

THE SYLOW THEOREMS

KEITH CONRAD

1. INTRODUCTION

The converse of Lagrange's theorem is false: if G is a finite group and $d \mid |G|$, then there may not be a subgroup of G with order d . The simplest example of this is the group A_4 , of order 12, which has no subgroup of order 6. The Norwegian mathematician Peter Ludwig Sylow [1] discovered that a converse result *is* true when d is a prime power: if p is a prime number and $p^k \mid |G|$ then G must contain a subgroup of order p^k . Sylow also discovered important relations among the subgroups with order the *largest* power of p dividing $|G|$, such as the fact that all subgroups of that order are conjugate to each other.

For example, a group of order $100 = 2^2 \cdot 5^2$ must contain subgroups of order 1, 2, 4, 5, and 25, the subgroups of order 4 are conjugate to each other, and the subgroups of order 25 are conjugate to each other. It is not necessarily the case that the subgroups of order 2 are conjugate or that the subgroups of order 5 are conjugate.

Definition 1.1. Let G be a finite group and p be a prime. A subgroup of G whose order is the highest power of p dividing $|G|$ is called a *p-Sylow subgroup* of G . A *p-Sylow subgroup* for some p is called a *Sylow subgroup*.

In a group of order 100, a 2-Sylow subgroup has order 4, a 5-Sylow subgroup has order 25, and a p -Sylow subgroup is trivial if $p \neq 2$ or 5.

In a group of order 12, a 2-Sylow subgroup has order 4, a 3-Sylow subgroup has order 3, and a p -Sylow subgroup is trivial if $p > 3$. Let's look at a few examples of Sylow subgroups in groups of order 12.

Example 1.2. In $\mathbf{Z}/(12)$, the only 2-Sylow subgroup is $\{0, 3, 6, 9\} = \langle 3 \rangle$ and the only 3-Sylow subgroup is $\{0, 4, 8\} = \langle 4 \rangle$.

Example 1.3. In A_4 there is one subgroup of order 4, so the only 2-Sylow subgroup is

$$\{(1), (12)(34), (13)(24), (14)(23)\} = \langle (12)(34), (14)(23) \rangle.$$

There are four 3-Sylow subgroups:

$$\begin{aligned} \{(1), (123), (132)\} &= \langle (123) \rangle, & \{(1), (124), (142)\} &= \langle (124) \rangle, \\ \{(1), (134), (143)\} &= \langle (134) \rangle, & \{(1), (234), (243)\} &= \langle (234) \rangle. \end{aligned}$$

Example 1.4. In D_6 there are three 2-Sylow subgroups:

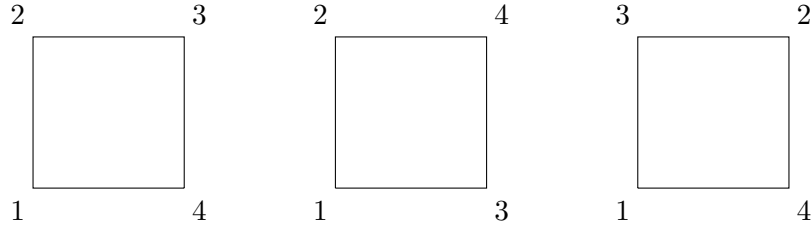
$$\{1, r^3, s, r^3s\} = \langle r^3, s \rangle, \quad \{1, r^3, rs, r^4s\} = \langle r^3, rs \rangle, \quad \{1, r^3, r^2s, r^5s\} = \langle r^3, r^2s \rangle.$$

The only 3-Sylow subgroup of D_6 is $\{1, r^2, r^4\} = \langle r^2 \rangle$.

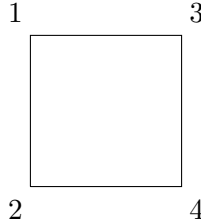
In a group of order 24, a 2-Sylow subgroup has order 8 and a 3-Sylow subgroup has order 3. Let's look at two examples.

Example 1.5. In S_4 , the 3-Sylow subgroups are the 3-Sylow subgroups of A_4 (an element of 3-power order in S_4 must be a 3-cycle, and they all lie in A_4). We determined the 3-Sylow subgroups of A_4 in Example 1.3; there are four of them.

There are three 2-Sylow subgroups of S_4 , and they are interesting to work out since they can be understood as *copies of D_4 inside S_4* . The number of ways to label the four vertices of a square as 1, 2, 3, and 4 is $4! = 24$, but up to rotations and reflections of the square there are really just three different ways of carrying out the labeling, as follows.



Every other labeling of the square is a rotated or reflected version of one of these three squares. For example, the square below is obtained from the middle square above by reflecting across a horizontal line through the middle of the square.



When D_4 acts on a square with labeled vertices, each motion of D_4 creates a permutation of the four vertices, and this permutation is an element of S_4 . For example, a 90-degree rotation of the square is a 4-cycle on the vertices. In this way we obtain a copy of D_4 inside S_4 . The three essentially different labelings of the vertices of the square above embed D_4 into S_4 as three different subgroups of order 8:

$$\{1, (1234), (1432), (12)(34), (13)(24), (14)(23), (13), (24)\} = \langle (1234), (13) \rangle,$$

$$\{1, (1243), (1342), (12)(34), (13)(24), (14)(23), (14), (23)\} = \langle (1243), (14) \rangle,$$

$$\{1, (1324), (1423), (12)(34), (13)(24), (14)(23), (12), (34)\} = \langle (1324), (12) \rangle.$$

These are the 2-Sylow subgroups of S_4 .

Example 1.6. The group $\mathrm{SL}_2(\mathbf{Z}/(3))$ has order 24. It is not isomorphic to S_4 since its center $\{\pm I_2\}$ is nontrivial. By explicit calculation, $\mathrm{SL}_2(\mathbf{Z}/(3))$ has only 8 elements with 2-power order:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \\ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}.$$

These form the only 2-Sylow subgroup, which is isomorphic to Q_8 by labeling the matrices in the first row as $1, i, j, k$ and the matrices in the second row as $-1, -i, -j, -k$.

There are four 3-Sylow subgroups of $\mathrm{SL}_2(\mathbf{Z}/(3))$: $\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \rangle$, $\langle \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \rangle$, $\langle \begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix} \rangle$, and $\langle \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix} \rangle$.

Here are the Sylow theorems. They are often given in three parts. The result we call Sylow III* is not always stated explicitly as part of the Sylow theorems.

Theorem 1.7 (Sylow I). *A finite group G has a p -Sylow subgroup for every prime p and each p -subgroup of G lies in some p -Sylow subgroup of G .*

Theorem 1.8 (Sylow II). *For each prime p , the p -Sylow subgroups of G are conjugate.*

Theorem 1.9 (Sylow III). *For each prime p , let n_p be the number of p -Sylow subgroups of G . Write $|G| = p^k m$, where p doesn't divide m . Then*

$$n_p \equiv 1 \pmod{p} \text{ and } n_p \mid m.$$

Theorem 1.10 (Sylow III*). *For each prime p , let n_p be the number of p -Sylow subgroups of G . Then $n_p = [G : N(P)]$, where P is a p -Sylow subgroup and $N(P)$ is its normalizer.*

The existence part of Sylow I has been illustrated in all the previous examples.

Sylow II says for two p -Sylow subgroups H and K of G that there is some $g \in G$ such that $gHg^{-1} = K$. This is illustrated in the table below, where Example 1.2 is skipped since $\mathbf{Z}/(12)$ is abelian.

Example	Group	Size	p	H	K	g
1.3	A_4	12	3	$\langle(123)\rangle$	$\langle(124)\rangle$	(243)
1.4	D_6	12	2	$\langle r^3, s \rangle$	$\langle r^3, rs \rangle$	r^2
1.5	S_4	24	2	$\langle(1234), (13)\rangle$	$\langle(1243), (14)\rangle$	(34)
1.6	$SL_2(\mathbf{Z}/(3))$	24	3	$\langle\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\rangle$	$\langle\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\rangle$	$\begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$

When trying to conjugate one cyclic subgroup to another cyclic subgroup, be careful: not all generators of the two groups have to be conjugate. For example, in A_4 the subgroups $\langle(123)\rangle = \{(1), (123), (132)\}$ and $\langle(124)\rangle = \{(1), (124), (142)\}$ are conjugate, but the conjugacy class of (123) in A_4 is $\{(123), (142), (134), (243)\}$, so there's no way to conjugate (123) to (124) by an element of A_4 ; we must conjugate (123) to (142). The 3-cycles (123) and (124) are conjugate in S_4 , but not in A_4 . Similarly, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ are conjugate in $GL_2(\mathbf{Z}/(3))$ but not in $SL_2(\mathbf{Z}/(3))$, so when Sylow II says the subgroups $\langle\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\rangle$ and $\langle\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\rangle$ are conjugate in $SL_2(\mathbf{Z}/(3))$ a conjugating matrix must send $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ to $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

Let's see what Sylow III tells us about the number of 2-Sylow and 3-Sylow subgroups of a group of order 12. For $p = 2$ and $p = 3$ in Sylow III, the divisibility conditions are $n_2 \mid 3$ and $n_3 \mid 4$ and the congruence conditions are $n_2 \equiv 1 \pmod{2}$ and $n_3 \equiv 1 \pmod{3}$. The divisibility conditions imply n_2 is 1 or 3 and n_3 is 1, 2, or 4. The congruence $n_2 \equiv 1 \pmod{2}$ tells us nothing new (1 and 3 are both odd), but the congruence $n_3 \equiv 1 \pmod{3}$ rules out the option $n_3 = 2$. Therefore n_2 is 1 or 3 and n_3 is 1 or 4 when $|G| = 12$.

If $|G| = 24$ we again find n_2 is 1 or 3 while n_3 is 1 or 4. (For instance, from $n_3 \mid 8$ and $n_3 \equiv 1 \pmod{3}$ the only choices are $n_3 = 1$ and $n_3 = 4$.) Therefore as soon as we find more than one 2-Sylow subgroup there must be three of them, and as soon as we find more than one 3-Sylow subgroup there must be four of them. The table below shows the values of n_2 and n_3 in the examples above.

Example	Group	Size	n_2	n_3
1.2	$\mathbf{Z}/(12)$	12	1	1
1.3	A_4	12	1	4
1.4	D_6	12	3	1
1.5	S_4	24	3	4
1.6	$SL_2(\mathbf{Z}/(3))$	24	1	4

2. PROOF OF THE SYLOW THEOREMS

Our proof of the Sylow theorems will use group actions, which we assume the reader knows. The table below is a summary. For each theorem the table lists a group, a set it acts on, and the action. Let $\text{Syl}_p(G)$ be the set of p -Sylow subgroups of G , so $n_p = |\text{Syl}_p(G)|$.

Theorem	Group	Set	Action
Sylow I	p -subgroup H	G/H	left mult.
Sylow II	p -Sylow subgroup Q	G/P	left mult.
Sylow III ($n_p \equiv 1 \pmod{p}$)	$P \in \text{Syl}_p(G)$	$\text{Syl}_p(G)$	conjugation
Sylow III ($n_p \mid m$)	G	$\text{Syl}_p(G)$	conjugation
Sylow III*	G	$\text{Syl}_p(G)$	conjugation

The two conclusions of Sylow III are listed separately in the table since they are proved using different group actions.

Our proofs will usually involve the action of a p -group on a set and use the fixed-point congruence for such actions: when X is a finite set being acted on by a finite p -group Γ ,

$$(2.1) \quad |X| \equiv |\text{Fix}_\Gamma(X)| \pmod{p},$$

where $\text{Fix}_\Gamma(X)$ is the set of fixed points of Γ in X .

Proof of Sylow I: Let p^k be the highest power of p in $|G|$. The result is obvious if $k = 0$, since the trivial subgroup is a p -Sylow subgroup, so we can take $k \geq 1$, hence $p \mid |G|$.

Our strategy for proving Sylow I is to **prove a stronger result**: G has a subgroup of order p^i for $0 \leq i \leq k$. More precisely, if $|H| = p^i$ and $i < k$, we will show there is a p -subgroup $H' \supset H$ with $[H' : H] = p$, so $|H'| = p^{i+1}$. Then, starting with H as the trivial subgroup, repeat this process with H' in place of H to create larger subgroups

$$\{e\} = H_0 \subset H_1 \subset H_2 \subset \cdots$$

where $|H_i| = p^i$, and after k steps we reach H_k , which is a p -Sylow subgroup of G . And if we start with H as a p -subgroup, we will have shown H is contained in a p -Sylow subgroup.

Consider the left multiplication action of H on the left cosets G/H (this need not be a group). This is an action of a finite p -group H on the set G/H , so by the fixed-point congruence (2.1) for actions of nontrivial p -groups,

$$(2.2) \quad |G/H| \equiv |\text{Fix}_H(G/H)| \pmod{p}.$$

Here is what it means for gH in G/H to be fixed by the group H acting by left multiplication:

$$\begin{aligned} hgH = gH \text{ for all } h \in H &\iff hg \in gH \text{ for all } h \in H \\ &\iff g^{-1}hg \in H \text{ for all } h \in H \\ &\iff g^{-1}Hg \subset H \\ &\iff g^{-1}Hg = H \text{ because } |g^{-1}Hg| = |H| \\ &\iff g \in N(H). \end{aligned}$$

Thus $\text{Fix}_H(G/H) = \{gH : g \in N(H)\} = N(H)/H$, so (2.2) becomes

$$(2.3) \quad [G : H] \equiv [N(H) : H] \pmod{p}.$$

Because $H \triangleleft N(H)$, $N(H)/H$ is a group.

When $|H| = p^i$ and $i < k$, the index $[G : H]$ is divisible by p , so (2.3) implies $[N(H) : H]$ is divisible by p , so $N(H)/H$ is a group with order divisible by p . Thus $N(H)/H$ has a subgroup of order p by Cauchy's theorem. All subgroups of the quotient group $N(H)/H$

have the form H'/H , where H' is a subgroup between H and $N(H)$. Therefore a subgroup of order p in $N(H)/H$ is H'/H such that $[H' : H] = p$, so $|H'| = p|H| = p^{i+1}$. This can be repeated until we reach a subgroup of order p^k , and we're done.

Proof of Sylow II: For p -Sylow subgroups P and Q , we want to show they are conjugate. Let Q act on G/P by left multiplication. Since Q is a finite p -group, (2.1) says

$$|G/P| \equiv |\text{Fix}_Q(G/P)| \pmod{p}.$$

The left side is $[G : P]$, which is nonzero modulo p since P is a p -Sylow subgroup. Thus $|\text{Fix}_Q(G/P)|$ can't be 0, so there is a fixed point in G/P . Call it gP . That is, $qgP = gP$ for all $q \in Q$. Equivalently, $qg \in gP$ for all $q \in Q$, so $Q \subset gPg^{-1}$. Therefore $Q = gPg^{-1}$, since Q and gPg^{-1} have the same size and we're done.

Proof of Sylow III: We will prove $n_p \equiv 1 \pmod{p}$ and then $n_p \mid m$.

To show $n_p \equiv 1 \pmod{p}$, let P act on $\text{Syl}_p(G)$ by conjugation. The size of $\text{Syl}_p(G)$ is n_p . Since P is a finite p -group, (2.1) says

$$n_p \equiv |\{\text{fixed points}\}| \pmod{p}.$$

Fixed points for P acting by conjugation on $\text{Syl}_p(G)$ are $Q \in \text{Syl}_p(G)$ such that $gQg^{-1} = Q$ for all $g \in P$. One choice for Q is P . For all such Q , $P \subset N(Q)$. Also $Q \subset N(Q)$, so P and Q are p -Sylow subgroups in $N(Q)$. Applying Sylow II to the group $N(Q)$, P and Q are conjugate in $N(Q)$. Since $Q \triangleleft N(Q)$, the only subgroup of $N(Q)$ conjugate to Q is Q , so $P = Q$. Thus P is the only fixed point when P acts on $\text{Syl}_p(G)$, so $n_p \equiv 1 \pmod{p}$.

To show $n_p \mid m$, consider the action of G by conjugation on $\text{Syl}_p(G)$. Since the p -Sylow subgroups are conjugate to each other (Sylow II), there is one orbit. A set on which a group acts with one orbit has size dividing the size of the group, so $n_p \mid |G|$. From $n_p \equiv 1 \pmod{p}$, the number n_p is relatively prime to p , so $n_p \mid m$ and we're done.

Proof of Sylow III*: Let P be a p -Sylow subgroup of G and let G act on $\text{Syl}_p(G)$ by conjugation. By the orbit-stabilizer formula,

$$n_p = |\text{Syl}_p(G)| = [G : \text{Stab}_{\{P\}}].$$

The stabilizer $\text{Stab}_{\{P\}}$ of the "point" P in $\text{Syl}_p(G)$ (viewing P as a point is why we write $\{P\}$) is

$$\text{Stab}_{\{P\}} = \{g : gPg^{-1} = P\} = N(P).$$

Thus $n_p = [G : N(P)]$ and we're done.

In the proof of Sylow I, we saw that if H is a p -subgroup of G that is not a p -Sylow subgroup then $N(H)$ is strictly larger than H . What can be said about $N(P)$ when P is a p -Sylow subgroup? It may or may not be larger than P , but we will show that taking the normalizer a second time will not give anything new.

Theorem 2.1. *Let P be a p -Sylow subgroup of a finite group G . Then $N(N(P)) = N(P)$. More generally, if H is a subgroup of G that contains $N(P)$ then $N(H) = H$.*

Proof. We will prove $H \subset N(H)$ and $N(H) \subset H$. The containment $H \subset N(H)$ is easy,

To prove $N(H) \subset H$ let $x \in N(H)$, so $xHx^{-1} = H$. Since $P \subset N(P) \subset H$ we have $xPx^{-1} \subset xHx^{-1} = H$, so P and xPx^{-1} are both p -Sylow subgroups of H . By Sylow II for the group H , there is $y \in H$ such that $xPx^{-1} = yPy^{-1}$. Thus $y^{-1}xP(y^{-1}x)^{-1} = P$, so $y^{-1}x \in N(P) \subset H$, so $x \in yH = H$. \square

3. HISTORICAL REMARKS

Sylow's proof of his theorems appeared in [1]. Here is what he showed (of course, without using the label "Sylow subgroup").

- (1) There exist p -Sylow subgroups. Moreover, $[G : N(P)] \equiv 1 \pmod p$ for each p -Sylow subgroup P .
- (2) Let P be a p -Sylow subgroup. The number of p -Sylow subgroups is $[G : N(P)]$. All p -Sylow subgroups are conjugate.
- (3) Each finite p -group G with size p^k contains an increasing chain of subgroups

$$\{e\} = G_0 \subset G_1 \subset G_2 \subset \cdots \subset G_k \subset G,$$

where each subgroup has index p in the next one. In particular, $|G_i| = p^i$ for all i .

Here is how Sylow phrased his first theorem (the first item on the above list):¹

Si p^α désigne la plus grande puissance du nombre premier p qui divise l'ordre du groupe G , ce groupe contient un autre H de l'ordre p^α ; si de plus $p^\alpha \nu$ désigne l'ordre du plus grand groupe contenu dans G dont les substitutions sont permutables à H , l'ordre de G sera de la forme $p^\alpha \nu (pm + 1)$.

In English, using current terminology, this says

If p^α is the largest power of the prime p which divides the size of the group G , this group contains a subgroup H of order p^α ; if moreover $p^\alpha \nu$ is the size of the largest subgroup of G that normalizes H , the size of G is of the form $p^\alpha \nu (pm + 1)$.

Sylow did not have the abstract concept of a group: all groups for him arose as subgroups of symmetric groups, so groups were always "groupes de substitutions." The condition that an element $x \in G$ is "permutable" with a subgroup H means $xH = Hx$, or in other words $x \in N(H)$. The end of the first part of his theorem says the normalizer of a Sylow subgroup has index $pm + 1$ for some m , which means the index is $\equiv 1 \pmod p$.

4. ANALOGUES OF THE SYLOW THEOREMS

There are analogues of the first two Sylow theorems and Theorem 2.1 for other types of subgroups.

- (1) A *Hall subgroup* of a finite group G is a subgroup H whose order and index are relatively prime. For example, in a group of order 60 a subgroup of order 12 has index 5 and thus is a Hall subgroup. A p -subgroup is a Hall subgroup if and only if it is a p -Sylow subgroup. In 1928 Philip Hall proved that in every *solvable* group of order n there is a Hall subgroup of each order d dividing n where $(d, n/d) = 1$ and two Hall subgroups with the same order are conjugate. (The proof of Hall's theorem uses the conjugacy of p -Sylow subgroups.) Also the normalizer of a Hall subgroup of a solvable group is its own normalizer. Conversely, Hall proved that a finite group of order n that contains a Hall subgroup of order d for each d dividing n when $(d, n/d) = 1$ has to be solvable.
- (2) In a compact connected Lie group G , *maximal tori* (maximal connected abelian subgroups of G) satisfy properties analogous to Sylow subgroups: they exist, every torus is in a maximal torus, and all maximal tori are conjugate. The proof of conjugacy uses the Lefschetz fixed point theorem. (This plays a role analogous to the the

¹We modify some of his notation: he wrote the subgroup as g , not H , and the prime as n , not p .

fixed-point congruence (2.1) in the proof of the Sylow theorems.) Like normalizers of Sylow subgroups, the normalizer of a maximal torus is its own normalizer. Unlike Sylow subgroups, maximal tori are always abelian and every element of G is in some maximal torus.

- (3) In a connected linear algebraic group, *maximal connected unipotent subgroups* are like Sylow subgroups: they exist, every connected unipotent subgroup is in a maximal connected unipotent subgroup, and all maximal connected unipotent subgroups are conjugate. The proof of conjugacy uses the Borel fixed point theorem. The normalizer of a maximal connected unipotent subgroup is called a Borel subgroup, and like normalizers of Sylow subgroups each Borel subgroup is its own normalizer.

REFERENCES

- [1] L. Sylow, Théorèmes sur les groupes de substitutions, *Mathematische Annalen* **5** (1872), 584–594. Translation into English by Robert Wilson at http://www.maths.qmul.ac.uk/~raw/pubs_files/Sylow.pdf.