and hence G is primitive. Suppose $N \trianglelefteq G$, $N \neq \{e\}, G$. Then, by Proposition 51, N is transitive on $\{1, \ldots, 7\}$. So 7 divides $|N|$, and so N contains at least one Sylow 7-subgroup of G. But N is normal, so it must contain all Sylow 7-subgroups of G.

We have that $n_7(G) \equiv 1 \bmod 7$ divides 168, so $n_7(G)$ is 1 or 8. Suppose $n_7(G) = 1$. Then G has a normal subgroup P of order 7. So $G \leq N_{S_7}(P)$. But we can easily calculate that for a subgroup $P \leq S_7$ of order 7, $|N_{S_7}(P)| = 42$. (There are 6! 7-cycles in S_7, and each 7-subgroup has 6 of them. So there are 5! such subgroups. They are Sylow 7-subgroups of S_7, so they are all conjugate. Since $N_{S_7}(P)$ is the stabilizer in the conjugacy action, we get $|N_{S_7}(P)| \times 5! = |S_7|$. So $|N_{S_7}(P)| = 7 \times 6 = 42$.) So it follows that $n_7(G) \neq 1$, so $n_7(G) = 8$.

Since all 8 Sylow 7-subgroups are contained in N, they are all conjugate in N. So 8 divides $|N|$. So $|N|$ is divisible by 56. Since there is no n such that $56|n$ and $n|168$ except $n = 56, 168$, we must have $|N| = 56$.

We have found 48 elements of order 7 in N. The remaining 8 must form a Sylow 2-subgroup Q of N. Note that N is primitive on $\{1, \ldots, 7\}$ and $Q \triangleleft N$. So Q is transitive on $\{1, \ldots, 7\}$ by Proposition 51. But this is impossible since 7 does not divide $|Q|$. So by contradiction, no such N exists. Hence G is simple.

Finally, we still have to show that S_7 has a subgroup of order 168. Let $G = \mathrm{GL}_3(2)$, the invertible 3×3 martices over $\mathbb{Z}_2$. The size of G is the number of (ordered) bases of $(\mathbb{Z}_2)^3$. We see that (v_1, v_2, v_3) is an ordered basis, provided $v_1 \neq 0$, $v_2 \notin \mathrm{Span}\{v_1\}$, $v_3 \notin \mathrm{Span}\{v_1, v_2\}$. Since $(\mathbb{Z}_2)^3$ has size 8, we see that the number of bases is $(8-1)(8-2)(8-4) = 168$.

Note that G acts on the non-zero vectors of $(\mathbb{Z}_2)^3$. This action gives a homomorphism $G \to S_7$. Since the kernel of the action is trivial, this homomorphism is injective, and so G is isomorphic to a subgroup of S_7 of order 168. By the claim above, this subgroup is simple.

We have now seen the two smallest non-abelian simple groups: A_5 of order 60 and this group of order 168.

In general, let F be a finite field. Then $\mathrm{GL}_n(F)$ is not generally simple. Obvious normal subgroups are $\mathrm{SL}_n(F)$ and $Z = \{\lambda I : \lambda \in F^\times\}$. The group $\mathrm{SL}_n(F)$ is not generally simple either, but

$$\mathrm{PSL}_n(F) = \frac{\mathrm{SL}_n(F)}{Z \cap \mathrm{SL}_n(F)},$$

the *projective special linear group*, is simple (for $n > 1$, not $\mathrm{PSL}_2(2)$, $\mathrm{PSL}_2(3)$).

Examples of Sylow subgroups

Sylow subgroups of S_n. First question: what power of p divides $n! = |S_n|$?

Write $\left\lfloor \frac{a}{b} \right\rfloor$ for the greatest integer $\leq \frac{a}{b}$. There are $\left\lfloor \frac{n}{p} \right\rfloor$ numbers in $\{1, \dots, n\}$ divisible by p, $\left\lfloor \frac{n}{p^2} \right\rfloor$ numbers divisible by p^2, and so on. So the p-power dividing $n!$ is p^a, where

$$a = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots$$

Take the p-ary expansion of n:

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_k p^k,$$

where $0 \leq a_i \leq p-1$. Then

$$\left\lfloor \frac{n}{p^i} \right\rfloor = a_i + a_{i+1} + \cdots + a_k p^{k-i},$$

so we can write:

$$\begin{aligned} a &= (a_1 + a_2 + \cdots + a_k p^{k-1}) + (a_2 + a_3 + \cdots + a_k p^{k-2}) + \cdots + (a_k) \\ &= a_1 + a_2(p+1) + a_3(p^2+p+1) + \cdots + a_k(p^{k-1} + \cdots + p + 1) \\ &= \tfrac{n-s}{p-1} \qquad \text{where } s = a_0 + a_1 + \cdots + a_k \end{aligned}$$

Divide our set of size n into non-intersecting subsets $N_1, \dots, N_s$ so that there are a_i subsets of size p^i for all i. Then S_n has a subgroup

$$H = \prod_{i=1}^{s} \mathrm{Sym}(N_i).$$

Example (3-subgroups of S_{16}). We have $16 = 3^2 + 2 \cdot 3 + 1$. So take $N = \{1, \dots, 9\}$, $N_2 = \{10, 11, 12\}$, $N_3 = \{13, 14, 15\}$, $N_4 = \{16\}$. Our subgroup consists of elements fixing the sets $\{1, \dots, 9\}$, $\{10, 11, 12\}$, $\{13, 14, 15\}$.

The power of p dividing $(p^i)!$ is $1 + p + \cdots + p^{i-1}$. So the power of p dividing $|H|$ is $\sum_i a_i(1+p+\cdots+p^{i-1})$, which is the same as the power of p dividing $n!$. So H contains a Sylow p-subgroup of S_n. A Sylow p-subgroup of H has the form $\prod_i P_i$, where $P_i \in \mathrm{Syl}_p(\mathrm{Sym}(N_i))$.

So if we can handle the case $n = p^i$ then we get a general solution.

Example $(n = p^2)$**.** We are looking for a subgroup of order p^{p+1}. It is easy to find one of order p^p, since the p-cycles $c_1 = (1 \dots p)$, $c_2 = (p+1 \dots 2p)$, …, $c_p = (p^2 - p + 1 \dots p^2)$ all commute, and generate a subgroup A isomorphic to $(C_p)^p$. Let g be the permutation

$$(1 \ p+1 \ 2p+1 \ \dots \ p^2-p+1)(2 \ p+2 \ 2p+2 \ \dots \ p^2-p+2)\dots(p \ 2p \ 3p \ \dots \ p^2)$$

(a product of p cycles of length p). Then it is easy to check that ${}^g c_i = c_{i+1}$ (and ${}^g c_p = c_1$). Hence $g \in N_{S_n}(A)$, so $\langle g \rangle A$ is a subgroup of order p^{p+1}, isomorphic to $A \rtimes_\varphi C_p$, where $C_p = \langle g \rangle$ and $\varphi_g(c_i) = c_{i+1}$.

This is an example of a *wreath product*. Suppose we have a group H and a permutation group $K \leq S_m$. Then the *wreath product* $H \wr K$ is defined as $H^m \rtimes_\varphi K$, where the action of K is given by

$$\varphi_k(h_1, \dots, h_m) = (h_{k^{-1}(1)}, h_{k^{-1}(2)}, \dots, h_{k^{-1}(m)})$$

This is a group of order $|H|^m|K|$. If $H \leq S_l$ then $H \wr K$ acts naturally on ml points.

Suppose H is a Sylow p-subgroup of S_{p^i}. Then $H \wr C_p$ acts on p^{i+1} points, and size

$$|H|^p p = (p^{1+p+\cdots+p^{i-1}})^p p = p^{1+p+\cdots+p^i}.$$

So $H \wr C_p$ is a Sylow p-subgroup of $S_{p^{i+1}}$.

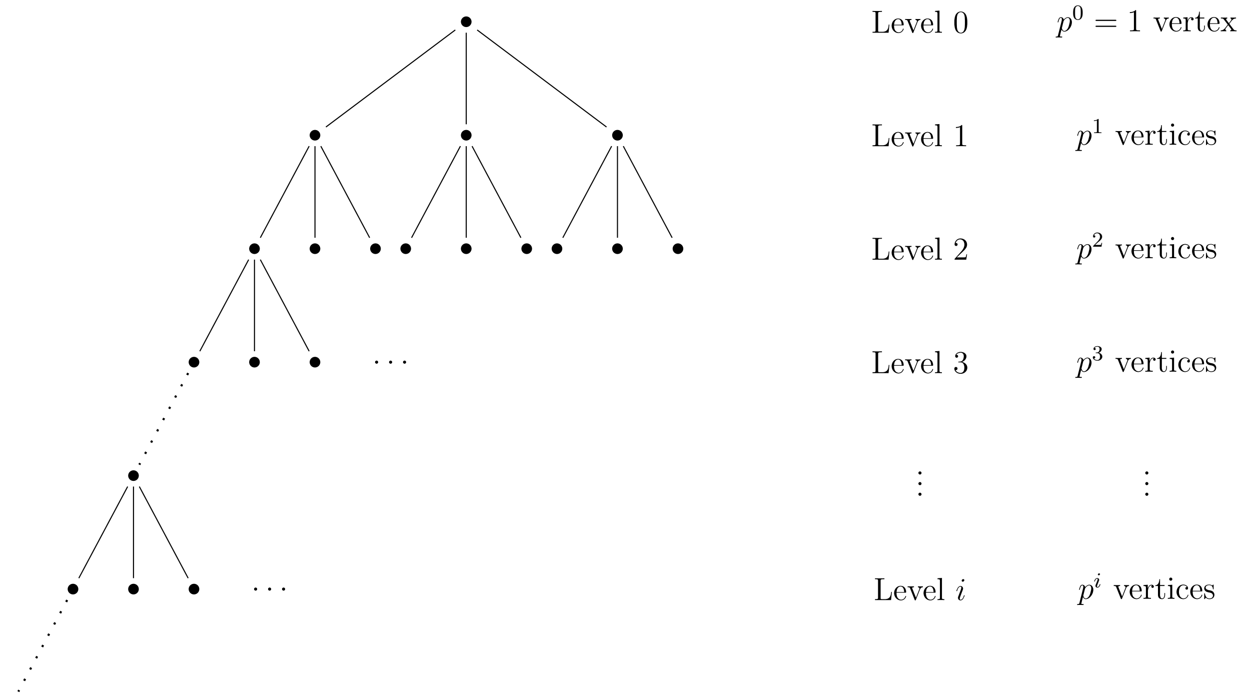
So the Sylow p-subgroup of S_{p^i} is the iterated wreath product (with i iterations)

$$((\dots((C_p \wr C_p) \wr C_p) \wr C_p \dots) \wr C_p),$$

sometimes written C_p^{2i}.

Visualizing the iterated wreath product.

We visualize the product by drawing an infinite tree. (In this picture, $p = 3$.)



At each vertex, there is a p-cycles that permutes the p branches below it. Let G be the group generated by all the cycles at the vertices.

For any i, G acts on the level i vertices, a set of size p^i.

For each i, we can define

$$L_i = \langle \text{level } j \text{ cycles for } j < i \rangle \cong (\text{Sylow } p\text{-subgroup of } S_{p^i}),$$

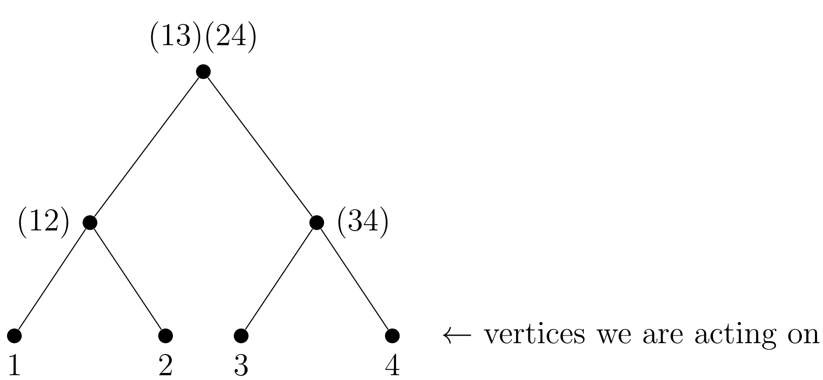
$$M_i = \langle \text{level } j \text{ cycles for } j \geq i \rangle = (\text{kernel of the action of } G \text{ on level } i \text{ vertices}) \trianglelefteq G.$$

It is not hard to see that if $g \in L_i$, $g \neq e$, then g moves a level i vertex. So $L_i \cap M_i = \{e\}$. Therefore:

$$G = M_i \rtimes L_i \cong G^{p^i} \rtimes L_i = G \wr L_i.$$

Keeping track of the levels, we obtain $L_j \cong L_{j-i} \wr L_i$ for any $i \leq j$.

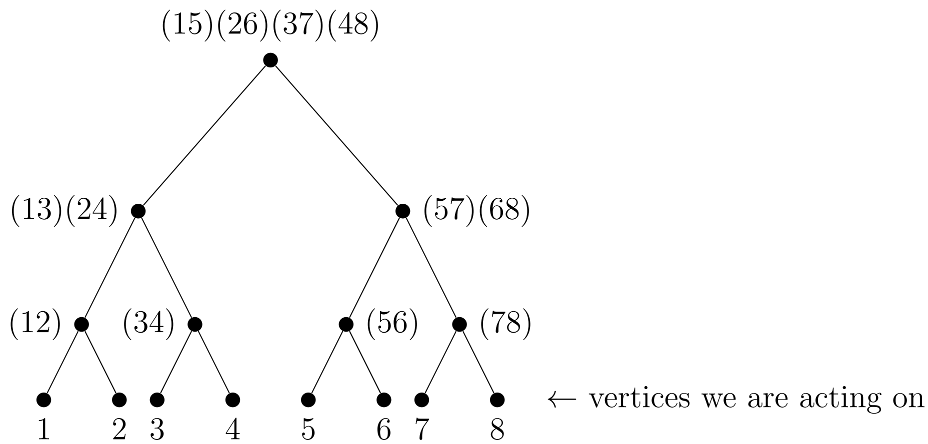
Example ($p = 2$)**.** To find the Sylow 2-subgroup of S_4, we draw 2 levels of the tree.



So a Sylow 2-subgroup of S_4 is generated by the 3 elements labelling the vertices. In fact, we only need the generators on a linear branch of the tree, so a Sylow 2-subgroup of S_4 is

$$P = \langle (12), (13)(24) \rangle.$$

To find the Sylow 2-subgroup of S_8, we draw 3 levels of the tree.



So a Sylow 2-subgroup of S_8 is generated by the 7 elements labelling the vertices. In fact, we only need the generators on a linear branch of the tree, so

$$P = \langle (12), (13)(24), (15)(26)(37)(48) \rangle.$$

From this construction, it is clear what generator we have to add to get the Sylow 2-subgroup of the next S_{2^i}.

Note that we get the other Sylow 2-subgroups by changing the numbering of the vertices at the bottom.

Sylow subgroups of $\mathrm{GL}_n(p)$. Note that:

$$|\mathrm{GL}_n(p)| = (p^n - 1)(p^n - p) \dots (p^n - p^{n-1}) = p^{\binom{n}{2}}(p^n - 1)(p^{n-1} - 1) \dots (p - 1)$$

Sylow p-subgroups are easy

$$P = \left\{ \begin{pmatrix} 1 & & & \bigstar \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} : \bigstar \text{ denotes anything} \right\}$$

Easy to check this is a subgroup. Note that $|P| = p^a$, where a is the number of $\bigstar$ entries. Clearly, $a = \binom{n}{2}$, so $P \in \mathrm{Syl}_p(\mathrm{GL}_n(p))$.

We say that an element of P has level $i + 1$ if there are i consecutive 0-diagonals above the main diagonal. So, for example,

$$\begin{pmatrix} 1 & 0 & 1 & 2 \\ & 1 & 0 & 1 \\ & & 1 & 0 \\ 0 & & & 1 \end{pmatrix}$$

has level 2. The leading diagonal is the first non-zero diagonal above all the consecutive 0-diagonals.

Facts.

(1) *Suppose* $\operatorname{level}(A) < \operatorname{level}(B)$. *Then the leading diagonal of* AB *is the same as that of* A.
(2) *The leading diagonal of* A^{-1} *is* (-1) *times the leading diagonal of* A.
(3) *If* $A, B \neq I$, *then the level of* $[A, B]$ *is strictly larger than the level of* A *and* B. *In fact, the LCS of* P *is*
$$P = L_1 > L_2 > \cdots > L_n = \{I\}$$
where L_i *is the subgroup of elements of* P *with level* $\geq i$. *It is easy to see that* $\frac{L_i}{L_{i+1}} \cong (C_p)^{n-i}$.
(4) *Note that* P *fixes each space* $\operatorname{Span}\{e_1, \ldots, e_i\}$ *for* $i = 1, \ldots, n$. *So* P *stabilizes a* flag, *a series of subspaces*
$$\{0\} < V_1 < V_2 < \cdots < V_n \qquad \textit{with } \dim V_i = i.$$
Choosing a different flag will give us a different Sylow p*-subgroup.*

Now, suppose r is a prime distinct from p. Then the Sylow r-subgroups of $\mathrm{GL}_n(p)$ behave very like the Sylow subgroups of symmetric groups. Suppose r divides
$$|\mathrm{GL}_n(p)| = p^{\binom{n}{2}}(p^n - 1)(p^{n-1} - 1)\ldots(p-1),$$
so r divides $p^a - 1$ for some $a \leq n$. Let a be the order of p in the multiplicative group modulo r. Then r divides $p^j - 1$ if and only if a divides j.

So the r-part of $|\mathrm{GL}_n(q)|$ divides
$$(p^a - 1)(p^{2a} - 1)\ldots(p^{ba} - 1)$$
where $b = \left\lfloor \frac{n}{a} \right\rfloor$. Suppose that r^c is the highest power of dividing $p^a - 1$. When is $p^{ja} - 1$ divisible by a higher power of r? We note that
$$\frac{p^{ja} - 1}{p^a - 1} = p^{(j-1)a} + p^{(j-2)a} + \cdots + p^a + 1 \equiv \underbrace{1 + 1 + \cdots + 1}_{j \text{ times}} = j \bmod r.$$

Fact. *The power of* r *dividing* $p^{ja} - 1$ *is* $c + i$, *where* i *is the power of* p *dividing* j. *So the total power of* r *dividing* $|\mathrm{GL}_n(q)|$ *is* $bc + [b!]_r$, *where* $[b!]_r$ *is the power of* r *dividing* $b!$.

Consider a vector space V of dimension n over $\mathbb{Z}_p$.

Fact. *The group* $\mathrm{GL}_a(p)$ *has an element* X *of order* r^c.

The set of block-diagonal matrices
$$\left\{ \begin{pmatrix} X^{s_1} & & & & 0 \\ & X^{s_2} & & & \\ & & \ddots & & \\ & & & X^{s_b} & \\ 0 & & & & I_{n-ab} \end{pmatrix} : s_1, \ldots, s_b \in \{0, \ldots, r-1\} \right\}$$

gives a subgroup $(C_{r^c})^b$ inside $\mathrm{GL}_n(q)$. If $b > r$, there is a linear transformation which permutes the a-subspaces. This gives a wreath product.

Proposition. *A Sylow r-subgroup of $\mathrm{GL}_n(p)$ has the form*

$$(\dots(((C_{r^c} \wr C_r) \wr C_r) \wr C_r \dots) \wr C_r$$

where the number of iterations of $\wr C_r$ is the highest j such that $r^j < \frac{n}{a}$. (Or, in other words, $C_{r^c} \wr P$ where P is a Sylow r-subgroup of S_b.)

For the mastery question, there will be a more detailed handout on the website about wreath products.

Appendix A. The alternating group A_n is simple for $n \geq 5$

In this appendix, we prove that A_n is simple for $n \geq 5$. While the proof was not included in the course, it was given as exercises in the assessed Homework 2 (base case: A_5 is simple) and Homework 4 (proof by induction on $n \geq 5$). The author's solutions to these exercises are included for completeness. The reader can also refer to the official solutions, which are posted on the course website.

Theorem 1. *The alternating group A_n is simple for $n \geq 5$.*

Before we can prove the theorem, we will prove a few lemmas. First, we note that any normal subgroup is a union of conjugacy classes.

Lemma 2. *A normal subgroup of a group is a union of conjugacy classes.*

Proof. Suppose $H \leq G$ is not a union of conjugacy classes. Then there is a conjugacy class, say ${}^G x$, such that $({}^G x) \cap H \neq \emptyset$ and $({}^G x) \cap (G \setminus H) \neq \emptyset$. So suppose without loss of generality (because we can change x to a conjugate of x) that $x \in H$, but $gxg^{-1} \notin H$, i.e. $x \notin g^{-1}Hg$. Then $H \neq g^{-1}Hg$, so H is not normal. $\square$

Therefore, to study conjugacy classes of A_n. We first note that the conjugacy classes of S_n are determined by the so-called cycle-structure.

Definition. We say that two elements of S_n have the same *cycle structure* if, when we write them in disjoint cycle notation, they have the same number of cycles of each length.

Lemma 3. *Two elements of S_n are conjugate if and only if they have the same cycle structure.*

Proof. We will first show that conjugation preserves cycle structure. Indeed, for any product of disjoint cycles $c_1 c_2 \dots c_m$ and any $\sigma \in S_n$ we have

$$\sigma c_1 c_2 \dots c_m \sigma^{-1} = (\sigma c_1 \sigma^{-1})(\sigma c_2 \sigma^{-1}) \dots (\sigma c_m \sigma^{-1})$$

Then we claim that $\sigma c_i \sigma^{-1}$ are disjoint cycles of the same length as c_i. Indeed, if $\sigma c_i \sigma^{-1}$ and $\sigma c_j \sigma^{-1}$ for $i \neq j$ are not disjoint, then they are both non-identity on some k, but this means c_i and c_j are both non-identity on $\sigma^{-1}k$, contradicting the fact that c_i and c_j are disjoint. Moroever, if c is a cycle of length l, then, without loss of generality, it acts by rotations on

$\{1,\dots,l\}$, and hence $\sigma c\sigma^{-1}$ acts by rotations on $\sigma\{1,\dots,l\}$, so $\sigma c\sigma^{-1}$ is a cycle of length l. Thus $c_1\dots c_m$ and $\sigma c_1\dots c_m\sigma^{-1}$ have the same cycle structure.

Now, we will show that if two elements of S_n have the same cycle structure, then they are conjugate. It is enough to show that a product of disjoint cycles $c=c_1\dots c_m$ with c_i cycle of length of l_i is conjugate to the product of disjoint cycles

$$d=\big(1\dots l_1\big)\big((l_1+1)\dots(l_1+l_2)\big)\dots\big((l_1+\dots+l_{m-1}+1)\dots(l_1+\dots+l_m)\big).$$

Suppose $c_i=(a_{i1}\dots a_{il_i})$ for $1\le i\le m$. Then define $\sigma\in S_n$ by

$$\sigma(a_{ij})=\sum_{k=1}^{i-1}l_k+j$$

for $1\le j\le l_i$, $1\le i\le m$, and σ is the identity everywhere else. We then have that

$$\sigma^{-1}d\sigma=c,$$

because for any $1\le j\le l_i$, $1\le i\le m$, we have that

$$\sigma^{-1}d\sigma(a_{ij})=\sigma^{-1}d\left(\sum_{k=1}^{i-1}l_k+j\right)$$

and we note that

$$d\left(\sum_{k=1}^{i-1}l_k+j\right)=\begin{cases}\sum_{k=1}^{i-1}l_k+j+1 & \text{if } 1\le j<l_i\\ \sum_{k=1}^{i-1}l_k+1 & \text{if } j=l_i\end{cases}$$

so

$$\sigma^{-1}d\left(\sum_{k=1}^{i-1}l_k+j\right)=\begin{cases}a_{i(j+1)} & \text{if } 1\le j<l_i\\ a_{i1} & \text{if } j=l_i\end{cases}$$

which shows that indeed $\sigma^{-1}d\sigma(a_{ij})=c(a_{ij})$. □

The conjugacy classes of A_n can be characterized using the conjugacy classes of S_n.

Lemma 4. *Let $g\in A_n$. If g commutes with an odd permutation, then ${}^{A_n}g={}^{S_n}g$; otherwise, $|{}^{A_n}g|=\frac{1}{2}|{}^{S_n}g|$.*

Proof. Fix $g\in A_n$. Note that for any odd permutation $h\in S_n$, we have that $h\notin A_n$, so we can write S_n as the disjoint union:

$$S_n=A_n\cup A_nh,$$

because the index of A_n in S_n is 2. Therefore:

$${}^{S_n}g=\{xgx^{-1}:x\in A_n\}\cup\{(xh)g(xh)^{-1}:x\in A_n\}={}^{A_n}g\cup{}^{A_n}(hgh^{-1}).$$

Now, if g commutes with an odd permutation h, then

$${}^{A_n}(hgh^{-1})=\{xhgh^{-1}x^{-1}:x\in A_n\}=\{xgx^{-1}:x\in A_n\}={}^{A_n}g,$$

so ${}^{S_n}g={}^{A_n}g$. Moreover, if ${}^{A_n}g\cap{}^{A_n}(hgh^{-1})\ne\emptyset$, then for some x,x' we have that

$$xgx^{-1}=x'hgh^{-1}(x')^{-1},$$

so $h^{-1}(x')^{-1}xg = gh^{-1}(x')^{-1}x$ and hence g commutes with an odd permutation $h^{-1}(x')^{-1}x$. Therefore, if g does not commute with any odd permutation, then for an odd permutation h

$$^{A_n}g \cap {}^{A_n}(hgh^{-1}) = \emptyset,$$

and hence

$$^{S_n}g = {}^{A_n}g \cup {}^{A_n}(hgh^{-1})$$

is a disjoint union. Note that $|^{A_n}(hgh^{-1})| = |^{A_n}g|$, because orbits under a transitive action have the same size. Therefore

$$|^{S_n}g| = 2|^{A_n}g|,$$

as requested. □

We can now prove that A_5 is simple.

Lemma 5. *The group A_5 is simple.*

Proof. First, we find the conjugacy classes in A_5. By Lemma 3, the conjugacy classes in S_5 are given by cycle shapes. In A_5, the possible cycle shapes are 1, 3, $(2,2)$, 5, and Lemma 3 shows that whether or not a conjugacy class splits (into two equal pieces) is given by whether or not its representative commutes with an odd permutation. We use it to prove it that the following table describes the conjugacy classes:

representative element	number of elements in the class
e	1
(123)	20
(12)(34)	15
(12345)	12
(13452)	12

- The conjugacy class of (123) in S_5 has 20 elements. Since (123) commutes with the odd permutation (45), the conjugacy class of (123) does not split in A_5—it also has 20 elements, and contains all the 3-cycles.
- The conjugacy class of (12)(34) in S_5 has 15 elements. Since (12)(34) commutes with the odd permutation (12), the conjugacy class of (12)(34) does not split in A_5—it also has 15 elements, and contains all the $(2,2)$-cycles.
- The conjugacy class of (12345) in S_5 has 24 elements. One can easily check that (12345) does not commute with any odd permutation. Therefore, the conjugacy class splits into two equal conjugacy classes in A_5, one represented by (12345), and the other by $(12)(12345)(12) = (13452)$.

To show that A_5 has no non-trivial, proper, normal subgroups, we just show that if a normal subgroup N contains one of the non-trivial conjugacy classes, then it contains all the other conjugacy classes, i.e. $N = A_5$. Trivially, $\{e\} \subseteq N$, so we only consider the non-trivial conjugacy classes. We denote by $[g]$ the conjugacy class of g.

(1) Suppose $[(123)] \subseteq N$, so all the 3-cycles are in N. Then:
  (a) $(12)(34) = (123)(234) \in N$, so N contains $[(12)(34)]$, i.e. all the $(2,2)$-cycles,
  (b) $(12345) = (145)(123) \in N$, so N contains $[(12345)]$,
  (c) $(13452) = (152)(134) \in N$, so N contains $[(13452)]$.

Thus $N = A_5$.

(2) Suppose $[(12)(34)] \subseteq N$, so all the $(2,2)$-cycles are in N. Then
$$(123) = (13)(12) = ((13)(45))((12)(45)) \in N,$$
so N contains $[(123)]$, and hence $N = A_5$ by (1).

(3) Suppose $[(12345)] \subseteq N$. Then $(124)(12345)(142) = (15243) \in [(12345)] \subseteq N$, and hence
$$(12345)(15243) = (253) \in N.$$
But this means that $[(253)] = [(123)] \subseteq N$, so $N = A_5$ by (1).

In all cases, $N = A_5$, and hence A_5 is simple. □

We note that A_n is generated by 3-cycles.

Lemma 6. *The group A_n is generated by 3-cycles.*

Proof. Let $\sigma \in A_n$ be any element. We will express σ as a product of 3-cycles. Since $\sigma \in S_n$, it can be expressed as a product of transpositions, and since $\operatorname{sgn}(\sigma) = 1$, the number of these transpositions has to be even. Explicitly, we can write
$$\sigma = (s_1t_1)(s_2t_2)\dots(s_{2m}t_{2m}).$$
We just have to show that pairs a product of two transpositions, $(st)(uv)$, is a product of 3-cycles. Indeed:

- If $|\{s,t,u,v\}| = 2$, then $(st) = (uv)$, so $(st)(uv) = 1$.
- If $|\{s,t,u,v\}| = 3$, then we may assume that $s \neq v$ and $t = u$, in which case
$$(st)(uv) = (st)(tv) = (stv).$$
- If $|\{s,t,u,v\}| = 3$, then
$$(st)(uv) = [(st)(tu)][(tu)(uv)] = (stu)(tuv).$$

This shows that A_n is generated by 3-cycles. □

Finally, we give a bound for the number of elements in a conjugacy class of A_n.

Lemma 7. *Let $n \geq 5$ and g be a non-identity element of S_n. Then $|{}^{S_n}g| \geq \binom{n}{2}$. In particular, if $g \in A_n$, then $|{}^{A_n}g| \geq n$.*

Proof. By Lemma 3, a conjugacy class of g is determined by the cycle structure of g. Let $N(g) = |{}^{S_n}g|$ be the number of elements in the conjugacy class of g.

Suppose first that g is a product of disjoint transpositions, i.e. g has cycle structure $\underbrace{(2,\dots,2)}_{k \text{ times}}$ for $k \leq \frac{n}{2}$. The number of elements with this cycle structure is
$$N(g) = \underbrace{\binom{n}{2}}_{\text{1st trans.}} \underbrace{\binom{n-2}{2}}_{\text{2nd trans.}} \cdots \underbrace{\binom{n-2k}{2}}_{k\text{th trans.}} \quad \underbrace{\frac{1}{k!}}_{\substack{\text{possible}\\ \text{permutations}}}$$

because the transpositions are disjoint and commute with each other. We only have to show that for $n \geq 5$, $k \leq \frac{n}{2}$, we have

$$\frac{1}{k!}\binom{n-2}{2}\cdots\binom{n-2k}{2} \geq 1.$$

Note that $\binom{n-2k}{2} \geq 1$, and hence:

$$\frac{1}{k!}\binom{n-2}{2}\cdots\binom{n-2k}{2} \geq \frac{(n-2)(n-3)\ldots(n-2k+2)(n-2k+1)}{2^{k-2}k!}$$

$$= \underbrace{\frac{n-2}{k}}_{\geq 1}\underbrace{\frac{n-3}{k-1}}_{\geq 1}\cdots\underbrace{\frac{n-k}{2}}_{\geq 1}\underbrace{\frac{n-k-1}{2}}_{\geq 1}\cdots\underbrace{\frac{n-2k+2}{2}}_{\geq 1}\underbrace{(n-2k+1)}_{\geq 1} \geq 1.$$

Therefore, we have shown that $N(g) \geq \binom{n}{2}$ in this case.

Now, suppose that the longest cycle in g has length $l \geq 3$, and g has $k \leq n/l$ such cycles. Then $N(g)$ is at least the number of distinct products of k disjoint cycles of length l; in other words, the number of elements with the cycle structure $\underbrace{(l,\ldots,l)}_{k \text{ times}}$. Hence

$$N(g) \geq \underbrace{\binom{n}{l}(l-1)!}_{\text{1st cycle}}\underbrace{\binom{n-l}{l}(l-1)!}_{\text{2nd cycle}}\cdots\underbrace{\binom{n-kl}{l}(l-1)!}_{k\text{th cycle}}\quad\underbrace{\frac{1}{k!}}_{\substack{\text{possible}\\ \text{permutations}}}$$

(because any l element subset has $(l-1)!$ permutations, leaving the first element fixed, that give rise to distinct l-cycles). We note that for $m \in \{0,1,\ldots,k-2\}$, since $k \leq n/l$, we have that

$$\binom{n-ml}{l}\frac{(l-1)!}{k-m} \geq \frac{(n-ml)!}{l!(n-(m+1)l)!}\frac{(l-1)!l}{(n-ml)} = \frac{(n-ml-1)!}{(n-(m+1)l)!}.$$

Altogether, we obtain:

$$N(g) \geq \underbrace{\frac{(n-1)!}{(n-l)!}}_{\geq (n-1)(n-2)}\underbrace{\left(\prod_{m=1}^{k-2}\frac{(n-ml-1)!}{(n-(m+1)l)!}\right)}_{\geq 1}\underbrace{\binom{n-(k-1)l}{l}(l-1)!\binom{n-kl}{l}(l-1)!}_{\geq 1} \geq (n-1)(n-2).$$

Finally, we note that for $n \geq 5$

$$(n-1)(n-2) - \frac{n(n-1)}{2} = \frac{2n^2-6n+4-n^2+n}{2} = \frac{n^2-5n+4}{2} = \frac{(n-4)(n-1)}{2} \geq 0,$$

and hence $N(g) \geq \binom{n}{2}$, as requested.

Finally, if $g \in A_n$, then by Lemma 4, we have that

$$|{}^{A_n}g| \geq \frac{1}{2}|{}^{S_n}g| \geq \frac{1}{2}\binom{n}{2} = \frac{n-1}{4}n \geq n$$

for $n \geq 5$. $\square$

Proof of Theorem 1. We will show that A_n is simple for $n \geq 5$ by induction on n. The base case, $n = 5$, is Lemma 5. Suppose A_{n-1} is simple for $n > 5$. We will show that A_n is simple.

Let $\{e\} \neq N \trianglelefteq A_n$ be a non-trivial normal subgroup of A_n. We will show that $N = A_n$. Note that N is a union of conjugacy classes of A_n by Lemma 2. Since $N \neq \{e\}$, N contains $\{e\}$ and a non-trivial conjugacy class, so by Lemma 7, we obtain $|N| \geq n + 1$. Fix any element $i \in \{1, \ldots, n\}$ and consider $H_i = \mathrm{Stab}_{A_n}\{i\}$. Note that $H_i \cong A_{n-1}$, because H_i contains all the 3-cycles that fix i, and we can apply Lemma 6. Since $|H_i| = (n-1)!$, there are n cosets of H_i in A_n. Hence one of these cosets has to contain at least 2 elements of N (since $|N| \geq n + 1$), so say $n_1, n_2 \in \sigma H_i$, $n_1 \neq n_2$, and write $n_1 = \sigma\tau_1$, $n_2 = \sigma\tau_2$ for $\tau_1, \tau_2 \in H_i$. Then $n = n_1^{-1} n_2 \neq e$ and

$$n = n_1^{-1} n_2 = \tau_1^{-1}\sigma^{-1}\sigma\tau_2 = \tau_1^{-1}\tau_2 \in H_i.$$

We have hence shown that $|N \cap H_i| \geq 2$, and so $N \cap H_i$ is a non-trivial normal subgroup of H_i. But by the inductive hypothesis $H_i \cong A_{n-1}$ is simple, which shows that $N \cap H_i = H_i$. Therefore, N contains the stabilizers of all the points $i \in \{1, \ldots, n\}$, and so N contains all the 3-cycles. Then, by Lemma 6, $N = A_n$. Therefore, A_n is simple. $\square$