Theorem Reference

Chen Xu

November 10, 2020

Theorem 1. A monomorphism is injective, and an epimorphism is surjective.

Definition 2. $\hookrightarrow$ is injective, $\twoheadrightarrow$ is surjective.

Definition 3. An exact sequence $G_i \xrightarrow{f_i} G_{i+1} \xrightarrow{f_{i+1}}$ is exact iff $(\forall i) \ker f_{i+1} = \operatorname{Im} f_i$

Definition 4. A short exact sequence is one of the form

$$1 \to G_1 \xrightarrow{f} G_2 \xrightarrow{g} G_3 \to 1$$

In other words:
1. f is a monomorphism
2. g is an epimorphism
3. $\operatorname{Im} f = \ker g$

Sometimes referenced as "s.e.s."

Theorem 5. If $H \triangleleft G$, then

$$1 \to H \to G \to G/H \to 1$$

is a short exact sequence, and all short exact sequences of this form for some H, G.

Theorem 6. First Isomorphism Theorem. If $f : G \to G'$ is a group homomorphism, then there exists a canonical isomorphism from $G/\ker f \xrightarrow{\sim} G'$.

Theorem 7. Second Isomorphism Theorem. If $H \triangleleft G$, and $K < G$, then:
1. $H \cap K \triangleleft K$
2. $H \cap K < H$
3. $H \triangleleft HK$
4. $K/(H \cap K) \cong HK/H$

Theorem 8. Third Isomorphism Theorem. For any $H \triangleleft G$
1. There exists a one to one mapping between subgroups of G containing H and subgroups of G/H. This mapping preserves normality and subgroup relations.
2. $K \triangleleft H \triangleleft G \implies (H/K) \triangleleft (G/K)$
3. $(G/K) \Big/ (H/K) \cong G/H$

Definition 9. An action is *faithful* if $G \to S(X)$ is a monomorphism.

Definition 10. An action is *transitive* if $\forall x, y \in X$, $\exists g \in G y = gx$.

Definition 11. The orbit of x is

$$Gx = \{gx | g \in G\}$$

Definition 12. The *isotropy group* or *stabilizer* of $x \in X$ is

$$\operatorname{Stab}_G(x) = \{g \in G | gx = x\}$$

Definition 13. The *centralizer* of $x \in X$, is simply the isotropy group under the conjugation action

$$G_x = C_G(x) = \{g | gxg^{-1} = x\}$$

Theorem 14. $G/G_x \xrightarrow{\sim} Gx$, where it's only a bijection of sets and not a homomorphism

Theorem 15. The *orbit stabilizer* theorem says that for any group action

$$[G : \mathrm{Stab}_G(g)] = \mathrm{Orb}_G(g)$$

Theorem 16. Class formula

$$|X| = \sum_{x_i \text{ orbit representatives}} [G : G_{x_i}]$$

Definition 17. The *normalizer* of H is the stabilizer of the conjugation action on subgroups of G

$$N_G(H) = \{g \in G | gHg^{-1} = H\}$$

Definition 18. The *center* of G is

$$Z(G) = \{h \in G | \forall g \in G, gh = hg\}$$

or the intersection of all centers of G ($\bigcap_{g \in G} C_G(g)$)

Definition 19. A p-group is a group of size p^k for some k.

Definition 20. A Sylow subgroup $H < G$ is a subgroup where $|H| = p^n$ and $(|H|, |G/H|) = 1$. In other words, it is the p-subgroup of "max order".

Theorem 21. *Sylow I*: If $p \mid |G|$, then there exists a p-Sylow subgroup in G.

Theorem 22. *Sylow II*

1. Any 2 Sylow subgroups are conjugate
2. Any p-subgroup is also a subgroup of a p-Sylow subgroup

Theorem 23. *Sylow III*: If N_p is the number of p-Sylow subgroups, then

1. $N_p \equiv 1 \pmod{p}$
2. $N_p \mid |G|$
3. $N_p = 1 \iff$ Sylow subgroup is normal

Definition 24. A short exact sequence (4)

$$1 \to H \to G \xrightarrow{p} K \to 1$$

is *split* if there exists some $i : K \to G$ such that $p \circ i = \mathrm{id}_K$.

Theorem 25. A short exact sequence given subgroups $H, K \triangleleft G$

$$1 \to H \to G \to K \to 1$$

splits if and only if $G \cong H \times K$.

Theorem 26. $G \equiv A \times B$ if and only if

1. $A, B \triangleleft G$
2. $A \cap B = \{e\}$
3. $AB = G$

Definition 27. For some implied group action $\varphi : H \to \mathrm{Aut}_{\mathrm{gr}}(N)$, we sometimes use the notation ${}^h n$ to denote $\varphi(h)(n)$

Definition 28. The *semidirect product* given groups N, H and a group action $\varphi : H \times N \to N$ is defined as the product except with the group operation as

$$(n_1, h_1) \cdot (n_2, h_2) = (n_1 \varphi(h_1, n_2), h_1 h_2)$$

and is denoted $N \rtimes_\varphi H$. Oftentimes the $_\varphi$ is dropped because it is implied (or is conjugation).

The inverse of (n, h) is $(\varphi(h^{-1}, n^{-1}), h^{-1})$

Theorem 29. If $G = N \rtimes_\varphi H$. Then $N \triangleleft G$, $H < G$, and $N \triangleleft H$. Additionally, conjugation by elements in H corresponds to the group action φ.

Theorem 30. If $N, H < G$. Then the following are equivalent:

1. $G \cong N \rtimes H$ considering the action when H acts on N via conjugation.
2. $N \triangleleft G$, $N \cap H = \{e\}$, $NH = G$
3. $\exists \pi : G \to H$ such that
$$H \xrightarrow{i_H} G \xrightarrow{\pi} H$$
and $\pi \circ i_H = \mathrm{id}$ and $N = \ker \pi$.
4. There exists a split short exact sequence
$$1 \to N \to G \xrightarrow{\pi} H \to 1$$

Definition 31. A *filtration* of a group G is a tower of subgroups

$$\cdots < G_2 < G_1 < G_0 = G$$

There are a couple different kinds of filtrations:

1. *Finite*: If $G_n = \{e\}$ for some n
2. *Normal*: If $G_i \triangleleft G_{i-1}$
3. *Abelian*: If normal and G_{i-1}/G_i is abelian

Definition 32. The *commutator* of two elements $x, y \in G$ is

$$[x, y] = xyx^{-1}y^{-1}$$

For two subgroups $G_1, G_2 < G$, the commutator subgroup is

$$[G_1, G_2] \coloneqq \left\langle [x, y] : x \in G_1, y \in G_2 \right\rangle$$

(Note that the set of commutators is not a subgroup in general)

Lemma 33. If $H, K \triangleleft G$, then $[H, K] \triangleleft G$.

Definition 34. $[G, G]$ is the *commutator subgroup*, or *derived subgroup* of G.

Theorem 35. $G/[G, G]$ is abelian.

Definition 36. Let $G^{(0)} = G$, and $G^{(1)} = [G, G]$. Define $G^{(i)} = [G^{(i-1)}, G^{(i-1)}]$. Then the *derived series* is the filtration

$$\cdots < G^{(2)} < G^{(1)} < G^{(0)} = G$$

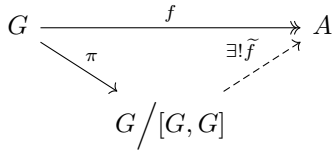
Note this is an abelian filtration, and $G^{(i)} \triangleleft G$.

Lemma 37. If G is a finite group with abelian filtration, then G has a normal filtration with cyclic quotients.

Theorem 38.

1. $G/[G, G]$ is abelian
2. $$[G, G] = \bigcap_{G/N \text{ is abelian}} N$$

3. The following universal property holds:



where A is an abelian group, and $f = \widetilde{f} \circ \pi$

Definition 39. 1. $G_{\text{ab}} = G/[G,G]$, and is known as the *abelization* of G.
2. G is *perfect* if $G = [G,G]$

Theorem 40. The following are equivalent
1. There exists a finite normal series in G with abelian quotients.
2. There exists a finite normal series in G with abelian quotients and $G_i \triangleleft G$, for each element G_i in the series.
3. The derived series terminates at e.

Definition 41. G is *solvable* if any of the equivalent conditions in theorem 40 are satisfied.

Definition 42. $H < G$ is *central* if $H < Z(G)$

Lemma 43. $H < K < G$. Then $H \triangleleft G$ and $K/H < Z(G/H)$ if and only if $[G,K] < H$.

Definition 44. A filtration $\cdots < G_i < G_{i-1} \cdots < G$ is *central* if it satisfies one of the following equivalent conditions:
1. $G_i \triangleleft G, G/G_{i+1} < Z(G_i/G_{i+1})$
2. $[G_i, G] < G_{i+1}$.

Definition 45. The *descending central series* for G is

$$\cdots < \Gamma_2 < \Gamma_1 < G$$

where $\Gamma_1 = [G,G]$ and $\Gamma_i = [\Gamma_{i-1}, G]$

Definition 46. The *ascending central series* for G is

$$e = Z_0 < Z_1 < \cdots < G$$

where $Z_1 = Z(G)$, and Z_i is the group such that $Z_i \triangleleft G$ and $Z_i/Z_{i-1} = Z(G/Z_{i-1})$.

Theorem 47. The following are equivalent
1. There exists a finite central series for G
2. The descending central series terminates at e
3. The ascending central series terminates at G

Definition 48. G is *nilpotent* if it satisfies one of the conditions of theorem 47

Theorem 49. If $G_1, \ldots, G_n$ is nilpotent, then so is $G_1 \times \cdots \times G_n$.

Theorem 50. If G is nilpotent, and $H \lneq G$, then $H \lneq N_G(H)$.

Theorem 51. If P is a Sylow subgroup of G, then $N_G(N_G(P)) = N_G(P)$

Theorem 52. If G is a p-group, then G is nilpotent.

Theorem 53. If G is nilpotent, then $G \cong P_1 \times P_2 \times \cdots \times P_n$ where $P_1, \ldots, P_n$ are the Sylow subgroups of G

Definition 54. G is *simple* if it does not have proper nontrivial normal subgroups.

Definition 55. A normal series

$$e = G_0 < G_1 < \cdots < G_n = G$$

is a *composition series* (or *Jordan Holder series*) if G_i/G_{i-1} is simple.

Remark 56. Solvable groups have composition series

Definition 57. Let $\cdots < G_i < G_{i+1} < \cdots < G$ is a normal series.

A *refinement* is then any normal series which contains $\ldots, G_i, G_{i+1}, \ldots$ in the same order, but with an additional subgroup different than all the G_i.

Definition 58. Two normal series are *equivalent* if there is a one to one correspondence between intermediate nontrivial factors such that the corresponding factors are isomorphic.

Theorem 59. *Jordan Holder*: Any two composition series are equivalent.

Lemma 60. *Zassenhaus*: If $H_1 \triangleleft H < G$ and $K_1 \triangleleft K < G$, then:

1. $H_1(H \cap K) \triangleright H_1(H \cap K_1)$
2. $K_1(H \cap K) \triangleright K_1(H_1 \cap K)$
3. $\frac{H_1(H \cap K)}{H_1(H \cap K_1)} \cong \frac{K_1(H \cap K)}{K_1(H_1 \cap K)}$

Theorem 61. *Schreier*: For a group G, any two normal series have equivalent refinements.

11/14/20
– 11/15/20

PRE-MIDTERM AND A FEW PRACTICE PROBLEMS, MATH 504, FALL 2020

General principle: the expectation is that you are well versed in everything covered in lectures and homework. In particular, know at least one solution to all and any homework problems, **including** all presentation problems. The list below is not claimed to be comprehensive but I tried to mention most of the topics we covered. If you notice an omission, let me know!

(1) Basic concepts:
   (a) Groups, subgroups, homomorphisms, cosets and double cosets, normal subgroups, factor groups
   (b) Group actions, stabilizers, centralizers, normalizers
   (c) Presentations by generators and relations
   (d) Exact sequences, split exact sequences for groups
   (e) p-groups, Sylow subgroups
   (f) Direct and semi-direct products
   (g) Center, commutator subgroup
   (h) Filtrations, derived series, central series, composition series
   (i) Solvable and nilpotent groups (several equivalent descriptions)

(2) Fundamental examples:
   (a) Symmetric groups (everything about them you learned from the worksheet),
   (b) dihedral groups (various presentations)
   (c) cyclic and abelian groups
   (d) groups of small order
   (e) matrix groups.

(3) Theorems:
   (a) Cayley
   (b) Lagrange
   (c) Three isomorphism theorems
   (d) Class formula
   (e) Jordan canonical form
   (f) Sylow theorems (two proofs for the first theorem)
   (g) Jordan-Holder theorem; Zassenhaus lemma

1. A few proctice problems

Disclaimer: this is not a comprehensive list. Once you are confident you've gone over all homework problems, you could try your hand at these, and then continue with an almost unlimited supply in Dummit and Foote.

Problem 1. Classify all finite groups of orders 1 through 10.

Problem 2. Classify all groups of order 2015.

Problem 1:

1. $\{e\}$

2, 3, 5, 7 are all prime and so cyclic $\Rightarrow \mathbb{Z}_2 \dots \mathbb{Z}_7$.

proof: consider $a \neq e$, $\langle a \rangle$ has a, e, so Lagrange $\Rightarrow |\langle a \rangle| = p$.

(9 same) 4. has order $p^2 \Rightarrow$ abelian (Homework 2 Problem 6). By Fundamental Thm Finitely Generated Abelian Groups $\Rightarrow \mathbb{Z}_4, \mathbb{Z}_2 \times \mathbb{Z}_2$

(10 same) 6. $pq = 3 \cdot 2$

normal one ✗

Homework 3 Problem 2

- $\mathbb{Z}_p \times \mathbb{Z}_q$
- $\mathbb{Z}_p \rtimes \mathbb{Z}_q$

$\varphi : K \to \mathrm{Aut}(H)$ (normal one)

$(h_1, h_2)(k_1, k_2) = (h_1(k_1 \cdot h_2), k_1 k_2)$

$h_1(k_1 \cdot h_2) \in H$

$[\varphi(k)](h_2)$

G has order pq.

n_3 = # of Sylow 3-subgroups

n_2 ——— 2-subgroups

Sylow III says $n_3 \equiv 1 \bmod 3$

$1, q$ ↳ $qx3 = 12 > 6$, impossible.

Thus, only 1 Sylow 3-subgroup $\Rightarrow$ it is normal.

$\mathbb{Z}_3$ Call this H. Call Sylow 2-subgroup $K \cong \mathbb{Z}_2$

Consider $H \rtimes K$, w/ $\varphi : K \to \mathrm{Aut}(H)$.

Trivial. φ(anything) = trivial automorphism of H.

$\mathrm{Aut}(\mathbb{Z}_3) \cong \mathbb{Z}_2$, so

$\varphi : \{0,1\} \to \mathrm{Aut}(\{0,1,2\})$

φ is homomorphism, so trivial → trivial.

Thus, 2 possibilities:

$\mathbb{Z}_6 = \mathbb{Z}_2 \times \mathbb{Z}_3$ ← $\varphi(1)$ = trivial

D_3 ← $\varphi(1)$ = flip.

0 ↦ 0
1 ↦ 2
2 ↦ 1

8. So $|G| = 8$.

If G abelian:

- $\mathbb{Z}_8$
- $\mathbb{Z}_2 \times \mathbb{Z}_4$
- $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$

If G not abelian: let $a \in G$.

- $|a| = 8 \Rightarrow$ abelian ※
- if all $|a| = 2$, then
- some $|a| = 4$. Then $\langle a \rangle = \{e, a, a^2, a^3\}$

$[G : \langle a \rangle] = 2 \Rightarrow \langle a \rangle$ normal.

Let b be any other other element.

Then $G = \{e, a, a^2, a^3, b, ab, a^2b, a^3b\}$

if $b^2 = e$, $\begin{cases} a \mapsto r \\ b \mapsto s \end{cases}$ is isomorphism from

$G \to D_4 = \{e, r, r^2, r^3, s, rs, r^2s, r^3s\}$

doing it again another way

Cases for b^2:

$b^2 \in \langle a \rangle$, b/c if not, then $b \in \langle a \rangle$.

If $b^2 = e$, then

$ba \notin \langle a \rangle$, b/c then $b = a^{\text{some power}} \in \langle a \rangle$

$ba \neq b$ b/c then $a = 1$, ※

3 last cases:

$ba = ab$. Then abelian ※

$ba = a^2b$. Then $a = b^{-1}a^2b$

$\Rightarrow a^2 = b^{-1}a^4b = e$. ※

$ba = a^3b$. $(ab)^2 = abab = aa^3bb = b^2 = 1$

7 cases

cuz a, a^3 have order 4.

$b^2 = a, a^3 \Rightarrow$ b has order 8.

last case: $b^2 = a^2$

b has order 4.

7 cases in green hold.

Last case: $ba = a^3b$

This is Q_8

$\{1, i, j, k, -1, -i, -j, -k\}$

$\langle i \rangle = \{e, i, i^2, i^3\}$ ($i^2 = -1$, $i^3 = -i$)

$\{e, i, -1, -i, j, ij\}$ ($ij = k$)

Problem 2

$2015 = 5 \cdot 13 \cdot 31$

$n_{31} \mid [G : \text{Sylow } 31\text{-subgrp}] = 65$

$n_{31} \equiv 1 \bmod 31.$

$32, 63$ do not divide $65 \Rightarrow n_{31} = 1$

$\Rightarrow H \trianglelefteq G.$ where $|H| = 31$

Similarly,

$n_{13} \mid 155$, $n_{13} \equiv 1 \bmod 13$

$\Rightarrow n_{13} = 1 \Rightarrow K \trianglelefteq G$, where $|K| = 13$

$Q = \text{Sylow } 5\text{-grp}$, $Q < G$, $KQ < G$.

and also $K \cap Q = 1$, $H \cap KQ = 1$ (H, K, Q are cyclic)

$\Downarrow$

$|KQ| = 13 \cdot 5 = 65$

$\hookrightarrow H \rtimes KQ$ w/ $\varphi : KQ \to \text{Aut}(H)$

φ trivial: $H \times KQ \cong \mathbb{Z}_{2015}$

$\uparrow \gcd(\cdots) = 1$, so cyclic blah blah

φ nontrivial: automorphisms of H take $a \mapsto a^k$.

$\text{Aut}(H) \cong \mathbb{Z}_{30}$.

$\varphi : KQ \to \mathbb{Z}_{30}$

K is cyclic, generated by k,

Q is cyclic, generated by q.

$\varphi(k) = a$

$\varphi(q) = b$

$\varphi(k^{13}) = 13a = 0$ no non-trivial solution in $\mathbb{Z}_{30}$.

$\varphi(q^5) = 5b = 0$ has soln $b = 6, 12, 18, 24$

Only possible nontrivial φ are

$\varphi(k) = 0$

$\varphi(q) = 6, 12, 18, 24.$

Consider $H \rtimes KQ$ w/ $[\varphi_6(q^r)](h) = h^{6r}$ or $[\varphi_{12}(q^r)](h) = h^{12r}$

For φ_6: $(h_1, k_1 q_1^{r_1}) \cdot (h_2, k_2 q_2)$ ($q_1 = q^{r_1}$)

$= (h_1 [\varphi_6(k_1 q_1^{r_1})](h_2), k_1 q_1 k_2 q_2)$

$= (h_1 (h_2)^{6r_1}, k_1 q_1 k_2 q_2)$

φ_{12}: $= (h_1 (h_2)^{12r_1}, k_1 q_1 k_2 q_2)$

So all $\varphi_6, \varphi_{12}, \varphi_{18}, \varphi_{24}$ isomorphic. $\Rightarrow$

$H \rtimes_{\varphi_6} KQ$

$2020 =$ seems too hard.

$2021 = 43 \cdot 47.$ (a pq-group).

$\mathbb{Z}_{2021}$

$\varphi : \mathbb{Z}_{43} \to \text{Aut}(\mathbb{Z}_{47}) \cong \mathbb{Z}_{46}$

$\varphi(0) = 0$

so $\varphi(1) = a$

$\Rightarrow \varphi(43) = 43a$

$\Rightarrow 43a = 0$ where $a \in \mathbb{Z}_{46}$, but the only solution to this is $a = 0$.

so trivial homomorphism is the only one we have to build semi direct product w/, i.e. $\mathbb{Z}_{2021} \cong \mathbb{Z}_{43} \times \mathbb{Z}_{47}$ is the only group of order 2021.

Actually, the analysis on the left for 2015 is not entirely satisfactory, b/c all we know about KQ is that it is $\cong \mathbb{Z}_{13} \rtimes \mathbb{Z}_5$ for some semi-direct product. But if we consider $G \cong HK \rtimes Q$ instead, we KNOW that $HK \cong H \times K \cong \mathbb{Z}_{31} \times \mathbb{Z}_{13} \cong \mathbb{Z}_{403}$, the answer is more clear.

Consider $\varphi : Q \to \text{Aut}(HK)$

$\cong \mathbb{Z}_5 \qquad \cong \text{Aut}(\mathbb{Z}_{403}) \cong (\mathbb{Z}/403\mathbb{Z})^*$

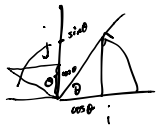
$\varphi(0) = 1$

$\varphi(1) = a$

$\varphi(5) = a^5 = 1$

i.e. we want to find roots of $a^5 - 1 \bmod 403$.

seems hard (!!)



Problem 3. Show that the dihedral group D_m of symmetries of regular m-gon is isomorphic to a subgoup of

(1) S_m,
(2) $\mathrm{GL}_2(\mathbb{C})$.

$r \mapsto (123\cdots m)$
$s \mapsto (1\,m)(2,m-1)\cdots$ until 1st ≥ 2nd.

Note: We now know at least 4 different presentations of D_m: as a semi-direct product of cyclic groups, by generators and relations, a permutation representation, and a matrix representation.

$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

$\theta = \frac{2\pi}{m}$

flip $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Problem 4. Let $B_n < \mathrm{GL}_n(\mathbb{R})$ be the subgroup of upper-triangular matrices, $T_n < B_n$ be the subgroup of diagonal matrices, and $U_n < B_n$ be the subgroup of upper-triangular matrices with 1's on the main diagonal. Assume $n \geq 2$. Show that

(a) U_n is nilpotent. What is the minimal length of its central series?
(b) B_n is solvable but not nilpotent.
(c) B_n is isomorphic to a semi-direct product of T_n and U_n.

$U_n \trianglelefteq B_n$, $U_n \cap T_n = \{I\}$, $U_n T_n = B_n$
$B_n \cong U_n \rtimes T_n$ (T_n acting by conjugation)

Note: U_n is called the unipotent subgroup, B_n - the Borel subgroup, and T_n is the torus (of $\mathrm{GL}_n(\mathbb{R})$). The statements are valid for any field of coefficients F, at least if characteristic is not 2, and so should be your proofs.

BTW: nilpotent ⇒ solvable
b/c central series is abelian,
and solvable ⇔ abelian series.

$\begin{cases} GL_n(\mathbb{R}) \\ B_n \text{ upper } \triangledown \\ T_n \text{ diagonal} \\ U_n \text{ upper } \triangledown \text{ w/ 1's on diagonal} \end{cases}$

$U_n \trianglelefteq B_n \leq GL_n(\mathbb{R})$
$T_n \leq B_n \leq GL_n(\mathbb{R})$

(a) $\mathcal{N}^k$ is set of matrices $\begin{bmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{bmatrix}$ (k super diags are 0, anything in $\mathbb{R}$) $k \in \{0,\dots,n-1\}$

$G_i = \{I+N : N \in \mathcal{N}^i\}$ is a group. — needs proof

Look @ elements of $[G_i, G]$:

$(I+N)(I+M)(I+N)^{-1}(I+M)^{-1}$ $\quad N \in \mathcal{N}^i$, $M \in \mathcal{N}^0$

$\mathcal{H}\ \mathcal{N}$

There is $N' \in \mathcal{N}^i$ s.t. $(I+N)^{-1} = (I+N')$, so — $N+N'+NN'=0$

$(I+N+M+NM)(I+N')(I+M)^{-1}$ But $NM \in \mathcal{N}^{i+1}$

$\equiv (I+N+M)(I+N')(I+M)^{-1} \pmod{\mathcal{N}^{i+1}}$

$= (I+N+M+N'+NN'+MN')(I+M)^{-1} \pmod{\mathcal{N}^{i+1}}$

$= (I+M)(I+M)^{-1} = I \pmod{\mathcal{N}^{i+1}}$

⇒ all commutators in "$I+\mathcal{N}^{i+1}$", i.e. G_{i+1}.

⇒ $[G_i, G] \leq G_{i+1}$

↳ descending central series (w/ indices going up :))

$I = G_{n-1} \leq \dots \leq G_0 = U_n$

length 2

(b) **Lemma:** if $N \trianglelefteq G$, and N and $\frac{G}{N}$ are solvable, then G is solvable.

Note that $U_n \trianglelefteq B_n$ (take homomorphism taking B_n to its diagonal, and U_n is kernel)

and $\frac{B_n}{U_n} \cong T_n$ ($\frac{G}{\ker\varphi} \cong \operatorname{Im}\varphi$)

which is abelian, and so $\frac{B_n}{U_n}$ is in particular solvable

and U_n is nilpotent (from (a)) and hence solvable.

Thus, B_n solvable.

But, not nilpotent: assume nilpotent. Then, any proper subgrp of B_n, say H, satisfies

proper $H < N_{B_n}(H)$. (Dank find pg later)

Let $H = T_n < B_n$. $N_{B_n}(T_n) = T_n$, hence contradiction.

→ proof: $N_{GL_n}(T_n)$ is the set of generalized permutation matrices.

(permutation matrix but non-zero entries don't have to be 1).

Alternate proof that B_n not nilpotent:

First consider $n=2$. Claim: $[U_2, B_2] \geq U_2$.

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}\begin{bmatrix} b_1 & b_2 \\ 0 & b_3 \end{bmatrix}\begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix}\frac{1}{b_1 b_3}\begin{bmatrix} b_3 & -b_2 \\ 0 & b_1 \end{bmatrix} =$$

$[B_n, B_n] = U_n$

Seems like a lot of work

Class Formula:

(In general for G acting on set X) $$|X| = \sum_{x_i \text{ orbit representatives}} [G : G_{x_i}]$$

$G_{x_i} = \{ g \cdot x_i : \forall g \in G \}$ (group action.)

aka $Orb_G(x_i)$

Orbit-Stabilizer Thm: $(\forall g \in G)$ $|G| = |Stab_G(g)|\,|Orb_G(g)|$

(for G acting on G by conjugation) $$|G| = Z(G) + \sum_{\text{non-trivial orbit reps } x} [G : G_x]$$

fix rep

Proof: $\phi : G \to Orb_G(r)$ s.t. $\phi(g) = g * r$ (* gp action.)

$\phi(g) = \phi(h) \Leftrightarrow g^{-1}h \in Stab_G(r)$

so $gStab_G(r) \mapsto g * r$ is bijection from $G/Stab_G(r) \to Orb_G(x)$.

Splitting

$1 \to H \to G \xrightarrow{f} K \to 1$ exact

and $K \xrightarrow{g} G$ s.t. fog is identity on K.

Homework 5 for 504, Fall 2020

Presentation problems only!

due Friday, November 13

Problem 1. Let G be a group. Prove that the following are equivalent:

(1) There exists a (finite) central series $\{e\} = G_0 < G_1 < \ldots < G_n = G$.

(2) The descending central series

$$\ldots < \Gamma_i = [\Gamma_{i-1}, G] < \Gamma_{i-1} < \ldots < \Gamma_1 < \Gamma_0 = G$$

terminates at $\Gamma_n = \{e\}$.

(3) The ascending central series $\{e\} = Z_0 < Z_1 < Z_2 \ldots$ (where $Z_i/Z_{i-1} = Z(G/Z_{i-1})$) terminates at $Z_n = G$.

Specifically:

Problem 1.1P: Prove the equivalence of (1) and (2).

Problem 1.2P: Prove the equivalence of (1) and (3)

Problem 2P. Let G be a finite group and $P < G$ be its Sylow subgroup. Show that $N_G(N_G(P)) = N_G(P)$.

Problem 1: $\{e\} = G_0 < \_\_\_\_ < G_n = G$,

$G_i = [G, G_{i+1}]$

is central series: $[G, G_{i+1}] \leq G_i$,

or rephrased $\frac{G_{i+1}}{G_i}$ is central in $\frac{G}{G_i}$.

$< Z(\quad)$

Follows from:

Claim: $[G,K] \leq H \iff [G/H, K/H] = e$

($\Rightarrow$) $gkg^{-1}k^{-1} = h \Rightarrow gHkHg^{-1}Hk^{-1}H = (gkg^{-1}k^{-1})H = H$

($\Leftarrow$) $\forall g \in G, k \in K$,

$gHkHg^{-1}Hk^{-1}H = H \Rightarrow gkg^{-1}k^{-1}H = H$

$\Rightarrow gkg^{-1}k^{-1} \in H$

(1) $\Rightarrow$ (2):

$\Gamma_n = G$

$\Gamma_{n-1} = [\Gamma_n, G]$

$\Gamma_{n-2} = [\Gamma_{n-1}, G]$

$\downarrow$

$\Gamma_0 = [\Gamma_1, G]$

Claim: $\Gamma_i = [\Gamma_{i+1}, G] \leq [G_{i+1}, G] \left(\leq G_i\right)$

Induction. Base case: $\Gamma_n = G$, $G_n = G$, $\Gamma_n \leq G_n$ ✓

Now assume $\Gamma_{i+1} \leq G_{i+1}$.

WTS: $[\Gamma_{i+1}, G] \leq [G_{i+1}, G] \Rightarrow \Gamma_i \leq G_i$

b/c any commutator is in b/c all elements of Γ_{i+1} are in G_{i+1}

$[\Gamma_{i+1}, G] = \left\langle \{[\gamma, g] : \gamma \in \Gamma_{i+1}\}_{g \in G} \right\rangle$ RED?

(2) $\Rightarrow$ (1): trivial.

---

(1) $\Rightarrow$ (3):

$\frac{Z_1}{Z_{i-1}} = Z\left(\frac{G}{Z_{i-1}}\right)$

$Z_0 = \{e\}$

$Z_1 = Z(G)$

$\vdots$

(3) $\Rightarrow$ (1)

trivial.

Claim: $Z_i \geq G_i$ (in fact normal)

Induction. Base case: $Z_0 = \{e\} \geq \{e\} = G_0$.

Assume $Z_{i-1} \geq G_{i-1}$.

Then $\frac{G_i}{G_{i-1}} \leq Z\left(\frac{G}{G_{i-1}}\right) \iff [G, G_i] \leq G_{i-1}$ but $G_{i-1} \leq Z_{i-1}$

$\iff \frac{G_i}{Z_{i-1}} \leq Z\left(\frac{G}{Z_{i-1}}\right) = \frac{Z_i}{Z_{i-1}}$.

↑ in center means abelian so in particular normal

3rd Iso says

$\frac{G}{H} \Big/ \frac{P}{H} \cong \frac{G}{P} \Rightarrow \frac{Z_i}{G_i}$ is a thing $\Rightarrow G_i \trianglelefteq Z_i$.

Bashir Abdel-Fattah
1777869

Assignment 1
Math 504

2020/11/15

Problem 1. Show that an identity element and an inverse in a group are unique.

Proof. Suppose that the elements e_1 and e_2 were both identities in a group G. Then, beginning with the expression e_1e_2, because e_1 is an identity we have $e_1e_2 = e_2$, and because e_2 is an identity we have $e_1e_2 = e_1$. Putting this together, we have

$$e_1 = e_1e_2 = e_2$$

thus proving that the identity in G is unique. Now let g be any element of G, and suppose that g had two inverse h_1 and h_2 (that is, $gh_1 = h_1g = gh_2 = h_2g = e$). Then

$$h_1 = h_1e = h_1(gh_2) = (h_1g)h_2 = eh_2 = h_2$$

thus demonstrating that the inverse of an arbitrary element $g \in G$ is also unique. ■

Problem 2. Let G be a set with two binary operations, denoted $*$ and $\circ$, and a fixed element e such that

1. e is the identity for both operations
2. $(a \circ b) * (c \circ d) = (a * c) \circ (b * d)$

Show that these two operations coincide and moreover that they are associative and commutative.

Proof. By taking $(a, b, c, d) = (g_1, e, e, g_2)$ in equation 2 above and using the fact that e is the identity with respect to both operations, we have

$$g_1 * g_2 = (g_1 \circ e) * (e \circ g_2) = (g_1 * e) \circ (e * g_2) = g_1 \circ g_2$$

for all $g_1, g_2 \in G$. Thus the operations $* : G \times G \to G$ and $\circ : G \times G \to G$ are exactly equal, and going forwards we will denote both operations with the symbol "$*$". By taking $(a, b, c, d) = (e, g_1, g_2, e)$ in equation 2, we have

$$g_1 * g_2 = (e * g_1) * (g_2 * e) = (e * g_2) * (g_1 * e) = g_2 * g_1$$

for all $g_1, g_2 \in G$, so the $*$ operation is commutative. Finally, taking $(a, b, c, d) = (g_1, g_2, e, g_3)$ in equation 2 gives us

$$(g_1 * g_2) * g_3 = (g_1 * g_2) * (e * g_3) = (g_1 * e) * (g_2 * g_3) = g_1 * (g_2 * g_3)$$

for all $g_1, g_2, g_3 \in G$, thus demonstrating that the $*$ operation is also associative. ■

Problem 3. *Cyclic groups.*

(1P). Let $a \in G$ be an element of order n. Show that if $a^m = e$ then $n \mid m$.

Proof. Suppose that $a^m = e$ and $|a| = n$. From the division algorithm in $\mathbb{Z}$, there are unique integers q and $0 \leq r < n$ such that

$$m = qn + r$$

Then we have that

$$e = a^m = a^{qn+r} = (a^n)^q a^r = e^q a^r = a^r$$

Were r nonzero, then the fact that $0 < r < n$ and $a^r = e$ would contradict the fact that $|a| = n$ is the smallest positive integer such that $a^n = e$, so we must have $r = 0$. In particular, $m = qn$ for some integer q, so $n \mid m$. ■

(2P). Prove that a subgroup of a cyclic group is cyclic.

Proof. Letting $G = \langle a \rangle$ be any cyclic group, then there exists a surjective group homomorphism $f : \mathbb{Z} \to G$ given by $f(n) = a^n$. That f is surjective is clear from the explicit characterization of G

$$G = \langle a \rangle = \{a^n : n \in \mathbb{Z}\} = f(\mathbb{Z})$$

and that f is a homomorphism follows from the equation

$$f(n)f(m) = a^n a^m = a^{n+m} = f(n+m)$$

for any $n, m \in \mathbb{Z}$. Then if H is any subgroup of G, then $f^{-1}(H)$ is a subgroup of $\mathbb{Z}$ (see lemma below). Since every subgroup of $\mathbb{Z}$ has the form $m\mathbb{Z}$ for some $m \in \mathbb{Z}$ (see the proof of question 4P), then because f is surjective we have

$$H = f(f^{-1}(H)) = f(m\mathbb{Z}) = \{a^k : k \in m\mathbb{Z}\} = \{a^{mn} : n \in \mathbb{Z}\} = \{(a^m)^n : n \in \mathbb{Z}\} = \langle a^m \rangle$$

That is, H is a cyclic group generated by the element a^m. ■

Lemma. Let $f : G_1 \to G_2$ be a group homomorphism, and let H_2 be a subgroup of G_2. Then $H_1 = f^{-1}(H_2)$ is a subgroup of G_1.

Proof. To begin, we note

$$\boxed{\begin{array}{c} H \text{ is a subgroup of } G\text{:} \\ e \in H \\ g_1, g_2 \in H \Rightarrow g_1 g_2 \in H \\ g \in H \Rightarrow g^{-1} \in H \end{array}} \iff \boxed{\begin{array}{c} H \text{ is a nonempty} \\ \text{subset of } G \text{ such that} \\ g_1, g_2 \in H \Rightarrow g_1 g_2^{-1} \in H \end{array}}$$

For the forwards implication, suppose that H is a subgroup of G. Then $e \in H$, so H is nonempty, and the fact that H is closed under inverses and multiplication means that

$$g_1, g_2 \in H \quad \Longrightarrow \quad g_1, g_2^{-1} \in H \quad \Longrightarrow \quad g_1 g_2^{-1} \in H$$

Now suppose that H is a nonempty subset of G such that $g_1, g_2 \in H \Longrightarrow g_1 g_2^{-1} \in H$. Because H is nonempty, then there exists some element $g_0 \in H$, in which case

$$g_0, g_0 \in H \quad \Longrightarrow \quad g_0 g_0^{-1} = e \in H$$

If $g \in H$, then because $e \in H$ we have

$$e, g \in H \quad \Longrightarrow \quad e g^{-1} = g^{-1} \in H$$

Finally, if $g_1, g_2 \in H$, then because $g_2 \in H \Longrightarrow g_2^{-1} \in H$ we have

$$g_1, g_2 \in H \quad \Longrightarrow \quad g_1, g_2^{-1} \in H \quad \Longrightarrow \quad g_1 (g_2^{-1})^{-1} = g_1 g_2 \in H$$

Therefore H satisfies the three properties of a subgroup, so we conclude that H is a subgroup of G.

Now, proceeding back to the main lemma, it is clear that H_1 is nonempty, because $f(e_{G_1}) = e_{G_2} \in H_2$ (because f is a homomorphism and H_2 is a subgroup), so $e_{G_1} \in f^{-1}(H_2)$. Now, suppose $g_1, g_2 \in H_1 = f^{-1}(H_2)$ (that is, suppose $f(g_1), f(g_2) \in H_2$). Then because H_2 is a subgroup, we have

$$f(g_1) f(g_2)^{-1} = f(g_1 g_2^{-1}) \in H_2$$

so $g_1 g_2^{-1} \in f^{-1}(H_2) = H_1$. Therefore H_1 is a subgroup of G_1. ■

(3P). Let $f : G \to G'$ be a group homomorphism, and assume that G is cyclic. Show that $\operatorname{im} f$ is a cyclic subgroup of G'.

Proof. Since G is cyclic, then there is some element $a \in G$ such that $G = \langle a \rangle = \{a^n : n \in \mathbb{Z}\}$, in which case

$$\operatorname{im} f = \{f(g) : g \in G\} = \{f(a^n) : n \in \mathbb{Z}\} = \{f(a)^n : n \in \mathbb{N}\} = \langle f(a) \rangle$$

That is, $\operatorname{im} f$ is the cyclic subgroup of G' generated by the element $f(a)$. ■

(4P). Prove that all subgroups of $\mathbb{Z}$ have the form $m\mathbb{Z}$ for $m \in \mathbb{Z}$.

Proof. Suppose that H is a subgroup of $\mathbb{Z}$. If H is the trivial subgroup $\{0\}$ then $H = 0\mathbb{Z}$, and otherwise H must contain at least one element $h \neq 0$. Either $h > 0$ or, if $h < 0$, then because H is itself a group under addition it must contain the additive inverse $-h > 0$. In either case, H contains at least one element greater than zero, and hence h contains a smallest element greater than zero by the well-ordering property of the natural numbers (that is, $H \cap \mathbb{N}$ is a nonempty subset of $\mathbb{N}$, so it contains a least element). Letting m be the minimal element of H that is greater than zero, we claim that $H = m\mathbb{Z}$. The reverse inclusion is evident: since $m \in H$, then H must also contain the subgroup generated by m

$$H \supset \langle m \rangle = \{nm : n \in \mathbb{N}\} = m\mathbb{Z}$$

To demonstrate the forwards inclusion, suppose for the sake of contradiction that $H \not\subset m\mathbb{Z}$: that is, that there exists some element $k \in H$ such that $k \notin m\mathbb{Z}$. Then by the division algorithm in the integers there are unique integers q and $0 \leq r < m$ such that

$$k = qm + r$$

and since $k \notin m\mathbb{Z}$ we must have that $r \neq 0$. Then since $k \in H$ and $-qm \in m\mathbb{Z} \subset H$ we conclude that H must also contain $k - qm = r$. However, the fact that $r \in H$ and $0 < r < m$ contradicts the choice of m to be the smallest positive element of H, so the assumption that there exists some element $k \in H$ such that $k \notin m\mathbb{Z}$ must have been false. Thus we have both $H \subset m\mathbb{Z}$ and $m\mathbb{Z} \subset H$, therefore $H = m\mathbb{Z}$. ■

(5P). Classify all subgroups of $\mathbb{Z}/m\mathbb{Z}$.

Proof. The subgroups of $\mathbb{Z}/m\mathbb{Z}$ are all isomorphic to $\mathbb{Z}/n\mathbb{Z}$ for some $n \mid m$. They must be isomorphic to $\mathbb{Z}/n\mathbb{Z}$ for some n because a subgroup of a cyclic group is cyclic by problem 3.2, and every cyclic group is isomorphic to either $\mathbb{Z}$ or $\mathbb{Z}/n\mathbb{Z}$ for some n (as discussed in the lecture). Since $\mathbb{Z}/m\mathbb{Z}$ is a finite group, then none of its subgroups can be isomorphic to the infinite group $\mathbb{Z}$, so it follows that the subgroup must be isomorphic to $\mathbb{Z}/n\mathbb{Z}$ for some n. This value of n must divide the value of m because the order of a subgroup must divide the order of the group (and the order of a group $\mathbb{Z}/n\mathbb{Z}$ is n).

■

(6). Show that all finitely generated non-trivial subgroups of $\mathbb{Q}$ are cyclic and isomorphic to $\mathbb{Z}$.

Proof. Let S be any finite subset of $\mathbb{Q}$, where we assume without loss of generality that $0 \notin S$ (because the group generated by $S \cup \{0\}$ is the same as the group generated by S, since S already contains 0 by virtue of being a subgroup and no smaller group contains S, so the smallest subgroup containing $S \cup \{0\}$ is $\langle S \rangle$) and that S is non-empty (because we only want to consider non-trivial subgroups of $\mathbb{Q}$, and the subgroup generated by the empty set is the smallest subgroup containing the empty set, i.e. the trivial subgroup). We can also assume without loss of generality that each element of S is positive, because for any element $s_i \in S = \{s_1, s_2, \ldots, s_k\}$ a subgroup of $\mathbb{Q}$ contains s_i if and only if it contains $-s_i$ (since if a subgroup contains an element, then it must also contain the inverse of that element), so the collection of subgroups of $\mathbb{Q}$ containing $\{s_1, \ldots, s_i, \ldots, s_k\}$ is exactly the same as the collection of subgroups containing $\{s_1, \ldots, -s_i, \ldots, s_k\}$ and hence both subsets generate the same subgroup. Therefore we can write

$$S = \{\frac{p_1}{q_1}, \frac{p_2}{q_2}, \ldots, \frac{p_k}{q_k}\}$$

where p_i and q_i are positive integers for each $1 \leq i \leq k$. Then, letting q be any integer such that q/q_i is an integer for each $1 \leq i \leq k$ (for instance $q = q_1 q_2 \ldots q_k$), the map $f : \langle S \rangle \to \mathbb{Z}$ given by $f(s) = qs$ is a group homomorphism. If we recall from the lecture that the subgroup generated by a subset T of a group is given (in additive notation) by

$$\langle T \rangle = \{n_1 t_2 + n_2 t_2 + \ldots + n_m t_m : m \in \mathbb{N},\ t_1, t_2, \ldots, t_m \in T,\ n_1, n_2, \ldots, n_m \in \mathbb{Z}\}$$

then it is clear that $\mathbb{Z}$ is indeed the codomain for f, because every element of $\langle S \rangle$ has the form

$$s = n_1(\frac{p_1}{q_1}) + n_2(\frac{p_2}{q_2}) + \ldots + n_k(\frac{p_k}{q_k}) \quad \text{for} \quad n_1, n_2, \ldots, n_k \in \mathbb{Z}$$

so it follows that

$$f(s) = qs = q\big(n_1(\frac{p_1}{q_1}) + n_2(\frac{p_2}{q_2}) + \ldots + n_k(\frac{p_k}{q_k})\big) = n_1 p_1(\frac{q}{q_1}) + n_2 p_2(\frac{q}{q_2}) + \ldots + n_k p_k(\frac{q}{q_k})$$

is an integer. Furthermore, f is a group homomorphism because for any $s, t \in \langle S \rangle$ we have

$$f(s+t) = q(s+t) = qs + qt = f(s) + f(t)$$

and f is injective because

$$qs = f(s) = f(t) = qt \quad \Longrightarrow \quad s = t$$

(since $q \neq 0$, so we can divide both sides by q). As a result, $\langle S \rangle$ is isomorphic to its image (it is already injective, and it is surjective onto $\operatorname{im} f \leq \mathbb{Z}$), and since every subgroup of $\mathbb{Z}$ has the form $m\mathbb{Z}$ for some integer m, it follows that

$$\langle S \rangle \cong \operatorname{im} f = m\mathbb{Z}$$

We must also have that $m \neq 0$, because $\langle S \rangle$ is a non-trivial subgroup (it contains S, which itself contains at least one non-zero element) and hence cannot be isomorphic to a trivial subgroup. Then it follows that $m\mathbb{Z} \cong \mathbb{Z}$ because the map $g : \mathbb{Z} \to m\mathbb{Z}$ given by

$$g(n) = mn$$

is an isomorphism: g is a homomorphism because for all $n_1, n_2 \in \mathbb{N}$ we have

$$g(n_1 + n_2) = m(n_1 + n_2) = mn_1 + mn_2 = g(n_1) + g(n_2),$$

g is injective because $m \neq 0$, so

$$mn_1 = g(n_1) = g(n_2) = mn_2 \quad \Longrightarrow \quad n_1 = n_2$$

(just by dividing by m), and g is surjective because

$$g(\mathbb{Z}) = \{g(n) : n \in \mathbb{Z}\} = \{mn : n \in \mathbb{Z}\} = m\mathbb{Z}.$$

Thus

$$\langle S \rangle \cong m\mathbb{Z} \cong \mathbb{Z}$$

■

Problem 4. *Quaternions.* Brushing up on your vector calculus might be helpful for this problem!
Let $V = \mathbb{R}^3$ and set $C = \mathbb{R} \times V$. (As a real vector space, C may be identified with $\mathbb{R}^4$.) Define a product on C by

$$(a, u)(b, v) = (ab - u \cdot v; av + bu + u \times v).$$

(C is called the quaternion algebra). Define the *norm* $N(a; u) = a^2 + |u|^2$.

1. Show that the product defined above is associative but not commutative.

Proof. In order to show that the given product is associative, we want to demonstrate the equality of the products

$$\begin{aligned}\big((a, \mathbf{u})(b, \mathbf{v})\big)(c, \mathbf{w}) &= (ab - \mathbf{u} \cdot \mathbf{v},\ a\mathbf{v} + b\mathbf{u} + \mathbf{u} \times \mathbf{v})(c, \mathbf{w}) \\ &= \big((ab - \mathbf{u} \cdot \mathbf{v})c - (a\mathbf{v} + b\mathbf{u} + \mathbf{u} \times \mathbf{v}) \cdot \mathbf{w},\ (ab - \mathbf{u} \cdot \mathbf{v})\mathbf{w} + c(a\mathbf{v} + b\mathbf{u} + \mathbf{u} \times \mathbf{v}) + (a\mathbf{v} + b\mathbf{u} + \mathbf{u} \times \mathbf{v}) \times \mathbf{w}\big)\end{aligned}$$

and

$$\begin{aligned}(a, \mathbf{u})\big((b, \mathbf{v})(c, \mathbf{w})\big) &= (a, \mathbf{u})(bc - \mathbf{v} \cdot \mathbf{w},\ b\mathbf{w} + c\mathbf{v} + \mathbf{v} \times \mathbf{w}) \\ &= \big(a(bc - \mathbf{v} \cdot \mathbf{w}) - \mathbf{u} \cdot (b\mathbf{w} + c\mathbf{v} + \mathbf{v} \times \mathbf{w}),\ a(b\mathbf{w} + c\mathbf{v} + \mathbf{v} \times \mathbf{w}) + (bc - \mathbf{v} \cdot \mathbf{w})\mathbf{u} + \mathbf{u} \times (b\mathbf{w} + c\mathbf{v} + \mathbf{v} \times \mathbf{w})\big)\end{aligned}$$

The equality of the scalar part follows because for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ we have

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \det \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

so

$$\begin{aligned}\operatorname{Re}\Big[\big((a, \mathbf{u})(b, \mathbf{v})\big)(c, \mathbf{w})\Big] &= (ab - \mathbf{u} \cdot \mathbf{v})c - (a\mathbf{v} + b\mathbf{u} + \mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} \\ &= abc - c\mathbf{u} \cdot \mathbf{v} - a\mathbf{v} \cdot \mathbf{w} - b\mathbf{u} \cdot \mathbf{w} - \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) \\ &= abc - a\mathbf{v} \cdot \mathbf{w} - b\mathbf{u} \cdot \mathbf{w} - c\mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) \\ &= a(bc - \mathbf{v} \cdot \mathbf{w}) - \mathbf{u} \cdot (b\mathbf{w} + c\mathbf{v} + \mathbf{v} \times \mathbf{w}) \\ &= \operatorname{Re}\Big[(a, \mathbf{u})\big((b, \mathbf{v})(c, \mathbf{w})\big)\Big]\end{aligned}$$

The equality of the vector part can be demonstrated using the BAC-CAB identity

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$$

for $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{R}^3$. We have

$$\begin{aligned}
\mathrm{Im}\Big[\big((a,\mathbf{u})(b,\mathbf{v})\big)(c,\mathbf{w})\Big] &= (ab-\mathbf{u}\cdot\mathbf{v})\mathbf{w} + c(a\mathbf{v}+b\mathbf{u}+\mathbf{u}\times\mathbf{v}) + (a\mathbf{v}+b\mathbf{u}+\mathbf{u}\times\mathbf{v})\times\mathbf{w}\\
&= ab\mathbf{w} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w} + ac\mathbf{v} + bc\mathbf{u} + c\mathbf{u}\times\mathbf{v} + a\mathbf{v}\times\mathbf{w} + b\mathbf{u}\times\mathbf{w} + (\mathbf{u}\times\mathbf{v})\times\mathbf{w}\\
&= ab\mathbf{w} + ac\mathbf{v} + bc\mathbf{u} + c\mathbf{u}\times\mathbf{v} + a\mathbf{v}\times\mathbf{w} + b\mathbf{u}\times\mathbf{w} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w} - \mathbf{w}\times(\mathbf{u}\times\mathbf{v})\\
&= ab\mathbf{w} + ac\mathbf{v} + bc\mathbf{u} + c\mathbf{u}\times\mathbf{v} + a\mathbf{v}\times\mathbf{w} + b\mathbf{u}\times\mathbf{w} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w} - \big(\mathbf{u}(\mathbf{w}\cdot\mathbf{v}) - \mathbf{v}(\mathbf{w}\cdot\mathbf{u})\big)\\
&= ab\mathbf{w} + ac\mathbf{v} + bc\mathbf{u} + c\mathbf{u}\times\mathbf{v} + a\mathbf{v}\times\mathbf{w} + b\mathbf{u}\times\mathbf{w} - \mathbf{w}(\mathbf{u}\cdot\mathbf{v}) - (\mathbf{v}\cdot\mathbf{w})\mathbf{u} + \mathbf{v}(\mathbf{u}\cdot\mathbf{w})\\
&= ab\mathbf{w} + ac\mathbf{v} + a\mathbf{v}\times\mathbf{w} + bc\mathbf{u} - (\mathbf{v}\cdot\mathbf{w})\mathbf{u} + b\mathbf{u}\times\mathbf{w} + c\mathbf{u}\times\mathbf{v} + \mathbf{v}(\mathbf{u}\cdot\mathbf{w}) - \mathbf{w}(\mathbf{u}\cdot\mathbf{v})\\
&= ab\mathbf{w} + ac\mathbf{v} + a\mathbf{v}\times\mathbf{w} + bc\mathbf{u} - (\mathbf{v}\cdot\mathbf{w})\mathbf{u} + b\mathbf{u}\times\mathbf{w} + c\mathbf{u}\times\mathbf{v} + \mathbf{u}\times(\mathbf{v}\times\mathbf{w})\\
&= a(b\mathbf{w}+c\mathbf{v}+\mathbf{v}\times\mathbf{w}) + (bc-\mathbf{v}\cdot\mathbf{w})\mathbf{u} + \mathbf{u}\times(b\mathbf{w}+c\mathbf{v}+\mathbf{v}\times\mathbf{w})\\
&= \mathrm{Im}\Big[(a,\mathbf{u})\big((b,\mathbf{v})(c,\mathbf{w})\big)\Big]
\end{aligned}$$

Thus

$$\big((a,\mathbf{u})(b,\mathbf{v})\big)(c,\mathbf{w}) = (a,\mathbf{u})\big((b,\mathbf{v})(c,\mathbf{w})\big)$$

for any elements $(a,\mathbf{u}),(b,\mathbf{v}),(c,\mathbf{w})\in C$, so the given multiplication operation is associative. To show that it isn't commutative, consider the elements

$$i = \big(0,(1,0,0)\big)$$

$$j = \big(0,(0,1,0)\big)$$

in C. Then

$$ij = \big(0,(1,0,0)\big)\big(0,(0,1,0)\big) = \big(0{\cdot}0-(1,0,0){\cdot}(0,1,0);\ 0{\cdot}(0,1,0)+0{\cdot}(1,0,0)+(1,0,0)\times(0,1,0)\big) = \big(0,(0,0,1)\big)$$

$$ji = \big(0,(0,1,0)\big)\big(0,(1,0,0)\big) = \big(0{\cdot}0-(0,1,0){\cdot}(1,0,0);\ 0{\cdot}(1,0,0)+0{\cdot}(0,1,0)+(0,1,0)\times(1,0,0)\big) = \big(0,(0,0,-1)\big)$$

That is, $ij\neq ji$, so we have found a particular example of non-commuting elements of C. ■

2. Show that $N(\alpha\beta) = N(\alpha)N(\beta)$ for $\alpha,\beta\in C$.

Proof. Let $\alpha=(a,\mathbf{u})$ and $\beta=(b,\mathbf{v})$. Then

$$\begin{aligned}
N(\alpha\beta) = N\big((a,\mathbf{u})(b,\mathbf{v})\big) &= N\big((ab-\mathbf{u}\cdot\mathbf{v},\ a\mathbf{v}+b\mathbf{u}+\mathbf{u}\times\mathbf{v})\big) = (ab-\mathbf{u}\cdot\mathbf{v})^2 + |a\mathbf{v}+b\mathbf{u}+\mathbf{u}\times\mathbf{v}|^2\\
&= a^2b^2 - 2ab\,\mathbf{u}\cdot\mathbf{v} + (\mathbf{u}\cdot\mathbf{v})^2 + (a\mathbf{v}+b\mathbf{u}+\mathbf{u}\times\mathbf{v})\cdot(a\mathbf{v}+b\mathbf{u}+\mathbf{u}\times\mathbf{v})\\
&= a^2b^2 - 2ab\,\mathbf{u}\cdot\mathbf{v} + (\mathbf{u}\cdot\mathbf{v})^2 + a^2|\mathbf{v}|^2 + ab\,\mathbf{v}\cdot\mathbf{u} + a\mathbf{v}\cdot(\mathbf{u}\times\mathbf{v}) + ba\,\mathbf{u}\cdot\mathbf{v} + b^2|\mathbf{u}|^2 + b\mathbf{u}\cdot(\mathbf{u}\times\mathbf{v})\\
&\quad + a(\mathbf{u}\times\mathbf{v})\cdot\mathbf{v} + b(\mathbf{u}\times\mathbf{v})\cdot\mathbf{u} + |\mathbf{u}\times\mathbf{v}|^2
\end{aligned}$$

Since the $\mathbf{u}\times\mathbf{v}$ is always orthogonal to both $\mathbf{u}$ and $\mathbf{v}$ (that is, the dot product of $\mathbf{u}\times\mathbf{v}$ with either $\mathbf{u}$ or $\mathbf{v}$ is always zero), we have

$$\begin{aligned}
N(\alpha\beta) &= a^2b^2 - 2ab\,\mathbf{u}\cdot\mathbf{v} + (\mathbf{u}\cdot\mathbf{v})^2 + a^2|\mathbf{v}|^2 + ab\,\mathbf{v}\cdot\mathbf{u} + ba\,\mathbf{u}\cdot\mathbf{v} + b^2|\mathbf{u}|^2 + |\mathbf{u}\times\mathbf{v}|^2\\
&= a^2b^2 - 2ab\,\mathbf{u}\cdot\mathbf{v} + a^2|\mathbf{v}|^2 + ab\,\mathbf{u}\cdot\mathbf{v} + ab\,\mathbf{u}\cdot\mathbf{v} + b^2|\mathbf{u}|^2 + (\mathbf{u}\cdot\mathbf{v})^2 + |\mathbf{u}\times\mathbf{v}|^2\\
&= a^2b^2 + a^2|\mathbf{v}|^2 + b^2|\mathbf{u}|^2 + (|\mathbf{u}||\mathbf{v}|\cos\theta)^2 + (|\mathbf{u}||\mathbf{v}|\sin\theta)^2 = a^2b^2 + a^2|\mathbf{v}|^2 + b^2|\mathbf{u}|^2 + |\mathbf{u}|^2|\mathbf{v}|^2\\
&= (a^2+|\mathbf{u}|^2)(b^2+|\mathbf{v}|^2) = N(\alpha)N(\beta)
\end{aligned}$$

■

3. Show that the set of nonzero elements of C is a group under multiplication. (Hence, (2) shows that $N: C^*\to\mathbb{R}$ is a homomorphism.)

Proof. First, we note that for $\alpha = (u_0, (u_1, u_2, u_3)) \in C$ we have

$$\alpha = (0, \mathbf{0}) \quad \iff \quad N(\alpha) = u_0^2 + u_1^2 + u_2^2 + u_3^2 = 0$$

(since if $\alpha = (0, \mathbf{0})$, then $N(\alpha) = 0^2 + 0^2 + 0^2 + 0^2 = 0$, and if $\alpha \neq (0, \mathbf{0})$ the $u_i \neq 0$ for some $0 \leq i \leq 3$, in which case $N(\alpha) = u_0^2 + u_1^2 + u_2^2 + u_3^2 \geq u_i^2 > 0$). Therefore the conditions $\alpha \neq (0, \mathbf{0})$ and $N(\alpha) \neq 0$ are exactly equivalent, so

$$C^* = \{\alpha \in C : \alpha \neq (0, \mathbf{0})\} = \{\alpha \in C : N(\alpha) \neq 0\}$$

Then by part 4.2 we have

$$\alpha, \beta \in C^* \quad \iff \quad N(\alpha), N(\beta) \neq 0 \quad \implies \quad N(\alpha\beta) = N(\alpha)N(\beta) \neq 0 \quad \iff \quad \alpha\beta \in C^*$$

so the product operation $C \times C \to C$ restricts to a product $C^* \times C^* \to C^*$. We already verified in part 4.1 that the product operation defined on C is associative, hence its restriction to C^* is also associative. The identity in C^* is given by the element $I = (1, \mathbf{0})$, because for any element $\alpha = (a, \mathbf{u}) \in C$ we have

$$I\alpha = (1, \mathbf{0})(a, \mathbf{u}) = (1a - \mathbf{0} \cdot \mathbf{u}, 1\mathbf{u} + a\mathbf{0} + \mathbf{0} \times \mathbf{u}) = (a - 0, \mathbf{u} + \mathbf{0} + \mathbf{0}) = (a, \mathbf{u}) = \alpha$$

$$\alpha I = (a, \mathbf{u})(1, \mathbf{0}) = (a1 - \mathbf{u} \cdot \mathbf{0}, a\mathbf{0} + 1\mathbf{u} + \mathbf{u} \times \mathbf{0}) = (a - 0, \mathbf{0} + \mathbf{u} + \mathbf{0}) = (a, \mathbf{u}) = \alpha$$

Finally, I claim that each element $\alpha = (a, \mathbf{u}) \in C^*$ has an inverse element, given by

$$\alpha^{-1} = (N(\alpha)^{-1}a, -N(\alpha)^{-1}\mathbf{u})$$

(note that this is well defined because $N(\alpha) \neq 0$ for every $\alpha \in C^*$, so $N(\alpha)^{-1}$ exists). Indeed, we have

$$\begin{aligned}
\alpha\alpha^{-1} &= (a, \mathbf{u})(N(\alpha)^{-1}a, -N(\alpha)^{-1}\mathbf{u}) \\
&= \big(aN(\alpha)^{-1}a - \mathbf{u} \cdot (-N(\alpha)^{-1}\mathbf{u}),\ a(-N(\alpha)^{-1}\mathbf{u}) + N(\alpha)^{-1}a\mathbf{u} + \mathbf{u} \times (-N(\alpha)^{-1}\mathbf{u})\big) \\
&= \big(N(\alpha)^{-1}a^2 + N(\alpha)^{-1}(\mathbf{u} \cdot \mathbf{u}),\ -N(\alpha)^{-1}a\mathbf{u} + N(\alpha)^{-1}a\mathbf{u} - N(\alpha)^{-1}(\mathbf{u} \times \mathbf{u})\big) \\
&= \big(N(\alpha)^{-1}(a^2 + \mathbf{u} \cdot \mathbf{u}),\ -N(\alpha)^{-1}(\mathbf{u} \times \mathbf{u})\big) = \big(N(\alpha)^{-1}(a^2 + |\mathbf{u}|^2),\ N(\alpha)^{-1}\mathbf{0}\big) \\
&= \big(N(\alpha)^{-1}N(\alpha), \mathbf{0}\big) = (1, \mathbf{0}) \\
&= I
\end{aligned}$$

$$\begin{aligned}
\alpha^{-1}\alpha &= (N(\alpha)^{-1}a, -N(\alpha)^{-1}\mathbf{u})(a, \mathbf{u}) \\
&= \big((N(\alpha)^{-1}a)a - (-N(\alpha)^{-1}\mathbf{u}) \cdot \mathbf{u},\ N(\alpha)^{-1}a\mathbf{u} + a(-N(\alpha)^{-1}\mathbf{u}) + (-N(\alpha)^{-1}\mathbf{u}) \times \mathbf{u}\big) \\
&= \big(N(\alpha)^{-1}a^2 + N(\alpha)^{-1}(\mathbf{u} \cdot \mathbf{u}),\ N(\alpha)^{-1}a\mathbf{u} - N(\alpha)^{-1}a\mathbf{u} - N(\alpha)^{-1}(\mathbf{u} \times \mathbf{u})\big) \\
&= \big(N(\alpha)^{-1}(a^2 + \mathbf{u} \cdot \mathbf{u}),\ -N(\alpha)^{-1}(\mathbf{u} \times \mathbf{u})\big) = \big(N(\alpha)^{-1}(a^2 + |\mathbf{u}|^2),\ N(\alpha)^{-1}\mathbf{0}\big) \\
&= \big(N(\alpha)^{-1}N(\alpha), \mathbf{0}\big) = (1, \mathbf{0}) \\
&= I
\end{aligned}$$

where we have used above the fact that for any $\mathbf{u} \in \mathbb{R}^3$ we have $\mathbf{u} \times \mathbf{u} = \mathbf{0}$. Therefore we have verified that C^* has an associative product operation, and identity, and inverses, so we conclude that C^* is a group. ■

4. Show that elements $\alpha \in C$ with integer coefficients and such that $N(\alpha) = 1$ form a subgroup of order 8 in C^*. This is the *quaternion group*.

Proof. From the proof of the lemma under problem 3.2, we know that to prove that the set

$$S = \{\alpha = (u_0, (u_1, u_2, u_3)) \in C : N(\alpha) = 1 \text{ and } u_i \in \mathbb{Z} \text{ for } 0 \leq i \leq 3\}$$

is a subgroup, it suffices to show that S is non-empty (which is evident because $(1, (0, 0, 0)) \in S$) and

$$\alpha, \beta \in S \quad \implies \quad \alpha\beta^{-1} \in S$$

As such, suppose that $\alpha = (u_0, (u_1, u_2, u_3))$ and $\beta = (v_0, (v_1, v_2, v_3))$ are any two elements of S. Then because the map $N : C^* \to \mathbb{R}$ is a group homomorphism by the previous part (and $S \subset C^*$, since S doesn't contain any elements of norm zero), it follows that

$$N(\alpha\beta^{-1}) = N(\alpha)\big(N(\beta)\big)^{-1} = 1 \cdot 1^{-1} = 1$$

and because $u_i, v_i \in \mathbb{Z}$ for all $0 \leq i \leq 3$ it follows that

$$\begin{aligned}\alpha\beta^{-1} &= (u_0, \mathbf{u})(N(\beta)^{-1}v_0, -N(\beta)^{-1}\mathbf{v}) = (u_0, \mathbf{u})(v_0, -\mathbf{v}) = \big(u_0v_0 - \mathbf{u}\cdot(-\mathbf{v}), u_0(-\mathbf{v}) + v_0\mathbf{u} + \mathbf{u}\times(-\mathbf{v})\big)\\ &= \big(u_0v_0 + (u_1,u_2,u_3)\cdot(v_1,v_2,v_3), -u_0(v_1,v_2,v_3) + v_0(u_1,u_2,u_3) - (u_1,u_2,u_3)\times(v_1,v_2,v_3)\big)\\ &= \big(u_0v_0 + u_1v_1 + u_2v_2 + u_3v_3, (-u_0v_1, -u_0v_2, -u_0v_3) + (u_1v_0, u_2v_0, u_3v_0) - (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)\big)\\ &= \big(u_0v_0 + u_1v_1 + u_2v_2 + u_3v_3, (u_1v_0 - u_0v_1 + u_3v_2 - u_2v_3, u_2v_0 - u_0v_2 + u_1v_3 - u_3v_1, u_3v_0 - u_0v_3 + u_2v_1 - u_1v_2)\big)\end{aligned}$$

In particular, because every coefficient of $\alpha\beta^{-1}$ is the sum/difference of products of integers, $\alpha\beta^{-1}$ also has integer coefficients. Since it also has norm 1, it follows that $\alpha\beta^{-1} \in S$, so S is indeed a subgroup. In order to find all of the elements contained within S, we want to find all of the solutions to the equation

$$N\big((u_0,(u_1,u_2,u_3))\big) = u_0^2 + u_1^2 + u_2^2 + u_3^2 = 1$$

for $u_i \in \mathbb{Z}$ $(0 \leq i \leq 3)$. We cannot have $|u_i| \geq 2$ for any of the u_i, or else we would have

$$u_0^2 + u_1^2 + u_2^2 + u_3^2 \geq u_i^2 = |u_i|^2 \geq 2^2 = 4 > 1$$

so we must have $u_i \in \{-1, 0, 1\}$ for all $0 \leq i \leq 3$. If two or more of the u_i are nonzero (i.e., $|u_i|, |u_j| \geq 1$ for some $0 \leq i, j \leq 3$ with $i \neq j$), then we would have

$$u_0^2 + u_1^2 + u_2^2 + u_3^2 \geq u_i^2 + u_j^2 = |u_i|^2 + |u_j|^2 \geq 1^2 + 1^2 = 2 > 1$$

so at most one of the u_i can be nonzero. It is also clear that at least one of them must be nonzero, because if they were all zero we would have $u_0^2 + u_1^2 + u_2^2 + u_3^2 = 0^2 + 0^2 + 0^2 + 0^2 = 0$, so we must have that exactly one of the u_i is nonzero. Therefore we must have $u_i \in \{-1, 1\}$ for some $0 \leq i \leq 3$ and $u_j = 0$ for all $j \neq i$, so the only possible elements of S are the elements

$$\begin{array}{cccc}(1,(0,0,0)) & (-1,(0,0,0)) & (0,(1,0,0)) & (0,(-1,0,0))\\ (0,(0,1,0)) & (0,(0,-1,0)) & (0,(0,0,1)) & (0,(0,0,-1))\end{array}$$

In fact, one can see that all of these solve the equation $u_0^2 + u_1^2 + u_2^2 + u_3^2 = 1$, so all of these are elements of S. As a result, S is a subgroup of C^* of order 8, with the 8 elements it contains listed above. ■

Problem 5P. Give examples:

1. of a group G with two minimal sets of generators of different cardinality,
2. of an infinite group generated by two elements.

Proof. Consider the group $G = \mathbb{Z} \times \mathbb{Z}$ under addition, and the sets

$$A = \{(1,0), (0,1)\}$$

$$B = \{(-1,2), (1,2), (2,1)\}$$

Then we claim that the infinite group G is generated by the two elements of A, and that A and B are minimal generating sets for G of different cardinality. To demonstrate this, we first recall that the subgroup generated by a set S is characterized (in additive notation) by

$$\langle S\rangle = \{n_1s_1 + n_2s_2 + \ldots + n_ks_k : k \in \mathbb{N},\ s_1, s_2, \ldots, s_m \in S,\ n_1, n_2, \ldots, n_m \in \mathbb{Z}\}$$

In particular, $G = \langle A\rangle$ because for any element $(n, m) \in G$ we have

$$(n,m) = n(1,0) + m(0,1) \in \langle A\rangle$$

and $G = \langle B\rangle$ because for any element $(n, m) \in G$ we have

$$(n,m) = (m-2n)(-1,2) + (3n-m)(1,2) + (m-2n)(2,1) \in \langle B\rangle$$

It is clear that A is a minimal generating group because, for instance $(1, 1) \in G$, yet

$$(1,1) \notin \langle(1,0)\rangle = \{n(1,0) : n \in \mathbb{Z}\} = \{(n,0) : n \in \mathbb{Z}\}$$

$$(1,1) \notin \langle(0,1)\rangle = \{m(0,1) : m \in \mathbb{Z}\} = \{(0,m) : m \in \mathbb{Z}\}$$

(that is, no subset of A generates all of G). Similarly, B is also a minimal generating set because none of the three elements of B is contained in the subgroup generated by the other two: by determining the unique solution to each system of linear equations

$$(-1,2) = x(1,2) + y(2,1) \quad \Longrightarrow \quad x = \frac{5}{3},\ y = -\frac{4}{3}$$

$$(1,2) = x(-1,2) + y(2,1) \quad \Longrightarrow \quad x = \frac{3}{5},\ y = \frac{4}{5}$$

$$(2,1) = x(-1,2) + y(1,2) \quad \Longrightarrow \quad x = -\frac{3}{4},\ y = \frac{5}{4}$$

we find that none of them admit integer solutions, so

$$(-1,2) \notin \langle (1,2),(2,1) \rangle = \{n(1,2) + m(2,1) : n, m \in \mathbb{Z}\}$$

$$(1,2) \notin \langle (-1,2),(2,1) \rangle = \{n(-1,2) + m(2,1) : n, m \in \mathbb{Z}\}$$

$$(2,1) \notin \langle (-1,2),(1,2) \rangle = \{n(-1,2) + m(1,2) : n, m \in \mathbb{Z}\}$$

■

Problem 6. Give an example of a tower of groups $K \lhd H \lhd G$ such that K is **not** normal in G.

Proof. Consider the subgroup of the S_4 (the group of permutations of the set $\{1,2,3,4\}$ under composition) generated by the elements

$$r = \begin{cases} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 4 \\ 4 \mapsto 1 \end{cases}$$

$$s = \begin{cases} 1 \mapsto 1 \\ 2 \mapsto 4 \\ 3 \mapsto 3 \\ 4 \mapsto 2 \end{cases}$$

These elements satisfy the relation $r^4 = s^2 = e$, which one can readily verify by composing the permutations:

$$r^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

$$r^4 = (r^2)^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = e$$

$$s^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = e$$

Another relation that we will use in the following calculations is that $rsrs = e$:

$$rs = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

$$rsrs = (rs)^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = e$$

Then consider the chain of subgroups

$$\langle s \rangle \leq \langle r^2, s \rangle \leq \langle r, s \rangle$$

(the inclusions are evident because $s \in \langle r^2, s \rangle \Longrightarrow \langle s \rangle \leq \langle r^2, s \rangle$ and $r^2, s \in \langle r, s \rangle \Longrightarrow \langle r^2, s \rangle \leq \langle r, s \rangle$, since the subgroup generated by a set is the intersection of every subgroup containing it and hence is a subgroup of any subgroup containing it). By the lemma below, $\langle s \rangle$ is normal in $\langle r^2, s \rangle$ because, conjugating each element of $\{s\}$ by each element of $\{r^2, s\}$, we have

$$r^2 s (r^2)^{-1} = r^2 s r^2 = r(rsr)r = r(s)r = s \in \langle s \rangle$$

$$s(s)s^{-1} = s \in \langle s \rangle$$

(note that above we have used the relations $(r^2)^2 = r^4 = e \implies r^2 = (r^2)^{-1}$, $s^2 = e \implies s = s^{-1}$, and $(rsr)s = e \implies rsr = s^{-1} = s$). Similarly, $\langle r^2, s \rangle$ is normal in $\langle r, s \rangle$ because if we conjugate each element of $\{r^2, s\}$ by each element of $\{r, s\}$ we find

$$r(r^2)r^{-1} = r^2 \in \langle r^2, s \rangle$$

$$s(r^2)s^{-1} = sr^2s = s(r^2s(r^2)^{-1})r^2 = s(s)r^2 = r^2 \in \langle r^2, s \rangle$$

$$rsr^{-1} = rsr^3 = (rsr)r^2 = (s)r^2 \in \langle r^2, s \rangle$$

$$s(s)s^{-1} = s \in \langle r^2, s \rangle$$

However, $\langle s \rangle$ is not normal in $\langle r, s \rangle$ because conjugating an element of $\langle s \rangle$ by an element of $\langle r, s \rangle$ does not necessarily yield another element of $\langle s \rangle$: since $s^2 = e$, it follows that $s^n = e$ if n is even and $s^n = s$ if n is odd, so

$$\langle s \rangle = \{s^n : n \in \mathbb{Z}\} = \{e, s\}$$

If we conjugate the element $s \in \langle s \rangle$ by the element $r \in \langle r, s \rangle$, then we have

$$rsr^{-1} = s(r^2) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

In particular, this element is not the identity, and it is distinct from

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

so $rsr^{-1} \notin \{e, s\} = \langle s \rangle$. Therefore

$$\langle s \rangle \lhd \langle r^2, s \rangle \lhd \langle r, s \rangle$$

yet

$$\langle s \rangle \not\lhd \langle r, s \rangle$$

■

Lemma. Let G be a group, and suppose that A and B are subsets of G such that $\langle A \rangle \leq \langle B \rangle$. If

$$(\forall b \in B)(\forall a \in A) : bab^{-1} \in \langle A \rangle \tag{1}$$

then $\langle A \rangle$ is normal in $\langle B \rangle$.

Proof. Given sets A and B satisfying property 1 above, first we show that

$$(\forall b \in B)(\forall a \in \langle A \rangle) : bab^{-1} \in \langle A \rangle \tag{2}$$

Since the group $\langle A \rangle$ is characterized by

$$\langle A \rangle = \{a_1 a_2 \dots a_k : k \in \mathbb{N},\ a_1, a_2, \dots, a_k \in A\}$$

(note that the elements a_i for $1 \leq i \leq k$ are not necessarily distinct), then for any $a \in \langle A \rangle$ and any $b \in B$ we can write

$$bab^{-1} = b(a_1 a_2 \dots a_k)b^{-1} = (ba_1b^{-1})(ba_2b^{-1}) \dots (ba_kb^{-1})$$

for some elements $a_i \in A$ $(1 \leq i \leq k)$. Since $(ba_ib^{-1}) \in \langle A \rangle$ for all $1 \leq i \leq k$ by property 1 and $\langle A \rangle$ is closed under multiplication (since it is a subgroup), it follows that

$$bab^{-1} = (ba_1b^{-1})(ba_2b^{-1}) \dots (ba_kb^{-1}) \in \langle A \rangle$$

for any $a \in \langle A \rangle$ and $b \in B$. Now we show that

$$(\forall b \in \langle B \rangle)(\forall a \in \langle A \rangle) : bab^{-1} \in \langle A \rangle \tag{3}$$

(this is one of the equivalent conditions for normality discussed in class). Since $\langle B\rangle$ can similarly be characterized by

$$\langle B\rangle = \{b_1 b_2 \dots b_k : k \in \mathbb{N},\ b_1, b_2, \dots, b_k \in B\}$$

then for any $a \in \langle A\rangle$ and any $b \in \langle B\rangle$ we can write

$$bab^{-1} = (b_1 b_2 \dots b_{k-1} b_k) a (b_1 b_2 \dots b_{k-1} b_k)^{-1} = b_1 b_2 \dots b_{k-1} (b_k a b_k^{-1}) b_{k-1}^{-1} \dots b_2^{-1} b_1^{-1}$$

for some $b_i \in B$ $(1 \leq i \leq k)$. By property 2 we have that $b_k a b_k^{-1} \in \langle A\rangle$, and similarly $b_{k-1}(b_k a b_k^{-1}) b_{k-1}^{-1} \in \langle A\rangle$ by another application of property 2, and so forth until we conclude that

$$bab^{-1} = b_1 (b_2 \dots b_k a b_k^{-1} \dots b_2^{-1}) b_1^{-1} \in \langle A\rangle$$

Therefore property 3 holds, so as claimed we have $\langle A\rangle \trianglelefteq \langle B\rangle$. ■

Problem 7P. Prove the second isomorphism theorem.

Proof. The statement of the second isomophism theorem in Dummit and Foote reads:

> Let G be a group, let A and B be subgroups of G, and assume $A \leq N_G(B)$. Then AB is a subgroup of G, $B \trianglelefteq AB$, $A \cap B \trianglelefteq A$, and $(AB)/B \cong A/(A \cap B)$.

First we demonstrate that

$$AB = \{ab : a \in A, b \in B\}$$

is a subgroup of G by showing that AB is a nonempty subset of G such that

$$a_1 b_1, a_2 b_2 \in AB \quad \Longrightarrow \quad a_1 b_1 (a_2 b_2)^{-1} \in AB$$

(see the proof of the lemma under problem 2P to know why this shoes that AB is a subgroup of G). Because A and B are both subgroups of G, then $e \in A$ and $e \in B$, so it follows that

$$e = ee \in \{ab : a \in A, b \in B\} = AB$$

and thus AB is nonempty. Next, given any $a_1 b_1, a_2 b_2 \in AB$ (where $a_1, a_2 \in A$ and $b_1, b_2 \in B$), the fact that A is a subgroup means that

$$a_1, a_2 \in A \quad \Longrightarrow \quad a_1 a_2^{-1} \in A$$

and the fact that B is a subgroups means that

$$b_1, b_2 \in B \quad \Longrightarrow \quad b_1 b_2^{-1} \in B$$

Because every element of A normalizes B (since $A \leq N_G(B)$), then

$$a_2 \in A,\ b_1 b_2^{-1} \in B \quad \Longrightarrow \quad a_2 (b_1 b_2^{-1}) a_2^{-1} \in B$$

Therefore

$$a_1 b_1 (a_2 b_2)^{-1} = a_1 b_1 b_2^{-1} a_2^{-1} = (a_1 a_2^{-1})(a_2 b_1 b_2^{-1} a_2^{-1}) \in AB$$

Thus AB is indeed a subgroup of G. Next, it is clear that B is a subgroup of AB because B is a subgroup of G (i.e., B contains the identity and is closed under inverses and multiplication) and a subset of AB: because $e \in A$,

$$B = \{eb : b \in B\} \subset \{ab : a \in A, b \in B\} = AB$$

That B is normal in AB follows because for every element $b \in B$ and any element $a_0 b_0 \in AB$ we have

$$(a_0 b_0) b (a_0 b_0)^{-1} = a_0 (b_0 b b_0^{-1}) a_0^{-1}$$

Because $b_0 b b_0^{-1} \in B$ (since B is a subgroup) and B is normalized by every element of A, it follows that $a_0 (b_0 b b_0^{-1}) a_0^{-1} \in B$, so B is also normalized by every element AB and hence B is normal in AB. As a result, the quotient group $(AB)/B$ is well-defined, so we can consider the map $f : A \to (AB)/B$ given by

$$f(a) = aB$$

for $a \in A$ (where $ae \in AB$ because $e \in B$, so $aB = aeB \in (AB)/B$). That f is a homomorphism follows from the induced multiplication in the quotient group

$$f(a_1)f(a_2) = (a_1B)(a_2B) = (a_1a_2)B = f(a_1a_2)$$

This homomorphism is also surjective, because for any coset $abB \in (AB)/B$ (with $a \in A$ and $b \in B$) we have

$$abB = aB = f(a)$$

(note that above we used the fact that $bB = B \iff b \in B$) Because the identity in $(AB)/B$ is the coset $eB = B$, the kernel of this map is

$$\ker f = \{a \in A : f(a) = e_{(AB)/B}\} = \{a \in A : aB = B\} = \{a \in A : a \in B\} = A \cap B$$

In particular, because $A \cap B$ is the kernel of the homomorphism $f : A \to (AB)/B$, it follows that $A \cap B$ is a normal subgroup of A, and by applying the first isomorphism theorem we have

$$A/(A \cap B) = A/\ker f \cong \operatorname{im} f = (AB)/B$$

■

Problem 8P. Prove the third isomorphism theorem.

Proof. The statement of the third isomorphism theorem in Dummit and Foote is

> Let G be a group and let H and K be normal subgroups of G with $H \leq K$. Then $K/H \trianglelefteq G/H$ and
>
> $$(G/H)/(K/H) \cong G/K$$

Since H and K are both normal subgroups of G, then we can form the quotient groups G/H and G/K and consider the map $f : (G/H) \to (G/K)$ defined by

$$f(gH) = gK$$

First, note that this map is well defined: because $H \leq K$ we have

$$g_1H = g_2H \iff g_1^{-1}g_2 \in H \implies g_1^{-1}g_2 \in K \iff g_1K = g_2K$$

Thus if $g_1H = g_2H$, then $f(g_1H) = g_1K = g_2K = f(g_2H)$. This map is also a group homomorphism, because for any $g_1H, g_2H \in (G/H)$ we have

$$f(g_1H)f(g_2H) = (g_1K)(g_2K) = g_1g_2K = f(g_1g_2H)$$

and it is clearly surjective because for any coset $gK \in (G/K)$ we have

$$gK = f(gH)$$

The kernel of this map is given by

$$\ker f = \{gH \in (G/H) : f(gH) = e_{G/K}\} = \{gH \in (G/H) : gK = K\} = \{gH \in (G/H) : g \in K\} = K/H$$

In particular, it follows that K/H is a normal subgroup of G/H, and by the first isomorphism theorem we have

$$(G/H)/(K/H) = (G/H)/\ker f \cong \operatorname{im} f = G/K$$

■

Problem 1P.

Let G be a p-group, and H be a non-trivial normal subgroup of G. Show that $H \cap Z(G) \neq 1$.

Proof. Let $|G| = p^k$ for some prime p and some integer $k \geq 1$, and consider the action of conjugation of G on H (that is, the map $* : G \times H \to H$ given by $g * h = ghg^{-1}$). Note that this map does indeed have the given codomain, because the fact that H is normal in G means that $ghg^{-1} \in H$ for all $h \in H$, and it is indeed a left action because for all $h \in H$ and for all $g_1, g_2 \in G$ we have

$$e * h = ehe^{-1} = h$$

$$g_1 * (g_2 * h) = g_1 * (g_2 h g_2^{-1}) = g_1 g_2 h g_2^{-1} g_1^{-1} = (g_1 g_2) h (g_1 g_2)^{-1} = (g_1 g_2) * h$$

Then, since the center of G is defined by

$$Z_G = \{x \in G : (\forall g \in G)\ gxg^{-1} = x\}$$

and the orbit of $h \in H$ is defined

$$\mathrm{Orb}(h) = \{g * h : g \in G\} = \{ghg^{-1} : g \in G\}$$

it follows that the orbit of $h \in H$ is trivial (that is, the $\mathrm{Orb}(h) = \{h\}$) if and only if $h \in (H \cap Z_G)$. Indeed,

$$\begin{aligned}\big\{h \in H : \mathrm{Orb}(h) = \{h\}\big\} &= \big\{h \in H : \{ghg^{-1} : g \in G\} = \{h\}\big\} \\ &= \big\{h \in H : (\forall g \in G)\ ghg^{-1} = h\big\} = \{h \in H : h \in Z_G\} = H \cap Z_G\end{aligned}$$

Then, if we let S be a set that contains exactly one representative of each nontrivial orbit, we have by the class formula

$$|H| = |H \cap Z_G| + \sum_{h \in S} |\,\mathrm{Orb}(h)| = |H \cap Z_G| + \sum_{h \in S} [G : \mathrm{Stab}(h)]$$

Because $\mathrm{Stab}(h)$ is a subgroup of G for each $h \in S$, then the fact that the order of $\mathrm{Stab}(h)$ must divide $|G| = p^k$ means that $|\,\mathrm{Stab}(h)| = p^n$ for some integer $0 \leq n \leq k$. Since the orbit of h is nontrivial for all $h \in S$, then $|\,\mathrm{Orb}(h)| = [G : \mathrm{Stab}(h)] = |G|/|\,\mathrm{Stab}(h)| \neq 1$, so we cannot have $|\,\mathrm{Stab}(h)| = |G| = p^k$ in particular, thus $0 \leq n < k$ and

$$[G : \mathrm{Stab}(h)] = \frac{|G|}{|\,\mathrm{Stab}(h)|} = \frac{p^k}{p^n} = p^{k-n}$$

That is, p divides $[G : \mathrm{Stab}(h)]$ because $k - n > 0$. On the other hand, because H is also a subgroup of G, then the order of $|H|$ also divides $|G| = p^k$ and thus $|H| = p^m$ for some $0 \leq m \leq k$. Since H is a nontrivial subgroup of G, then $|H| \neq 1$, so $0 < m \leq k$ and in particular p divides $|H| = p^m$. Then because p divides all of the terms on the right side of the equation

$$|H \cap Z_G| = |H| - \sum_{h \in S} [G : \mathrm{Stab}(h)]$$

it follows that p divides $|H \cap Z_G|$, and in particular $|H \cap Z_G| \neq 1$. ■

Problem 2.

Let p, q be prime numbers.

1. (2.1P) Show that $\mathrm{Aut}(\mathbb{Z}/qZ) \simeq \mathbb{Z}/(q-1)\,\mathbb{Z}$.

   Proof. To begin, we note that every group homomorphism $f : \mathbb{Z}/q\,\mathbb{Z} \to \mathbb{Z}/q\,\mathbb{Z}$ is determined entirely by its action on the element $[1] \in \mathbb{Z}/q\,\mathbb{Z}$, because for any element $[n] \in \mathbb{Z}/q\,\mathbb{Z}$ we have

   $$f([n]) = f(n[1]) = nf([1])$$

   ■

2. (2.2P) Classify all groups of order pq.

   Proof. ■

Problem 3.

(Sylow subgroups of symmetric groups)

(1). What is the order of a Sylow 2-subgroup of S_{2^n}?

Proof. The order of the Sylow 2-subgroup of S_{2^n} is 2^{k_n}, where k_n is the largest natural number such that 2^{k_n} divides $|S_{2^n}| = (2^n)!$, or equivalently the number of factors of 2 contained in the prime factorization of $(2^n)!$. Then, viewing k_n as a function $k : \mathbb{N} \to \mathbb{N}$ that maps the integer n to precisely the aforementioned value of k_n, we note that

$$k_{n+1} = 2^n + k_n$$

This follows because the number of factors of 2 in the prime factorization of $(2^{n+1})!$ the same as the number of factors of 2 in the prime factorization of only the even factors of $(2^{n+1})!$ (since the odd factors contain no powers of 2): that is, the number of factors of 2 in

$$\prod_{1 \le 2j \le 2^{n+1}} (2j) = \prod_{j=1}^{2^n} (2j) = 2^{2^n} \prod_{j=1}^{2^n} j = 2^{2^n} (2^n)!$$

This contains 2^n factors of 2 plus the number of factors of 2 in $(2^n)!$, thus $k^{n+1} = 2^n + k_n$ as claimed. We further claim that

$$k_n = 2^n - 1$$

which follows readily from induction. In the base case of $n = 1$, then $(2^1)! = 2! = 2$, so clearly $k_1 = 1$ because there is exactly one factor of 2 present, and if $k_n = 2^n - 1$ for some $n \in \mathbb{N}$ then we have that

$$k_{n+1} = 2^n + k_n = 2^n + (2^n - 1) = 2(2^n) - 1 = 2^{n+1} - 1$$

Therefore the order of the Sylow 2-subgroup of S_{2^n} is

$$2^{k_n} = 2^{2^n - 1}$$

■

(2). Give an explicit description of a Sylow 2-subgroup of S_{2^n}.

Proof. First, note that S_{2^n} can be naturally represented via the group P_{2n} of $2^n \times 2^n$ permutation matrices (matrices that contain exactly one unit element in each row and column and zeros everywhere else) under matrix multiplication. Indeed, we can form a map $\rho : S_{2^n} \to P_{2^n}$ by

$$\rho(f)_{ij} = \mathbb{I}[i = f(j)] \qquad (1 \le i, j \le 2^n)$$

where $f \in S_{2^n}$ is a bijection of the set $\{1, 2, \ldots, 2^n\}$ with itself and $\mathbb{I}[P]$ is the indicator function which, given a proposition P, equals 1 if P is true and 0 if P is false. Note that this map does indeed have the claimed codomain: for each column j_0 there is exactly one value of i satisfying $i = f(j_0)$ because the function must be single-valued, and because f is a bijection and thus has an inverse, for each row i_0 there is exactly one value of j satisfying $i_0 = f(j)$, namely $j = f^{-1}(i_0)$. Thus for any $f \in S_{2^n}$ we have that each row and each column has exactly one entry equal to 1 and all other entries equal to 0, thus $\rho(f) \in P_{2^n}$. Then, taking any $f, g \in S_{2^n}$, let us verify that this is a group homomorphism:

$$\big(\rho(f)\rho(g)\big)_{ij} = \sum_{k=1}^{2^n} \big(\rho(f)\big)_{ik} \big(\rho(g)\big)_{kj} = \sum_{k=1}^{2^n} \mathbb{I}[i = f(k)]\, \mathbb{I}[k = g(j)]$$

Note that the product of the indicator functions $\mathbb{I}[P]$ and $\mathbb{I}[Q]$ is the same as the indicator function $\mathbb{I}[P \wedge Q]$, because

$$\mathbb{I}[P]\, \mathbb{I}[Q] = 1 \quad \iff \quad \mathbb{I}[P] = 1 \text{ and } \mathbb{I}[Q] = 1 \quad \iff \quad P \wedge Q \quad \iff \quad \mathbb{I}[P \wedge Q]$$

Therefore

$$\big(\rho(f)\rho(g)\big)_{ij} = \sum_{k=1}^{2^n} \mathbb{I}[i = f(k)]\, \mathbb{I}[k = g(j)] = \sum_{k=1}^{2^n} \mathbb{I}[i = f(k) \wedge k = g(j)]$$

Now suppose for some particular i and j that $i = f(g(j))$. Then $k = g(j)$ is the unique value of k satisfying $i = f(k)$ and $k = g(j)$, in which case there is exactly one nonzero term in the sum and thus $\sum_{k=1}^{2^n} \mathbb{I}[i = f(k) \wedge k = g(j)] = 1$. On the other hand, if $i \neq f(g(j))$ for some i and j, then there is no value of k satisfying $i = f(k)$ and $k = g(j)$ without contradicting that $i \neq f(g(j))$, so every term in the sum is zero and thus $\sum_{k=1}^{2^n} \mathbb{I}[i = f(k) \wedge k = g(j)] = 0$. That is,

$$\big(\rho(f)\rho(g)\big)_{ij} = \sum_{k=1}^{2^n} \mathbb{I}[i = f(k) \wedge k = g(j)] = \begin{cases} 1 & \text{if } i = f(g(j)) \\ 0 & \text{if } i \neq f(g(j)) \end{cases} = \mathbb{I}[i = f(g(j))] = \mathbb{I}[i = (f \circ g)(j)] = \rho(f \circ g)_{ij}$$

One can also verify that this homomorphism is injective by showing that the kernel is trivial. Suppose that $\rho(f) = I$ for some $f \in S_{2^n}$: that is, suppose

$$\mathbb{I}[i = f(j)] = \rho(f)_{ij} = I_{ij} = \delta_{ij} = \mathbb{I}[i = j]$$

From the equality of the indicators on the far left and far right of the above equation, we have for all $1 \leq i, j \leq 2^n$ that

$$i = f(j) \quad \iff \quad i = j$$

In other words, $f(j) = j$ for all $1 \leq j \leq 2^n$, so f is exactly the identity permutation on $\{1, 2, \ldots, 2^n\}$. From this, we can also conclude that ρ is surjective, because the cardinality of S_{2^n} and P_{2^n} are the same (and both finite) and ρ is an injective map from S_{2^n} to P_{2^n}. To see this, note that, if we build each permutation matrix column by column, there are exactly 2^n choices for the location of the 1 in the first column, and because we cannot place the 1 in the second column in the same row as the 1 in the first column there are $2^n - 1$ choices of where to put the n in the second column, and $2^n - 2$ choices for the third column and so forth. Thus

$$|P_{2^n}| = (2^n)(2^n - 1)(2^n - 2) \ldots (3)(2)(1) = (2^n)! = |S_{2^n}|$$

In fact, that ρ is surjective and respects the multiplication law in S_{2^n} retroactively justifies the assertion that P_{2^n} is a group, because for any matrix $A \in P_{2^n}$ we have that $A = \rho(f)$ for some $f \in S_{2^n}$, in which case

$$\big(\rho(f)\rho(f^{-1})\big)_{ij} = \rho(f \circ f^{-1})_{ij} = \rho(\mathrm{id})_{ij} = \mathbb{I}[i = \mathrm{id}(j)] = \mathbb{I}[i = j] = \delta_{ij} = I_{ij}$$

$$\big(\rho(f^{-1})\rho(f)\big)_{ij} = \rho(f^{-1} \circ f)_{ij} = \rho(\mathrm{id})_{ij} = \mathbb{I}[i = \mathrm{id}(j)] = \mathbb{I}[i = j] = \delta_{ij} = I_{ij}$$

(that is, the matrix $\rho(f^{-1})$ is the inverse of the matrix $\rho(f)$). Then since $S_{2^n} \cong P_{2^n}$, it suffices to find a Sylow 2-subgroup of P_{2^n}. Then define what we will call a 2^n block matrix recursively as follows:

1. A 2^1 block matrix is a matrix of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

2. For integers $n > 1$, a 2^n block matrix is a matrix of the form

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$$

   where A and B are 2^{n-1} block matrices.

Note that a 2^n block matrix is a $2^n \times 2^n$ matrix (a 2^1 block matrix is a 2×2 matrix, and the dimensions of a 2^n block matrix are twice the dimension of a 2^{n-1} block matrix). We can also see inductively that these are indeed permutation matrices: clearly every 2^1 block matrices are have exactly one unit element in each row and column and zeros elsewhere, and assuming that every 2^{n_0} block matrix has exactly one unit element in each row and column and zeros elsewhere for some $n_0 \in \mathbb{N}$, it is clear that the same is true of any 2^{n_0+1} block matrices

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$$

(because every column contains exactly one column of one of A and B and zeros everywhere else, and every row contains exactly one row of one of A and B and zeros everywhere else, and each column and row of A and B contains exactly a single 1 and zeros everywhere else by the induction hypothesis). Furthermore, it

follows by induction that multiplying two 2^n block matrices produces another 2^n block matrix: in the case of $n = 1$, we have

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$
$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$$

and suppose that for some $n_0 \in \mathbb{N}$ that we have that every product of two 2^{n_0} block matrices is a 2^{n_0} block matrix. Then any product between two 2^{n_0+1} block matrices must have one of the following forms:

$$\begin{bmatrix}A & 0\\0 & B\end{bmatrix}\begin{bmatrix}C & 0\\0 & D\end{bmatrix} = \begin{bmatrix}AC & 0\\0 & BD\end{bmatrix} \qquad \begin{bmatrix}0 & A\\B & 0\end{bmatrix}\begin{bmatrix}0 & C\\D & 0\end{bmatrix} = \begin{bmatrix}AD & 0\\0 & BC\end{bmatrix}$$
$$\begin{bmatrix}A & 0\\0 & B\end{bmatrix}\begin{bmatrix}0 & C\\D & 0\end{bmatrix} = \begin{bmatrix}0 & AC\\BD & 0\end{bmatrix} \qquad \begin{bmatrix}0 & A\\B & 0\end{bmatrix}\begin{bmatrix}C & 0\\0 & D\end{bmatrix} = \begin{bmatrix}0 & AD\\BC & 0\end{bmatrix}$$

The matrices AC, AD, BC, and BD are all 2^{n_0} block matrices by the induction hypothesis, so all of the above products are 2^{n_0+1} block matrices by definition. In particular, because the set B_{2^n} of 2^n block matrices is a set of $2^n \times 2^n$ permutation matrices, we have that $B_{2^n} \subset P_{2^n}$, and because B_{2^n} is a nonempty subset of the finite group P_{2^n} that is closed under multiplication, it follows that B_{2^n} is a subgroup of P_{2^n} by the lemma below. Now we claim that the order of B_{2^n} is exactly 2^{2^n-1}, which we prove by induction. In the case of $n = 1$, there are only two 2^1 block matrices, namely

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$$

thus

$$|B_{2^1}| = 2 = 2^{2^1-1}$$

Next, suppose that $|B_{2^{n_0}}| = 2^{2^{n_0}-1}$ for some $n_0 \in \mathbb{N}$. Then the set $B_{2^{n_0+1}}$ can be partitioned in matrices of the form

$$\begin{bmatrix}A & 0\\0 & B\end{bmatrix}$$

and matrices of the form

$$\begin{bmatrix}0 & A\\B & 0\end{bmatrix}$$

for $A, B \in B_{2^{n_0}}$. The number matrices of the former type is exactly equal to the number of possible combinations of choices for A and B in $B_{2^{n_0}}$, namely

$$|B_{2^{n_0}}||B_{2^{n_0}}| = 2^{2^{n_0}-1}2^{2^{n_0}-1} = 2^{2(2^{n_0}-1)} = 2^{2^{n_0+1}-2}$$

and the number of matrices of the latter type is equal to the same. Thus the total number of 2^{n_0+1} block matrices is given by

$$|B_{2^{n_0+1}}| = 2^{2^{n_0+1}-2} + 2^{2^{n_0+1}-2} = 2(2^{2^{n_0+1}-2}) = 2^{2^{n_0+1}-1}$$

Thus B_{2^n} is a subgroup of $P_{2^n} \cong S_{2^n}$ with cardinality 2^{2^n-1}, so B_{2^n} is exactly the Sylow 2-subgroup of P_{2^n} and hence is isomorphic to the Sylow 2-subgroup of S_{2^n}.

■

Lemma. Let G be a finite group, and suppose that S is a nonempty subset of G that is closed under multiplication. Then S is a subgroup of G.

Proof. Since we already know that S is closed under multiplication, all that remains in order to show that S is a subgroup is demonstrating that it is closed under inverses and contains the identity. In order to show the former, let a be any element of S (at least one such element exists by the fact that S is nonempty). Then the order of a is finite (otherwise the cyclic subgroup generated by a in G would contain infinitely many elements, contradicting the fact that G is finite), so there is some $n \geq 1$ such that $a^n = e$. Because S is closed under multiplication, it follows that $e = a^n \in S$ because $a \in S$. Thus we know that S contains e and hence $e^{-1} = e$, so in order to show that S contains the inverse of every other element as well, suppose that $a \neq e$. Then $n \geq 2$, in which case $n - 1 \geq 1$ and thus a^{n-1} is also contained S, where a^{n-1} satisfies

$$a^{n-1}a = aa^{n-1} = a^n = e$$

That is, $a^{-1} = a^{n-1} \in S$. Thus S is a subgroup of G.

■

Problem 4.

Let $\mathbb{F}_p$ be a finite field of p elements. (Here, p is a prime. The finite field $\mathbb{F}_p$ coincides with $\mathbb{Z}/p\mathbb{Z}$ as an abelian group. The non-zero elements in $\mathbb{Z}/p\mathbb{Z}$ also form a group with respect to multiplication mod p. Hence, in $\mathbb{F}_p$ we can both add and multiply, and these operations are "nicely" compatible. This allows to consider matrices over $\mathbb{F}_p$ the same way we consider matrices over $\mathbb{C}$.)

(1). Find the order of the finite group $\mathrm{GL}_n(\mathbb{F}_p) \stackrel{\text{def}}{=} \{\text{invertible } n \times n \text{ matrices over } \mathbb{F}_p\}$.

Proof. Let $[n]$ denote the equivalence class $n + p\mathbb{Z}$ in the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, and let $M_n(\mathbb{F}_p)$ denote the set of all $n \times n$ matrices with entries in $\mathbb{F}_p$. Note that there exists a matrix that acts as the identity under the operation of matrix multiplication in $M_n(\mathbb{F}_p)$, namely the matrix

$$I_{ij} = [\delta_{ij}]$$

for $1 \leq i, j \leq n$, where δ_{ij} is the Kronecker delta

$$\delta_{ij} = \mathbb{I}[i = j] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

For any matrix $A \in M_n(\mathbb{F}_p)$ with entries $[a_{ij}]$ for $1 \leq i, j \leq n$, the matrix I satisfies

$$(IA)_{ij} = \sum_{k=1}^{n} I_{ik}A_{kj} = \sum_{k=1}^{n} [\delta_{ik}][a_{kj}] = \sum_{k=1}^{n} [\delta_{ik}a_{kj}] = \big[\sum_{k=1}^{n} \delta_{ik}a_{kj}\big] = [a_{ij}] = A_{ij}$$

$$(AI)_{ij} = \sum_{k=1}^{n} A_{ik}I_{kj} = \sum_{k=1}^{n} [a_{ik}][\delta_{kj}] = \sum_{k=1}^{n} [a_{ik}\delta_{kj}] = \big[\sum_{k=1}^{n} a_{ik}\delta_{kj}\big] = [a_{ij}] = A_{ij}$$

For any column vector $\mathbf{a} \in \mathbb{F}_p^n$ with entries $[a_j]$ for $1 \leq j \leq n$, the identity matrix also satisfies

$$(I\mathbf{a})_i = \sum_{j=1}^{n} I_{ij}\mathbf{a}_j = \sum_{j=1}^{n} [\delta_{ij}][a_j] = \sum_{j=1}^{n} [\delta_{ij}a_j] = \big[\sum_{j=1}^{n} \delta_{ij}a_j\big] = [a_i] = \mathbf{a}_i$$

Then we claim that a matrix $A \in M_n(\mathbb{F}_p)$ is invertible (i.e., there is a matrix $A^{-1} \in M_n(\mathbb{F}_p)$ such that $A^{-1}A = AA^{-1} = I$) if and only if its column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{F}_p^n$ are linearly independent (that is, if and only if $[\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_n]\mathbf{a}_n = \mathbf{0} \implies [\lambda_1] = [\lambda_2] = \cdots = [\lambda_n] = [0]$). Of course, if we write $\mathbf{x} = \big[[\lambda_1], [\lambda_2], \ldots, [\lambda_n]\big]^T \in \mathbb{F}_p^n$, then

$$A\mathbf{x} = [\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_n]\mathbf{a}_n$$

so the columns of A are linearly independent if and only if

$$A\mathbf{x} = \mathbf{0} \quad \implies \quad \mathbf{x} = \mathbf{0}$$

For the forwards direction, suppose that A is an invertible matrix: then for any $\mathbf{x} \in \mathbb{F}_p^n$ we have

$$A\mathbf{x} = \mathbf{0} \quad \implies \quad \mathbf{x} = I\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{0} = \mathbf{0}$$

and hence the columns of A are linearly independent. For the reverse direction, suppose that columns of A are linearly independent: then the map $T : \mathbb{F}_p^n \to \mathbb{F}_p^n$ given by $T(\mathbf{x}) = A\mathbf{x}$ is injective, since

$$T(\mathbf{x}) = T(\mathbf{y}) \quad \implies \quad A\mathbf{x} = A\mathbf{y} \quad \implies \quad A\mathbf{x} - A\mathbf{y} = A(\mathbf{x} - \mathbf{y}) = \mathbf{0} \quad \implies \quad \mathbf{x} - \mathbf{y} = \mathbf{0} \quad \implies \quad \mathbf{x} = \mathbf{y}$$

Since $\mathbb{F}_p^n$ is a finite set (it has cardinality $|\,\mathbb{F}_p^n\,| = |\,\mathbb{F}_p\,|^n = p^n$), then any injection from $T : \mathbb{F}_p^n \to \mathbb{F}_p^n$ is necessarily also a surjection, so in particular for each column $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in \mathbb{F}_p^n$ of the identity matrix there is some vector $\mathbf{b}_i \in \mathbb{F}_p^n$ such that

$$A\mathbf{b}_i = \mathbf{e}_i$$

Then we can construct the matrix $B = [\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_n] \in M_n(\mathbb{F}_p)$ with the property

$$AB = A[\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_n] = [A\mathbf{b}_1\ A\mathbf{b}_2\ \cdots\ A\mathbf{b}_n] = [\mathbf{e}_1\ \mathbf{e}_2\ \cdots\ \mathbf{e}_n] = I$$

Note that B also has linearly independent columns, since

$$B\mathbf{x} = \mathbf{0} \quad \Longrightarrow \quad \mathbf{x} = I\mathbf{x} = (AB)\mathbf{x} = A(B\mathbf{x}) = A\mathbf{0} = \mathbf{0}$$

so we can construct a right inverse C for B as well using the same procedure. Then, since $BC = I$, we have

$$BA = BAI = BA(BC) = B(AB)C = BIC = BC = I$$

so B is the left inverse of A as well, thus $B = A^{-1}$ and hence A is invertible.

Since a matrix $A \in M_n(\mathbb{F}_p)$ is invertible if and only if its columns are linearly independent, then the number of invertible matrices $|\operatorname{GL}_n(F_p)|$ is exactly the number of ways that we can choose the n columns of a matrix so that they are linearly independent. To do this, we note that if $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}$ is a nonempty set of linearly independent vectors in $\mathbb{F}_p^n$ and $\mathbf{a}_{m+1}$ is any arbitrary vector in $\mathbb{F}_p^n$, then

$$\{\mathbf{a}_1, \mathbf{a_2}, \ldots, \mathbf{a}_m, \mathbf{a}_{m+1}\} \text{ is linearly independent} \quad \Longleftrightarrow \quad \mathbf{a}_{m+1} \notin \operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}$$

where we define

$$\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\} = \{[\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_m]\mathbf{a}_m : [\lambda_1], [\lambda_2], \ldots, [\lambda_m] \in \mathbb{F}_p\}$$

First, we prove

$$\{\mathbf{a}_1, \ldots, \mathbf{a}_{m+1}\} \text{ is not linearly independent} \quad \Longleftarrow \quad \mathbf{a}_{m+1} \in \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$$

If $\mathbf{a}_{m+1} \in \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, then there exist some coefficients $[\lambda_1], [\lambda_2], \ldots, [\lambda_m] \in \mathbb{F}_p$ such that

$$\mathbf{a}_{m+1} = [\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_m]\mathbf{a}_m$$

Then we have that

$$[\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_m]\mathbf{a}_m + [-1]\mathbf{a}_{m+1} = \mathbf{0}$$

while not all of the coefficients $[\lambda_1], \ldots, [\lambda_m], [-1]$ are zero, so in particular the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_{m+1}\}$ are not linearly independent. Next we prove

$$\{\mathbf{a}_1, \ldots, \mathbf{a}_{m+1}\} \text{ is not linearly independent} \quad \Longrightarrow \quad \mathbf{a}_{m+1} \in \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$$

Because the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_{m+1}\}$ are not linearly independent, then there is some nontrivial solution to the equation

$$[\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_m]\mathbf{a}_m + [\lambda_{m+1}]\mathbf{a}_{m+1} = \mathbf{0}$$

(that is, with not all of the coefficients $[\lambda_1], \ldots, [\lambda_m], [\lambda_{m+1}]$ equal to zero). Then we cannot have $[\lambda_{m+1}] = [0]$, or else it would be true that the coefficients $[\lambda_1], \ldots, [\lambda_m]$ are not all zero, yet satisfy

$$[\lambda_1]\mathbf{a}_1 + [\lambda_2]\mathbf{a}_2 + \cdots + [\lambda_m]\mathbf{a}_m = \mathbf{0}$$

contradicting the fact that the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ are linearly independent (i.e., the only solution to the above equation is the trivial solution of $[\lambda_1] = [\lambda_2] = \cdots = [\lambda_m] = [0]$). Thus $[\lambda_{m+1}]$ is nonzero, so it has a multiplicative inverse and thus we can write

$$[\lambda_{m+1}]\mathbf{a}_{m+1} = -[\lambda_1]\mathbf{a}_1 - [\lambda_2]\mathbf{a}_2 - \cdots - [\lambda_m]\mathbf{a}_m$$

$$\Downarrow$$

$$\mathbf{a}_{m+1} = -[\lambda_{m+1}]^{-1}[\lambda_1]\mathbf{a}_1 - [\lambda_{m+1}]^{-1}[\lambda_2]\mathbf{a}_2 - \cdots - [\lambda_{m+1}]^{-1}[\lambda_m]\mathbf{a_m}$$

and so $\mathbf{a}_{m+1} \in \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$. Also note that every (nonempty) subset of a linearly independent set is linearly independent and every superset of a (nonempty) linearly dependent set is linearly dependent. Since both of these statements are logically equivalent, we prove the latter: suppose that the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly dependent set of vectors in $\mathbb{F}_p^n$, and consider any superset $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_m\}$. Since the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ are linearly dependent, then there are some coefficients $[\lambda_1], [\lambda_2], \ldots, [\lambda_k] \in \mathbb{F}_p$ that are not all equal to zero such that

$$[\lambda_1]\mathbf{v}_1 + [\lambda_2]\mathbf{v}_2 + \cdots + [\lambda_k]\mathbf{v}_k = \mathbf{0}$$

Then we have

$$[\lambda_1]\mathbf{v}_1 + [\lambda_2]\mathbf{v}_2 + \cdots + [\lambda_k]\mathbf{v}_k + [0]\mathbf{v}_{k+1} + [0]\mathbf{v}_{k+2} + \cdots + [0]\mathbf{v}_m = \mathbf{0}$$

where not all of the coefficients $[\lambda_1], [\lambda_2], \ldots, [\lambda_k], [0], [0], \ldots, [0]$ are zero, which would be impossible if the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_m\}$ was linearly independent. What all this demonstrates is that every set n linearly independent vectors can be chosen one at a time so that all of the vectors chosen thus far are linearly independent at each step (which is the case if and only if the vector we add at each step is not contained within the span of all of the previous vectors), and in fact it is necessary that the vectors previously chosen be linearly independent at each step in order for the resulting set of n vectors to be linearly independent. The number of possible ways of choosing the first vector $\mathbf{v}_1 \in \mathbb{F}_p^n$ is given by

$$|\,\mathbb{F}_p^n \setminus \{\mathbf{0}\}| = |\,\mathbb{F}_p^n\,| - |\{0\}|$$

(since the zero vector is necessarily linearly independent by itself, because, for instance, $[1]\mathbf{0} = \mathbf{0}$ yet $[1] \neq [0]$), the number of possible ways of choosing the second vector $\mathbf{v}_2 \in \mathbb{F}_p^n$ is given by

$$|\,\mathbb{F}_p^n \setminus \operatorname{span}\{\mathbf{v}_1\}| = |\,\mathbb{F}_p^n\,| - |\operatorname{span}\{\mathbf{v}_1\}|$$

and so forth, with the number of way of choosing the mth vector (for $m \leq n$) is given by

$$|\,\mathbb{F}_p^n \setminus \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m-1}\}| = |\,\mathbb{F}_p^n\,| - |\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m-1}\}|$$

We claim that given any set of linearly independent vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m-1}\}$, the function

$$f : \mathbb{F}_p^{m-1} \to \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}\}$$

given by

$$f([\lambda_1], [\lambda_2], \ldots, [\lambda_{m-1}]) = [\lambda_1]\mathbf{v}_2 + [\lambda_2]\mathbf{v}_2 + \ldots + [\lambda_{m-1}]\mathbf{v}_{m-1}$$

is a bijection of sets. Note that f is manifestly surjective, because by the definition of the span we have

$$\begin{aligned}\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m-1}\} &= \{[\lambda_1]\mathbf{v}_1 + [\lambda_2]\mathbf{v}_2 + \cdots + [\lambda_{m-1}]\mathbf{v}_{m-1} : [\lambda_1], [\lambda_2], \ldots, [\lambda_{m-1}] \in \mathbb{F}_p\} \\ &= \{[\lambda_1]\mathbf{v}_1 + \cdots + [\lambda_{m-1}]\mathbf{v}_{m-1} : \big([\lambda_1], \ldots, [\lambda_{m-1}]\big) \in \mathbb{F}_p^{m-1}\} = f(\mathbb{F}_p^{m-1})\end{aligned}$$

Its injectivity follows from the linear independence of the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}\}$: if we have

$$f([\lambda_1], \ldots, [\lambda_{m-1}]) = f([\mu_1], \ldots, [\mu_{m-1}])$$

for some $([\lambda_1], \ldots, [\lambda_{m-1}]), ([\mu_1], \ldots, [\mu_{m-1}]) \in \mathbb{F}_p^{m-1}$, then we have

$$\begin{gathered}[\lambda_1]\mathbf{v}_1 + [\lambda_2]\mathbf{v}_2 + \ldots + [\lambda_{m-1}]\mathbf{v}_{m-1} = [\mu_1]\mathbf{v}_1 + [\mu_2]\mathbf{v}_2 + \ldots + [\mu_{m-1}]\mathbf{v}_{m-1} \\ \Downarrow \\ ([\lambda_1] - [\mu_1])\mathbf{v}_1 + ([\lambda_2] - [\mu_2])\mathbf{v}_2 + \ldots + ([\lambda_{m-1}] - [\mu_{m-1}])\mathbf{v}_{m-1} = \mathbf{0} \\ \Downarrow \\ [\lambda_1] - [\mu_1] = [\lambda_2] - [\mu_2] = \cdots = [\lambda_{m-1}] - [\mu_{m-1}] = [0] \\ \Downarrow \\ ([\lambda_1], [\lambda_2], \ldots, [\lambda_{m-1}]) = ([\mu_1], [\mu_2], \ldots, [\mu_{m-1}])\end{gathered}$$

Therefore

$$|\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m-1}\}| = |\,\mathbb{F}_p^{m-1}\,|$$

and so

$$|\,\mathbb{F}_p^n\,| - |\{\mathbf{0}\}| = |\,\mathbb{F}_p\,|^n - |\{\mathbf{0}\}| = p^n - 1$$

and

$$|\,\mathbb{F}_p^n\,| - |\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}\} = |\,\mathbb{F}_p^n\,| - |\,\mathbb{F}_p^{m-1}\,| = |\,\mathbb{F}_p\,|^n - |\,\mathbb{F}_p\,|^{m-1} = p^n - p^{m-1}$$

Thus the total number of ways of choosing the n columns of a linearly independent matrix is given by

$$(p^n - 1)(p^n - p)(p^n - p^2)\cdots(p^n - p^{n-2})(p^n - p^{n-1}) = \prod_{i=0}^{n-1}(p^n - p^i)$$

which is the same as the cardinality of $|\,\mathrm{GL}_n(\mathbb{F}_p)|$.

■

(2). Show that the subgroup of strictly upper triangular matrices, $U_n(\mathbb{F}_p)$, is a Sylow p-subgroup in $\mathrm{GL}_n(\mathbb{F}_p)$. Is it unique?

Proof. Every strictly upper triangular matrix has the form

$$\begin{bmatrix} [1] & [a_{12}] & [a_{13}] & [a_{14}] & \cdots & [a_{1(n-1)}] & [a_{1n}] \\ [0] & [1] & [a_{23}] & [a_{24}] & \cdots & [a_{2(n-1)}] & [a_{2n}] \\ [0] & [0] & [1] & [a_{34}] & \cdots & [a_{3(n-1)}] & [a_{3n}] \\ [0] & [0] & [0] & [1] & \cdots & [a_{4(n-1)}] & [a_{4n}] \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ [0] & [0] & [0] & [0] & \cdots & [1] & [a_{(n-1)n}] \\ [0] & [0] & [0] & [0] & \cdots & [0] & [1] \end{bmatrix}$$

for some arbitrary entries a_{ij} for $1 \leq i < j \leq n$. There are $(n^2 - n)/2 = n(n-1)/2$ pairs of natural numbers (i, j) satisfying $1 \leq i < j \leq n$ (because the number of entries strictly above the diagonal of an $n \times n$ matrix is half of the number of entries that off of the main diagonal, and there are $n \times n = n^2$ total entries and n entries on the diagonal), and each such entry could take any value in the field $\mathbb{F}_p$, so the number of matrices of this form is given by

$$|\,\mathbb{F}_p\,|^{\frac{n(n-1)}{2}} = p^{\frac{n(n-1)}{2}}$$

Note that every matrix of this form is necessarily invertible, because its columns are linearly independent, because each column is not contained in the span of the preceding columns. Each of the column vectors $\mathbf{a}_1, \mathbf{a_2}, \ldots, \mathbf{a}_{m-1}$ only have nonzero entries in the places 1 through $m-1$, so the same goes for any vector in their span (that is, any linear combination of them). The vector $\mathbf{a}_m$, however, has a nonzero element in the mth entry, so it cannot be contained in the span of the previous vectors. Thus the order of set of strictly upper triangular $U_n(\mathbb{F}_p)$ is in fact a subset of $\mathrm{GL}_n(\mathbb{F}_p)$ with cardinality

$$|U_n(\mathbb{F}_p)| = p^{\frac{n(n-1)}{2}}$$

As for demonstrating that this is the highest power of p

$$|\,\mathrm{GL}_n(\mathbb{F}_p)| = \prod_{i=0}^{n-1}(p^n - p^i)$$

note that the highest power of p dividing each of the factors $(p^n - p^i) = p^i(p^{n-i} - 1)$ is exactly p^i (since $(p^{n-i} - 1)$ doesn't contain any additional factors of p for $0 \leq i \leq n-1$, because p divides p^{n-i} and so $(p^{n-i} - 1) = -1 \neq 0 \mod p$), so the number of factors of p in $|\,\mathrm{GL}_n(\mathbb{F}_p)|$ is

$$\sum_{i=0}^{n-1} i = \frac{n(n-1)}{2}$$

Thus in order to conclude that $U_n(\mathbb{F}_p)$ is a Sylow p-subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$, all that remains to prove is that it is in fact a subgroup. We already know that it is a nonempty subset of the finite group $\mathrm{GL}_n(\mathbb{F}_p)$, so by the lemma below problem 3.2 it suffices to prove that $U_n(\mathbb{F}_p)$ is closed under multiplication. To that end, let A and B be two matrices in $U_n(\mathbb{F}_p)$ with entries $[a_{ij}]$ and $[b_{ij}]$, respectively, satifying

$$[a_{ii}] = [b_{ii}] = [1] \text{ for all } 1 \leq i \leq n$$

$$[a_{ij}] = [b_{ij}] = [0] \text{ for all } 1 \leq j < i \leq n$$

Then for all $1 \leq i \leq n$ we have

$$\begin{aligned}(AB)_{ii} = \sum_{k=1}^{n} A_{ik}B_{ki} = \sum_{k=1}^{n}[a_{ik}][b_{ki}] &= \Big(\sum_{k=1}^{i-1}[a_{ik}][b_{ki}]\Big) + [a_{ii}][b_{ii}] + \Big(\sum_{k=i+1}^{n}[a_{ik}][b_{ki}]\Big) \\ &= \Big(\underbrace{\sum_{k=1}^{i-1}[0][b_{ki}]}_{\substack{\text{because } [a_{ik}]=0 \\ \text{for } i>k}}\Big) + [1][1] + \Big(\underbrace{\sum_{k=i+1}^{n}[a_{ik}][0]}_{\substack{\text{because } [b_{ki}]=0 \\ \text{for } k>i}}\Big) = [1]\end{aligned}$$

Similarly, for all $1 \leq j < i \leq n$ we have

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{n}[a_{ik}][b_{kj}] = \big(\sum_{k=1}^{j}[a_{ik}][b_{ki}]\big) + \big(\sum_{k=j+1}^{n}[a_{ik}][b_{ki}]\big)$$

$$= \big(\underbrace{\sum_{k=1}^{j}[0][b_{ki}]}_{\substack{\text{because } [a_{ik}]=0 \\ \text{for } i>j\geq k}}\big) + \big(\underbrace{\sum_{k=j+1}^{n}[a_{ik}][0]}_{\substack{\text{because } [b_{kj}]=0 \\ \text{for } k>j}}\big) = [0]$$

Therefore AB is also strictly upper triangular, so $U_n(\mathbb{F}_p)$ is closed under multiplication and hence is a subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$. Since the order of a Sylow p-subgroup of $|\mathrm{GL}_n(\mathbb{F}_p)|$ is $p^{\frac{n(n-1)}{2}}$, we conclude that $U_n(\mathbb{F}_p)$ is a Sylow p-subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$.

■

(3). Let A, B be two commuting $n \times n$ matrices over $\mathbb{F}_p$ of order p (that is, $A^p = B^p = \mathrm{Id}_n$ and $[A, B] = 0$). Show that there exists a matrix $S \in \mathrm{GL}_n(\mathbb{F}_p)$ such that both SAS^{-1} and SBS^{-1} are upper-triangular (in other words, A, B can be simultaneously conjugated into upper-triangular form).

Proof. Because A and B both have order p, then the cyclic subgroups $\langle A\rangle$ and $\langle B\rangle$ both have order p, and hence the order of the subgroup

$$\langle A\rangle \cap \langle B\rangle$$

must divide p because it is a subgroup of both $\langle A\rangle$ and $\langle B\rangle$. Since p is prime, the only natural numbers dividing p are 1 and p. In the case that $|\langle A\rangle \cap \langle B\rangle| = p$ we must have that

$$\langle A\rangle = (\langle A\rangle \cap \langle B\rangle) = \langle B\rangle$$

(since $\langle A\rangle \cap \langle B\rangle$ is a subset of the finite set $\langle A\rangle$ with the same cardinality, and similarly it is a subset of the finite set $\langle B\rangle$ with the same cardinality). In particular, $B \in \langle B\rangle = \langle A\rangle$, so we must have

$$\langle A, B\rangle = \langle A\rangle$$

(that $\langle A, B\rangle \supset \langle A\rangle$ follows because $\langle A\rangle$ is a subgroup of any subgroup containing A, and that $\langle A, B\rangle \subset \langle A\rangle$ follows because $\langle A, B\rangle$ is a subgroup of any subgroup containing both A and B). Therefore

$$|\langle A, B\rangle| = |\langle A\rangle| = p$$

if $|\langle A\rangle \cap \langle B\rangle| = p$, and $\langle A, B\rangle$ is a p-group.

Now consider the case of $|\langle A\rangle \cap \langle B\rangle| = 1$ (that is, $\langle A\rangle \cap \langle B\rangle = \{I\}$). Since A and B commute, then $\langle A, B\rangle$ is characterized by

$$\langle A, B\rangle = \{A^i B^j : i, j \in \mathbb{N}\}$$

(we can move all of the factors of A in any finite product of the matrices A and B to the front without loss of generality). Then we claim that the map $f : (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z}) \to \langle A, B\rangle$ given by

$$f : ([i], [j]) \mapsto A^i B^j$$

is an isomorphism. First, we verify that this map is well defined. Suppose $[i_1] = [i_2]$ and $[j_1] = [j_2]$ for some coset representatives i_1, i_2 and j_1, j_2. Then

$$i_1 + p\mathbb{Z} = i_2 + p\mathbb{Z} \implies -i_1 + i_2 + p\mathbb{Z} = p\mathbb{Z} \implies -i_1 + i_2 \in p\mathbb{Z} \implies -i_1 + i_2 = np \text{ for some } n \in \mathbb{N}$$

$$j_1 + p\mathbb{Z} = j_2 + p\mathbb{Z} \implies -j_1 + j_2 + p\mathbb{Z} = p\mathbb{Z} \implies -j_1 + j_2 \in p\mathbb{Z} \implies -j_1 + j_2 = mp \text{ for some } m \in \mathbb{N}$$

so

$$f([i_2], [j_2]) = A^{i_2}B^{j_2} = A^{i_1+(-i_1+i_2)}B^{j_1+(-j_1+j_2)} = A^{i_1+np}B^{j_1+mp}$$
$$= A^{i_1}A^{np}B^{j_1}B^{mp} = A^{i_1}(A^p)^n B^{j_1}(B^p)^m = A^{i_1}I^n B^{j_1}I^m = A^{i_1}B^{j_1} = f([i_1], [j_1])$$

We also verify that f is indeed a homomorphism: for any elements $([i_1],[j_1]),([i_2],[j_2]) \in (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})$, because A and B commute we have

$$\begin{aligned} f\big(([i_1],[j_1]) + ([i_2],[j_2])\big) &= f([i_1+i_2],[j_1+j_2]) = A^{i_1+i_2}B^{j_1+j_2} \\ &= A^{i_1}A^{i_2}B^{j_1}B^{j_2} = A^{i_1}B^{j_1}A^{i_2}B^{j_2} = f([i_1],[j_1])f([i_2],[j_2]) \end{aligned}$$

Next, we note that f is manifestly surjective, because for any element $A^iB^j \in \langle A,B\rangle$ we have $f([i],[j]) = A^iB^j$. Finally, we verify injectivity by demonstrating that the kernel of f is trivial. Suppose that

$$f([i],[j]) = A^iB^j = I$$

Then

$$A^i = IB^{-j} = B^{-j}$$

is contained within both $\langle A\rangle$ and $\langle B\rangle$, so because $\langle A\rangle \cap \langle B\rangle = \{I\}$ we have

$$A^i = B^{-j} = I$$

Because $|A| = |B| = p$, we must then have that $i = np$ for some $n \in \mathbb{N}$ and $-j = mp$ for some $m \in \mathbb{N}$, in which case $[i] = [np] = [0]$ and $[j] = [-mp] = [0]$, thus the kernel is trivial as claimed. Therefore, since an isomorphism is also a bijection of the underlying sets, we have

$$|\langle A,B\rangle| = |(\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})| = |\mathbb{Z}/p\mathbb{Z}||\mathbb{Z}/p\mathbb{Z}| = |\mathbb{Z}/p\mathbb{Z}|^2 = p^2$$

In particular, $\langle A,B\rangle$ is a p-group in this case as well.

Since $\langle A,B\rangle$ is a p-group in every case, by the second Sylow theorem we have that there is always some Sylow p-group $P < \mathrm{GL}_n(\mathbb{F}_p)$ containing $\langle A,B\rangle$. Because $U_n(\mathbb{F}_p)$ and P are both p-Sylow subgroups, they are conjugate to each other: that is, there exists some matrix $S \in \mathrm{GL}_n(\mathbb{F}_p)$ such that

$$U_n(\mathbb{F}_p) = SPS^{-1}$$

Then, since

$$A, B \in \langle A,B\rangle \subset P$$

we have that

$$SAS^{-1}, SBS^{-1} \in SPS^{-1} = U_n(\mathbb{F}_p)$$

In particular, for the matrix $S \in \mathrm{GL}_n(\mathbb{F}_p)$ we have that SAS^{-1} and SBS^{-1} are both in upper triangular form.

■

(4). Is this still true if A, B do not commute?

Proof. It is not true in general if A and B do not commute. Consider the group $\mathrm{GL}_2(\mathbb{F}_2)$ (that is, taking $n = 2$ and $p = 2$), and the matrices

$$A = \begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix} \qquad\qquad B = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}$$

Noting that for any element $n \in \mathbb{F}_2 = (\mathbb{Z}/2\mathbb{Z})$ we have $[n]+[n] = [2n] = [2][n] = [0][n] = [0]$ (that is, every element is its own additive inverse), we can readily verify the following relations

$$A^2 = \begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [0] & [1]\end{bmatrix} = I$$

$$B^2 = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}\begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [0] & [1]\end{bmatrix} = I$$

$$AB = \begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix}\begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix} = \begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix}$$

$$BA = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix} = \begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix}$$

That is, $A^p = B^p = I$ and $AB \neq BA$. Note that by problem 4.1, $\mathrm{GL}_2(\mathbb{F}_2)$ only contains $(2^2-1)(2^2-2) = (4-1)(4-2) = 3 \cdot 2 = 6$ elements, which one can confirm are those listed below:

$$I = \begin{bmatrix}[1] & [0] \\ [0] & [1]\end{bmatrix} \qquad A = \begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix}$$

$$B = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix} \qquad C = \begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix}$$

$$D = \begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix} \qquad E = \begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix}$$

That all of these elements are invertible follows because

$$A^2 = I \implies A^{-1} = A$$

$$B^2 = I \implies B^{-1} = B$$

$$C^2 = \begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix}\begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [0] & [1]\end{bmatrix} = I \implies C^{-1} = C$$

$$DE = \begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [0] & [1]\end{bmatrix} = I \quad \text{and} \quad ED = \begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix}\begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [0] & [1]\end{bmatrix} = I$$

$$\Downarrow$$

$$D^{-1} = E \quad \text{and} \quad E^{-1} = D$$

Then we can simply show that for every matrix $X \in \mathrm{GL}_2(\mathbb{F}_2)$, at least one of XAX^{-1} and XBX^{-1} is not upper diagonal. Consider conjugating B by I, A, and B and conjugating A by C, D, and E:

$$IBI^{-1} = B = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}$$

$$ABA^{-1} = (AB)A = \begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix}$$

$$BBB^{-1} = B = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}$$

$$CAC^{-1} = CAC = \begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix}\begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix}\begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix} = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}$$

$$DAD^{-1} = DAE = \begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix} = \begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix}\begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix} = \begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix}$$

$$EAE^{-1} = EAD = \begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix}\begin{bmatrix}[1] & [1] \\ [0] & [1]\end{bmatrix}\begin{bmatrix}[0] & [1] \\ [1] & [1]\end{bmatrix} = \begin{bmatrix}[1] & [1] \\ [1] & [0]\end{bmatrix}\begin{bmatrix}[1] & [0] \\ [1] & [1]\end{bmatrix} = \begin{bmatrix}[0] & [1] \\ [1] & [0]\end{bmatrix}$$

None of the resulting matrices are upper triangular, so we have provided a counterexample. ■

Problem 1.

Prove that $\binom{p^a m}{p^a}$ is not divisible by p, a prime, if $(m, p) = 1$, and a is a positive integer.

Proof. The binomial coefficient $\binom{p^a m}{p^a}$ is given by

$$\binom{p^a m}{p^a} = \frac{(p^a m)!}{(p^a)!(p^a m - p^a)!} = \frac{(p^a m)(p^a m - 1)(p^a m - 2)\cdots(p^a m - p^a + 1)}{(p^a)(p^a - 1)(p^a - 2)\cdots(p^a - p^a + 1)} = \prod_{i=0}^{p^a - 1} \frac{p^a m - i}{p^a - i}$$

If one considers the prime factorization of both the numerator and denominator of the above fraction, one sees that p divides $\binom{p^a m}{p^a}$ if and only if there are some factors of p in the prime factorization of the numerator that are left un-canceled by the prime factorization of the denominator. However, there are no excess factors of p in the numerator relative to the denominator, because for each of the terms $(p^a m - i)/(p^a - i)$ the same factor of p divides both the numerator and denominator. First, we note that p^n divides neither $p^a m - i$ nor $p^a - i$ for any $n > a$: in the case of $p^a - i$, clearly p^n cannot divide $p^a - i$ because

$$p^a - i \leq p^a < p^n$$

In the case of $p^a m - i$, in order to have that p^n divides $p^a m - i$, then for some integer $k \in \mathbb{Z}$ we must have

$$p^a m - i = k p^n = k p^{n-a} p^a$$

$$\Downarrow$$

$$p^a (m - k p^{n-a}) = i$$

That is, since $m - k p^{n-a}$ is an integer (since $n > a$), we have that p^a divides i. The product only runs over the indices $0 \leq i \leq p^a - 1$, and if $1 \leq i \leq p^a - 1$ then clearly p^a doesn't divide i, so this is only possible if $i = 0$. However, p^n doesn't divide $p^a m - i = p^a m$ when $i = 0$ because $a < n$ and m is coprime to p (it doesn't contain any factors of p in its prime factorization). Therefore p^n doesn't divide $p^a - i$ or $p^a m - i$ for any $n > a$, so now consider the case of $1 \leq n \leq a$. Then both $p^a m$ and p^a are zero mod p^n, so we have

$$p^a m - i \ \equiv\ -i \ \equiv\ p^a - i \qquad \mod p^n$$

and thus p^n divides $p^a - i$ (that is, $p^a - i \equiv 0 \mod p^n$) if and only if p^n divides $p^a m - i$ (that is, $p^a m - i \equiv 0 \mod p^n$). That is to say, exactly the same power of p divides both the $p^a - i$ and $p^a m - i$, so the factors of p in the numerator of $\binom{p^a m}{p^a}$ are in one-to-one correspondence with the factors of p in the denominator. Therefore

$$p \ \nmid\ \binom{p^a m}{p^a}$$

whenever $\gcd(n, p) = 1$. ∎

Problem 2.

Let $H < G$ be a subgroup. Show that the normalizer $N_G(H) = \{g \in G \,|\, gHg^{-1} = H\}$ is the maximal subgroup of G in which H is normal.

Proof. First, we show that $N_G(H)$ is a subgroup of G. We have that $e \in N_G(H)$ because $eHe^{-1} = H$, and because

$$gHg^{-1} = H \quad \iff \quad gH = Hg \quad \iff \quad H = g^{-1}Hg$$

then for any $g \in N_G(H)$ we have that

$$g^{-1}H(g^{-1})^{-1} = g^{-1}Hg = H$$

and thus $g^{-1} \in N_G(H)$. Finally, if $g_1, g_2 \in N_G(H)$ (that is, $g_1 H g_1^{-1} = H$ and $g_2 H g_2^{-1} = H$), then

$$(g_1 g_2) H (g_1 g_2)^{-1} = g_1 (g_2 H g_2^{-1}) g_1^{-1} = g_1 H g_1^{-1} = H$$

so $g_1 g_2 \in H$. Therefore $N_G(H)$ contains the identity and is closed under both multiplication and inverses, it follows that $N_G(H)$ is a subgroup of G. We also note that H is contained in $N_G(H)$, because for each element $h \in H$ we have

$$hH = H = Hh$$
$$\Downarrow$$
$$hHh^{-1} = H$$

and thus $h \in N_G(H)$.

Now consider any subgroup K of G such that $H \trianglelefteq K$. Then, by the definition of normality, for all $k \in K$ we have

$$kHk^{-1} = H$$

so in particular $k \in N_G(H)$. Therefore $K \subset N_G(H)$, so because $N_G(H)$ is a subgroup that H is normal in that contains every other subgroup that H is normal in, $N_G(H)$ is the maximal subgroup that in which H is normal. ■

Problem 3.

Let H be a subgroup of G of index 2. Show that H is normal.

Proof. In order to demonstrate that H is normal in G, we wish to show that for all $g \in G$ we have

$$gH = Hg$$

To do so, we note that the set of left cosets partitions the group, so since we know that $eH = H$ is one of the left cosets and there are only two left cosets in total (since the index of H in G is 2), then the other left coset must be $G \setminus H$. Similarly, the set of right cosets must consist of $He = H$ and $G \setminus H$. Now let g be any element of the group G. Since we have that

$$gH = H \quad \iff \quad g \in H$$

and similarly for right cosets

$$Hg = H \quad \iff \quad g \in H$$

it follows that if $g \in H$ we have

$$gH = H = Hg$$

and if $g \notin H$ we have

$$gH = G \setminus H = Hg$$

Thus in every case we have $gH = Hg$, so H is normal in G.

■

Problem 4.

Let G be a finite group, and $H \lhd G$ be a normal subgroup. Suppose $(|H|, [G : H]) = 1$. Show that H is the ONLY subgroup of G of order $|H|$.

Proof. Given a finite group G and a normal subgroup $H \trianglelefteq G$ such that $\gcd(|H|, [G : H]) = 1$, let $K \leq G$ be any subgroup of order $|H|$. Let $\pi : G \to G/H$ be the canonical projection $g \mapsto gH$, and consider $\pi(K)$. Since $\pi(K)$ is the image of a subgroup under a group homomorphism, it follows that $\pi(K)$ is a subgroup of G/H, so by Lagrange's theorem we have that $|\pi(K)|$ divides $|G/H| = [G : H]$. Furthermore, consider the restriction $\pi|_K : K \to \pi(K)$, which is a surjective group homomorphism from K to $\pi(K)$. Because the kernel of the canonical projection $\pi : G \to G/H$ is exactly H (since $\pi(g) = e_{G/H} \iff gH = H \iff g \in H$), then kernel of $\pi|_K$ is given by

$$\ker \pi|_K = \{k \in K : \pi|_K(k) = e_{\pi(K)}\} = \{k \in K : \pi(k) = e_{G/H}\} = \{k \in K : k \in \ker \pi = H\} = K \cap H$$

By the first isomorphism theorem, it follows that

$$\frac{K}{K \cap H} = \frac{K}{\ker \pi|_K} \cong \operatorname{im} \pi|_K = \pi(K)$$

$$\Downarrow$$

$$\Big|\frac{K}{K \cap H}\Big| = |\pi(K)|$$

Applying Lagrange's theorem again and making the substitution $|K| = |H|$, we have

$$|\pi(K)| = \frac{|K|}{|K \cap H|} = \frac{|H|}{|K \cap H|}$$

In particular, because $|H| = |K \cap H||\pi(K)|$, where $|K \cap H|$ is an integer, it follows that $|\pi(K)|$ also divides $|H|$. Thus we have that $|\pi(K)|$ divides both $|H|$ and $[G : H]$, so because $\gcd(|H| : [G : H]) = 1$ we must have $|\pi(K)| = 1$. In particular, $\pi(K)$ must be the trivial subgroup $e_{G/H} \leq G/H$ and hence $K \subset \ker \pi = H$. Since K is a subset of H with the same cardinality as H (where $|H|$ is finite, i.e. no proper subset of H has the same cardinality as H), then it follows that $K = H$. Thus given any subgroup $K \leq G$ with $|K| = |H|$, we have shown that $K = H$, so in particular H is the only subgroup of G of order $|H|$. ■

Problem 5.

(5.1). Consider the "standard" action of $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^n$ (that is, an $n \times n$ invertible matrix acts on an $n \times 1$ vector by matrix multiplication). Describe the orbits and isotropy groups (up to conjugation) of this action.

Proof. For any vector $x \in \mathbb{R}^n$, the orbit of x is given by

$$\mathrm{Orb}(x) = \{Ax : A \in GL_n(\mathbb{R})\}$$

If $x = 0$, then $Ax = 0$ for all matrices A, so the orbit of 0 contains only itself

$$\mathrm{Orb}(0) = \{0\}$$

Now, considering the case of any nonzero vector x, we claim

$$\mathrm{Orb}(x) = \{y \in \mathbb{R}^n : y \neq 0\} = \mathbb{R}^n \setminus \{0\}$$

To demonstrate this, consider any two elements $x, y \in \mathbb{R}^n \setminus \{0\}$. Since the sets $\{x\}$ and $\{y\}$ are linearly independent (the single vector that each consists of is nonzero), then both can be expanded into bases $\{x_1 = x, x_2, x_3, \ldots, x_n\}$ and $\{y_1 = y, y_2, y_3, \ldots, y_n\}$ of $\mathbb{R}^n$. Then consider the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ defined for any vector $u = \sum_{i=1}^n u_i x_i$ in $\mathbb{R}^n$ by

$$T(u) = T(\sum_{i=1}^n u_i x_i) = \sum_{i=1}^n u_i y_i$$

Note that this function is well defined because each vector in $\mathbb{R}^n$ can be represented as a unique linear combination of the basis vectors, and it is linear because for any vectors $u, v \in \mathbb{R}^n$ and any scalars $a, b \in \mathbb{R}$ we have

$$\begin{aligned} T(au + bv) &= T\big(a\sum_{i=1}^n u_i x_i + b\sum_{i=1}^n v_i x_i\big) = T\big(\sum_{i=1}^n (au_i + bv_i)x_i\big) = \sum_{i=1}^n (au_i + bv_i)y_i = a\sum_{i=1}^n u_i y_i + b\sum_{i=1}^n v_i y_i \\ &= aT(u) + bT(v) \end{aligned}$$

This linear transformation has the property that

$$T(x) = T(x_1) = y_1 = y$$

It also doesn't send any nonzero vectors to the zero vector, because the fact that each basis consists of linearly independent vectors means that

$$u = 0 \iff \sum_{i=1}^n u_i x_i = 0 \iff u_i = 0 \text{ for all } 1 \leq i \leq n \iff \sum_{i=1}^n u_i y_i = 0 \iff T(u) = 0$$

This means that if we represent T by a matrix A such that $T(u) = Au$ for all vectors $u \in \mathbb{R}^n$, for instance $A = [T(e_1) \;\; T(e_2) \;\; T(e_3) \;\; \cdots \;\; T(e_n)]$ (where e_i is the ith standard basis vector), then this matrix A is invertible because the dimension of its null space is zero. Thus we have found a matrix $A \in \mathrm{GL}_n(\mathbb{R})$ such that $Ax = y$, so $y \in \mathrm{Orb}(x)$ as claimed.

On the other hand, the stabilizer of an element $x \in \mathbb{R}^n$ is given by

$$\mathrm{Stab}(x) = \{A \in \mathrm{GL}_n(\mathbb{R}) : Ax = x\}$$

As mentioned before, for any $A \in \mathrm{GL}_n(\mathbb{R})$ we have that A applied to the zero vector always equals the zero vector, so

$$\mathrm{Stab}(0) = \mathrm{GL}_n(\mathbb{R})$$

For any two nonzero vectors $x, y \in \mathbb{R}^n \setminus \{0\}$, we note that because they are in the same orbit, their stabilizers are conjugates of each other: because there exists some $A \in \mathrm{GL}_n(\mathbb{R})$ such that $Ax = y \iff A^{-1}y = x$, then we can define inverse maps $f : \mathrm{Stab}(x) \to \mathrm{Stab}(y)$ and $f^{-1} : \mathrm{Stab}(y) \to \mathrm{Stab}(x)$ by $f : B \mapsto ABA^{-1}$ and $f : C \mapsto A^{-1}CA$. One can verify that if $B \in \mathrm{Stab}(x)$, then $f(B) \in \mathrm{Stab}(y)$ because

$$f(B)y = ABA^{-1}y = ABx = Ax = y$$

and similarly if $C \in \mathrm{Stab}(y)$, then $f^{-1}(C) \in \mathrm{Stab}(x)$ because

$$f^{-1}(C)x = A^{-1}CAx = A^{-1}Cy = Ay = x$$

One can also verify that the maps f and f^{-1} are indeed inverses, because

$$f^{-1}(f(B)) = f^{-1}(ABA^{-1}) = A^{-1}ABA^{-1}A = B$$

$$f(f^{-1}(C)) = f(A^{-1}CA) = AA^{-1}CAA^{-1} = C$$

Even besides being a bijection of sets, the maps f and f^{-1} are also isomorphisms of groups, because

$$f(B_1)f(B_2) = (AB_1A^{-1})(AB_2A^{-1}) = AB_1B_2A^{-1} = f(B_1B_2)$$

$$f^{-1}(C_1)f^{-1}(C_2) = (A^{-1}C_1A)(A^{-1}C_2A) = A^{-1}C_1C_2A = f^{-1}(C_1C_2)$$

Therefore we can specify the stabilizers up to conjugation for the elements of the orbit $\mathbb{R}^n \setminus \{0\}$ just by specifying the orbit for one particular representative of the orbit. Then consider the element $e_1 = [1, 0, 0, 0, \cdots, 0]^T$: the orbit of that particular element is the set of all matrices $(A_{ij})_{1 \le i,j \le n}$ such that the first column of A is e_1 and the $(n-1) \times (n-1)$ sub-matrix $(A_{ij})_{2 \le i,j \le n}$ has nonzero determinant. This follows because A is in the stabilizer of e_1 if and only if $A \in \mathrm{GL}_n(\mathbb{R})$ and $Ae_1 = e_1$. Since Ae_1 is the first column of A, then $Ae_1 = e_1$ if and only if the first column of A is e_1, in which case, doing the cofactor expansion down the first column of A, we find that

$$\det(A_{ij})_{1 \le i,j \le n} = \det(A_{ij})_{2 \le i,j \le n}$$

Therefore A is invertible if and only if $\det(A) \neq 0$, which is true if and only if the determinant of the sub-matrix $(A_{ij})_{2 \le i,j \le n}$ is nonzero. ■

(5.2P). Let $\mathrm{GL}_n(\mathbb{R})$ act on the set of all $n \times n$ complex matrices by conjugation. Determine the orbits.

Proof. Every matrix is related to its Jordan canonical form via a similarity transformation, and the Jordan canonical form of a matrix is uniquely determined (up to permutation of the eigenvalues) by the eigenvalues of that matrix (with multiplicity) and the geometric multiplicity. If two matrices have the same Jordan canonical form J, then they can be related via a similarity transformation because being similar is an equivalence relation. If their Jordan canonical forms are distinct, then the Jordan canonical forms cannot be related via a similarity transformation, so neither can the original matrices be related via a similarity transformation. We can test whether two matrices A and B have the same Jordan canonical form by testing their eigenvalues and multiplicities using $\det(A - \lambda I) = \det(B - \lambda I)$ and testing that they have the same geometric multiplicities with $\dim \mathrm{null}(A - \lambda I) = \dim \mathrm{null}(B - \lambda I)$, thus we can write the orbit of A as

$$\mathrm{Orb}(A) = \{B \in \mathrm{GL}_n(\mathbb{R}) : \det(B - \lambda I) = \det(A - \lambda I) \text{ and } \dim \mathrm{null}(B - \lambda I) = \dim \mathrm{null}(A - \lambda I)\}$$

■

Problem 6P.

Show that any group of order p^2 for p prime is abelian.

Proof. Let G be our given group of order p^2, and consider the class formula (from G acting on itself by conjugation)

$$|G| = |Z| + \sum_{x \in S} [G : \mathrm{Stab}_G(x)]$$

where S is some set containing exactly one representative for each nontrivial orbit and Z is the center

$$Z = \{z \in G : gzg^{-1} = z \text{ for all } g \in G\} = \{z \in G : gz = zg \text{ for all } g \in G\}$$

Since each stabilizer $\mathrm{Stab}_G(x)$ is a subgroup of G, by Lagrange's theorem we have

$$|\mathrm{Stab}_G(x)| \,\Big|\, |G| = p^2$$

so in particular we must have that the order of $\mathrm{Stab}_G(x)$ is either 1, p, or p^2. Since S only contains representatives for nontrivial orbits, then for each $x \in S$ we have that $\mathrm{Stab}(x) \neq G$, so the possible orders of $\mathrm{Stab}(x)$ are 1 and p and the possible values of the indices are

$$[G : \mathrm{Stab}(x)] = \frac{|G|}{|\mathrm{Stab}(x)|} = \frac{p^2}{1} = p^2$$

$$[G : \mathrm{Stab}(x)] = \frac{|G|}{|\mathrm{Stab}(x)|} = \frac{p^2}{p} = p$$

In either case we have

$$p \,\Big|\, [G : \mathrm{Stab}_G(x)]$$

Since p also divides $|G| = p^2$, we have

$$|Z| = |G| - \sum_{x \in S} [G : \mathrm{Stab}_G(x)]$$

$$\Downarrow$$

$$p \,\Big|\, |Z|$$

(since p divides everything on the right hand side). Because Z is also a subgroup of G, then the order of Z has to divide $|G| = p^2$ as well, so we must have either $|Z| = p$ or $|Z| = p^2$. If $|Z| = p^2$, then Z is all of G and thus for all $z \in G = Z$ we have

$$gz = zg \text{ for all } g \in G$$

In particular, G is abelian, so we are done. Suppose instead that $|Z| = p$. Then because Z is normal (for any $z \in Z$ and $g \in G$ we have that $gz = zg \implies gzg^{-1} = z$, so for any $g \in G$ we have $gZg^{-1} = Z$), we can form the quotient group G/Z of order

$$|G/Z| = \frac{|G|}{|Z|} = \frac{p^2}{p} = p$$

Since every group of prime order is cyclic, we have that there is some element $a \in G$ such that

$$G/Z = \langle aZ \rangle = \{a^n Z : n \in \mathbb{N}\}$$

Thus for every $g_1, g_2 \in G$ we have that $g_1 \in a^n Z$ and $g_2 \in a^m Z$ for some $n, m \in \mathbb{N}$ (since the collection of cosets partitions the space), in which case $g_1 = a^n z_1$ and $g_2 = a^m z_2$ for some $z_1, z_2 \in Z$. Then because z_1 and z_2 commute with every element of G by the definition of the center, we have

$$g_1 g_2 = a^n z_1 a^m z_2 = a^n a^m z_1 z_2 = a^m a^n z_2 z_1 = a^m z_2 a^n z_1 = g_2 g_1$$

Since this holds for all $g_1, g_2 \in G$, we have that G is abelian. ∎

Classification of Groups of Order pq

October 30, 2020

Main Idea

We will use these two theorems.

Theorem (Direct Products)

Let $H, K \leq G$. Then $G \cong H \times K$ if and only if

- $H, K \lhd G$
- $H \cap K = \{e\}$
- $HK = G$.

Theorem (Semi-direct Products)

Let $H, K \leq G$ Then $G \cong H \ltimes K$ if and only if

- $K \lhd G$
- $K \cap H = \{e\}$
- $KH = G$.

Sylow Theorems

$|G| = pq \implies$ (by Cauchy's theorem) that there exist $H, K \leq G$ such that $|H| = p, |K| = q$.

Suppose WLOG $p < q$. Then by Sylow's third theorem,

- $n_q = kq + 1$ $(k \in \mathbb{N})$ and $n_q | p \implies k = 0 \implies n_q = 1$, so K is normal in G.

What about H?

Similarly for H,

$$n_p = mp + 1 \ (k \in \mathbb{N}) \text{ and } n_p | q. \tag{1}$$

So $n_p = 1$ or $n_p = q$.

If $n_p \neq 1$, then

$$mp + 1 = n_p = q \implies p|(q-1). \tag{2}$$

In the case that $n_p = 1$

We show that the conditions of the Direct Product theorem are satisfied.

- $H \cap K = \{e\}$: we have that $H \cap K \leq H$ and $H \cap K \leq K$, hence by Lagrange's theorem

$$|H \cap K|\,|H| = p \quad \text{and} \quad |H \cap K|\,|K| = q. \tag{3}$$

  As p, q are prime, $|H \cap K| = 1$.

- As K, H are normal in G, we have that $HK \leq G$, and as

$$|HK| = \frac{|H||K|}{|H \cap K|} = |H||K| = pq, \tag{4}$$

  we have that $HK = G$.

Therefore in this case

$$G \cong H \times K. \tag{5}$$

In the case that $n_p = q$

We already showed that $n_p = q \implies p|(q-1)$. The same arguments

used in the previous slide show that as K is normal,

$$G \cong H \ltimes K. \tag{6}$$

Recall that a semi-direct product is defined w.r.t to some homomorphism

$$\phi : H \to \mathrm{Aut}(K). \tag{7}$$

To finish the classification we will show that there is a unique choice for the map ϕ.

What are the possible homomorphisms?

From Problem 2.1, $Aut(K)$ is cyclic of order $q-1$, and by Lagrange's theorem,

$$p|(q-1) \implies \exists! P \leq Aut(K) \text{ with } |P| = p. \tag{8}$$

As H and P have prime order, they are cyclic. Let $H = \langle h \rangle$ and $P = \langle z \rangle$, then

$$\phi : H \to Aut(K) \implies \phi(H) = P. \tag{9}$$

To see why, we must have that $\phi(x)$ generates $Im(\phi)$ (by properties of hom.). So $Im(\phi)$ is cyclic of order p, hence equal to P.

So there are p possible maps, given by

$$\phi_i(h) = z^i, \ 0 \leq i \leq p-1. \tag{10}$$

If $i = 0$, then the map is trivial, and we end up with the direct product, as before.

Showing that each homomorphism yields an isomorphic group

We show that the choice of map ϕ_i for $0 < i \leq p-1$ only changes the generator of P.
Fix i, then there is some $x \in H$ such that

$$\phi_i(x) = z. \tag{11}$$

We show that $H = \langle x \rangle$.
Let h_0 be arbitrary in H, then $h_0 = h^k$ for some k. Then

$$\phi_i(h_0) = \phi(h^k) = \phi(h)^k = z^{ik} = \phi_i(x)^{ik} = \phi(x^{ik}), \tag{12}$$

so $h_0 = x^{ik}$ (this map is injective because it is a map between cyclic groups).
Thus $H = \langle x \rangle$.

Sketch of Isomorphism Proof

$$\textit{Cyclic groups} \implies \textit{Map is determined by } \phi_i(h)$$

$$K \rtimes_{\phi_1} H \qquad\qquad K \rtimes_{\phi_2} H$$

$$(k,h)(k_1,h_1) \qquad\qquad (k,h)(k_1,h_1)$$

$$\left(k\,{}^{h}k_1,\ hh_1\right) = (k(\phi_1(h)(k)),\ hh_1) \qquad \left(k\,{}^{h}k_1,\ hh_1\right) = (k(\phi_2(h)(k)),\ hh_1)$$

$$\phi_1,\ \phi_2 \textit{ are bijections from } H \to P \textit{ so } \exists\, \phi_2^{-1} : P \to H;$$
$$\textit{Take } \varphi : H \to H : \ \varphi(h) = \phi_2^{-1}(\phi_1(h)). \textit{ Then :}$$

$$f : K \rtimes_{\phi_1} H \to K \rtimes_{\phi_2} H \quad \textit{given by } \ f((k,h)) = (k,\varphi(h)) \textit{ is an iso.}$$

Conclusion

So we conclude that if $p \not| (q-1)$, then

$$G \cong H \times K \cong \mathbb{Z}_p \times \mathbb{Z}_q. \tag{13}$$

If $p|(q-1)$ then there are two cases:

- If we choose ϕ_1 on the previous slides, we get the trivial map, and

$$G \cong H \times K \cong \mathbb{Z}_p \times \mathbb{Z}_q, \tag{14}$$

  as above.
- Otherwise, if we choose ϕ_i, $0 < i < p$, then we conclude that

$$G \cong H \ltimes K \cong \mathbb{Z}_p \ltimes \mathbb{Z}_q. \tag{15}$$

Problem Problem 1.

Let G be a group. Prove that the following are equivalent:

(1) There exists a (finite) central series $\{e\} = G_0 < G_1 < \ldots < G_n = G$.

(2) The descending central series

$$\ldots < \Gamma_i = [\Gamma_{i-1}, G] < \Gamma_{i-1} < \ldots < \Gamma_1 < \Gamma_0 = G$$

terminates at $\Gamma_n = \{e\}$.

(3) The ascending central series $\{e\} = Z_0 < Z_1 < Z_2 < \ldots$ (where $Z_i/Z_{i-1} = Z(G/Z_{i-1})$) terminates at $Z_n = G$.

- Problem 1.1P: Prove the equivalence of (1) and (2).

  Proof. Suppose that there exists a finite central series

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

  By the definition of a central series, we have that

  $$[G, G_i] < G_{i+1}$$

  for all i. Then consider the descending central series

  $$G = \Gamma_0 > \Gamma_1 > \Gamma_2 > \cdots$$

  which is defined by $\Gamma_{i+1} = [G, \Gamma_i]$ (note that this series might not terminate a priori). Then we claim that for all i, $\Gamma_i < G_i$, which we prove by induction. In the case of $i = 0$, it follows immediately that $\Gamma_0 < G_0$ because $\Gamma_0 = G_0 = G$. Now suppose that we have that $\Gamma_i < G_i$ for some particular i: then

  $$\Gamma_{i+1} = [G, \Gamma_i] < [G, G_i] < G_{i+1}$$

  Thus the result follows for all i, and in particular $\Gamma_n < G_n = \{e\}$, so the descending central series terminates. ■

- Problem 1.2P: Prove the equivalence of (1) and (3).

Problem Problem 2P.

Let G be a finite group and $P < G$ be its Sylow subgroup. Show that $N_G(N_G(P)) = N_G(P)$.

Proof. Note that because a subgroup is always contained within its own normalizer, we have that

$$P \leq N_G(P) \leq N_G(N_G(P)) \leq G$$

Suppose that $|P| = k$, so that $p^k \mid |G|$ but $p^{k+1} \nmid |G|$. Then because P is a subgroup of $N_G(P)$, it follows that $|P| = p^k$ divides $|N_G(P)|$, and because $N_G(P)$ is a subgroup of G, it follows that p^{k+1} does not divide $|N_G(P)|$ (if it did, then because $|N_G(P)| \mid |G|$ we would have that p^{k+1} divides $|G|$). Therefore P is a subgroup of $N_G(P)$ whose order is the maximum power of P dividing $|G|$, so P is also a Sylow p-subgroup of $N_G(P)$, and because P is normal in $N_G(P)$ it follows that P is the only Sylow p-subgroup of $N_G(P)$ (all Sylow p-subgroups of $N_G(P)$ must be conjugate to P by some element in $N_G(P)$, but P is fixed under conjugation by every element of $N_G(P)$). Now consider any element $x \in N_G(N_G(P))$: because $N_G(P)$ is normal in $N_G(N_G(P))$, we have that $xN_G(P)x^{-1} = N_G(P)$, so

$$xPx^{-1} \leq xN_G(P)x^{-1} = N_G(P)$$

Since xPx^{-1} is a Sylow p-subgroup of $N_G(P)$, it follows that $xPx^{-1} = P$ (since there is only one Sylow p-subgroup of $N_G(P)$), so $x \in N_G(P)$. Since this holds for any $x \in N_G(N_G(P))$, it follows that $N_G(N_G(P)) \subset N_G(P)$ and hence

$$N_G(N_G(P)) = N_G(P)$$

■