PROOF OF THE SYLOW THEOREMS

1. Introduction

The Sylow theorems concern subgroups with (maximal) prime power size. They are named after Ludwig Sylow, who proved them in a short paper on which his entire fame rests.

Definition 1.1. Let G be a finite group and p a prime. Write $\#G = p^k m$, where p does not divide m. A subgroup of G with size p^k is called a *p-Sylow subgroup* of G.

A subgroup that is a p-Sylow subgroup for some p is called a *Sylow subgroup.*

Example 1.2. In $\mathbf{Z}/(12)$, whose size is $12 = 4 \cdot 3$, a 2-Sylow subgroup has size 4 and a 3-Sylow subgroup has size 3. There is one 2-Sylow subgroup, $\{0, 3, 6, 9\}$, and one 3-Sylow subgroup, $\{0, 4, 8\}$. For $p \neq 2, 3$, the p-Sylow subgroups of $\mathbf{Z}/(12)$ are trivial.

Example 1.3. In A_4, whose size is also 12, there is one 2-Sylow subgroup,

$$\{(1), (12)(34), (13)(24), (14)(23)\},$$

but there are four 3-Sylow subgroups:

$$\{(1), (123), (132)\}, \{(1), (124), (142)\}, \{(1), (134), (143)\}, \{(1), (234), (243)\}.$$

Example 1.4. Another group of size 12 is D_6. There are three 2-Sylow subgroups:

$$\{1, r^3, s, r^3 s\}, \{1, r^3, rs, r^4 s\}, \{1, r^3, r^2 s, r^5 s\}.$$

The only elements of order 3 in D_6 are r^2 and r^4, so $\{1, r^2, r^4\}$ is the only 3-Sylow subgroup of D_6.

Example 1.5. Since $\#S_4 = 24 = 8 \cdot 3$, a 2-Sylow subgroup of S_4 has size 8 and a 3-Sylow subgroup has size 3. The 2-Sylow subgroups are interesting to work out, since they can be understood as copies of D_4 inside S_4. If we label the four vertices of a square in different ways as 1, 2, 3, and 4, and let D_4 act on the square then D_4 permutes the vertices. These vertex permutations provide an embedding of D_4 in S_4. For example, the counterclockwise rotation r in D_4 turns into a 4-cycle in S_4. There are essentially three different ways of labelling the four vertices:

$$\begin{array}{ccc}1 & & 2\\ & \square & \\ 3 & & 4\end{array}, \quad \begin{array}{ccc}1 & & 2\\ & \square & \\ 4 & & 3\end{array}, \quad \begin{array}{ccc}1 & & 3\\ & \square & \\ 4 & & 2\end{array}.$$

Any other vertex labelling is obtained from exactly one of these by applying an element of D_4. These three labellings embed D_4 as three different subgroups of S_4, which are the 2-Sylow subgroups.

There are four 3-Sylow subgroups of S_4, namely the 3-Sylow subgroups of A_4 from Example 1.3.

Example 1.6. Let $G = \mathrm{SL}_2(\mathbf{Z}/(3))$. Then $\#G = 24 = 8 \cdot 3$. A 2-Sylow subgroup of $\mathrm{SL}_2(\mathbf{Z}/(3))$ has size 8. An explicit tabulation shows there are only 8 elements with 2-power order in $\mathrm{SL}_2(\mathbf{Z}/(3))$:

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

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

These form the only 2-Sylow subgroup. This subgroup is isomorphic to Q_8 by labelling the top row as $1, i, j, k$ and the bottom row as $-1, -i, -j, -k$.

There are four 3-Sylow subgroups: $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$, $\langle \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) \rangle$, $\langle \left(\begin{smallmatrix} 0 & 1 \\ 2 & 2 \end{smallmatrix}\right) \rangle$, and $\langle \left(\begin{smallmatrix} 0 & 2 \\ 1 & 2 \end{smallmatrix}\right) \rangle$.

Here are the Sylow theorems. They are often given in three parts.

Theorem 1.7 (Sylow I)**.** *A finite group G has a p-Sylow subgroup for every prime p and any p-subgroup of G lies in a 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)**.** *Let n_p be the number of p-Sylow subgroups of G. Write $\#G = p^k m$, where p doesn't divide m. Then*

$$n_p | m \text{ and } n_p \equiv 1 \bmod p.$$

Theorem 1.10 (Sylow III*)**.** *Let n_p be the number of p-Sylow subgroups of G. Then $n_p = [G : \mathrm{N}(P)]$, where P is any p-Sylow subgroup and $\mathrm{N}(P)$ is its normalizer.*

By Sylow II, a p-subgroup of G that is a normal subgroup of G must lie in every p-Sylow subgroup. For example, when m is even the center of D_m is $\{1, r^{m/2}\}$ and this must lie in every 2-Sylow subgroup. We saw this in Example 1.4 when $m = 6$.

The divisibility conditions on n_p in Sylow III tell us when $\#G$ is 12 that $n_2|3$ and $n_3|4$. The congruence conditions in Sylow III are $n_2 \equiv 1 \bmod 2$ and $n_3 \equiv 1 \bmod 3$. The first one tells us nothing new (any factor of 3 is odd), but the second one rules out $n_3 = 2$. Therefore n_2 is 1 or 3 and n_3 is 1 or 4. If $\#G = 24$ the same conclusions follow. The table below shows some possibilities which arise, based on examples above.

Group	Size	n_2	n_3
$\mathbf{Z}/(12)$	12	1	1
A_4	12	1	4
D_6	12	3	1
S_4	24	3	4
$\mathrm{SL}_2(\mathbf{Z}/(3))$	24	1	4

The theorem we call Sylow III* is not always stated explicitly as part of the Sylow theorems. It follows from the proof of Sylow III.

Each of the proofs of the Sylow theorems will use a group action. The following table is a summary. For each theorem the table lists a group, a set it acts on, and the action. In the row for Sylow I, H is a p-subgroup of G. We write $\mathrm{Syl}_p(G)$ for the set of p-Sylow subgroups of G, so $n_p = \# \mathrm{Syl}_p(G)$. The two conclusions of Sylow III are listed separately in the table since they are proved using different group actions. Sylow III* will be proved in the course of proving $n_p|m$ in Sylow III.

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

2. Proof of Sylow I

We will argue recursively: for any p-subgroup $H \subset G$ which is not yet a p-Sylow subgroup (so $[G : H]$ is divisible by p), we will show $H \subset H'$ for some subgroup H' of G where $[H' : H] = p$. Repeating this argument for H' in the role of H, we can keep enlarging a p-subgroup to a subgroup p times bigger as long as we have not yet reached a p-Sylow subgroup. Eventually this process stops (we are in a finite group), and at that point we have reached a p-Sylow subgroup. This settles the existence of p-Sylow subgroups (by starting with H as the trivial subgroup) and also shows the containment of any p-subgroup in a p-Sylow subgroup (by starting with any p-subgroup H).

If p doesn't divide $\#G$ then the only p-subgroup of G is the trivial subgroup and Sylow I is trivial, so we may suppose $p|\#G$. If H is trivial then we can take for H' any subgroup of G with order p; such H' exists by Cauchy's theorem. Thus we may suppose $p|\#G$ and H is nontrivial.

Consider the action of H on the left coset space G/H by left multiplication. This is the action of a nontrivial p-group H on a finite set G/H, so by the fixed-point congruence for actions of p-groups,

$$\#(G/H) \equiv \mathrm{Fix}_H(G/H) \bmod p.$$

Let's unravel what it means for a coset gH to be a fixed point of H by left multiplication:

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

Thus $\mathrm{Fix}_H(G/H) = \{gH : g \in \mathrm{N}(H)\} = \mathrm{N}(H)/H$. This has absolutely nothing to do with H being a p-group. It describes the fixed points for left multiplication of any subgroup on the left cosets of that subgroup. When H is a (nontrivial) p-subgroup of G we can feed this into the fixed-point congruence:

$$[G : H] \equiv [\mathrm{N}(H) : H] \bmod p.$$

When H is not yet a Sylow subgroup, $[G : H]$ is divisible by p, so $[\mathrm{N}(H) : H]$ is divisible by p. Since $H \lhd \mathrm{N}(H)$, $[\mathrm{N}(H) : H] = \#(\mathrm{N}(H)/H)$, so $\mathrm{N}(H)/H$ is a group with order divisible by p. Thus $\mathrm{N}(H)/H$ has a subgroup of order p (Cauchy's theorem). This subgroup must have the form H'/H where $H \subset H' \subset \mathrm{N}(H)$, so $[H' : H] = p$ and we are done.

3. Proof of Sylow II

For the proof of Sylow II, pick two p-Sylow subgroups P and Q. We may suppose these are nontrivial subgroups, as otherwise the result is trivial.

Consider the action of Q on G/P by left multiplication. Since Q is a non-trivial p-group,

$$\#(G/P) \equiv \#\,\mathrm{Fix}_Q(G/P) \bmod p.$$

The left side is non-zero modulo p since P is a p-Sylow subgroup. Therefore there is a fixed point in G/P. Call it gP. That is, $qgP = gP$ for every $q \in Q$. Equivalently, $g^{-1}Qg \subset P$, so $Q \subset gPg^{-1}$. Both Q and gPg^{-1} have the same size, so $Q = gPg^{-1}$.

4. Proof of Sylow III

To show $n_p|m$, consider the action of G by conjugation on the set $\mathrm{Syl}_p(G)$ of p-Sylow subgroups of G. Then $\#\,\mathrm{Syl}_p(G) = n_p$, the number of p-Sylow subgroups of G. By Sylow II, this action has a single orbit.

Pick a p-Sylow of G, say P. Then, by the orbit-stabilizer formula,

$$n_p = \#\,\mathrm{Syl}_p(G) = [G : \mathrm{Stab}_{\{P\}}].$$

The stabilizer $\mathrm{Stab}_{\{P\}}$ is

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

Thus $n_p = [G : \mathrm{N}(P)]$ (which proves Sylow III*). Since $P \subset \mathrm{N}(P) \subset G$ and $m = [G : P]$, we have $n_p|m$.

To show $n_p \equiv 1 \bmod p$, consider the action of P (not G!) on $\mathrm{Syl}_p(G)$ by conjugation. If P is trivial (*i.e.*, p does not divide G) then $n_p = 1$ and we're done. Now suppose P is non-trivial. Then

$$n_p \equiv \#\{\text{fixed points}\} \bmod p.$$

The fixed points for P acting by conjugation on $\mathrm{Syl}_p(G)$ are the p-Sylows Q in G such that $gQg^{-1} = Q$ for every $g \in P$. This means $P \subset \mathrm{N}(Q)$. Also $Q \subset \mathrm{N}(Q)$, so P and Q are p-Sylow subgroups in $\mathrm{N}(Q)$. Applying Sylow II to the group $\mathrm{N}(Q)$, P and Q are conjugate in $\mathrm{N}(Q)$. Since $Q \lhd \mathrm{N}(Q)$, the only subgroup of $\mathrm{N}(Q)$ conjugate to Q is Q, so $P = Q$. Thus P is the only fixed point, so $n_p \equiv 1 \bmod p$.