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. Any 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 $\mathbb{Z}/(12)$, 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:

$$\{(1), (123), (132)\} = \langle (123) \rangle, \quad \{(1), (124), (142)\} = \langle (124) \rangle,$$

$$\{(1), (134), (143)\} = \langle (134) \rangle, \quad \{(1), (234), (243)\} = \langle (234) \rangle.$$

**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.

Any 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:

\[
\begin{align*}
&\{1, (1234), (1324), (1432), (12)(34), (14)(23), (13), (24)\} = \langle (1234), (13) \rangle, \\
&\{1, (1243), (1342), (12)(34), (13)(24), (14)(23), (14), (23)\} = \langle (1243), (14) \rangle, \\
&\{1, (1324), (1423), (12)(34), (13)(24), (14)(23), (12), (34)\} = \langle (1324), (12) \rangle.
\end{align*}
\]

These are the 2-Sylow subgroups of $S_4$.

**Example 1.6.** The group $\text{SL}_2(\mathbb{Z}/(3))$ has order 24. An explicit tabulation of the elements of this group reveals that there are only 8 elements in the group with 2-power order:

\[
\begin{align*}
&\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \\
&\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}.
\end{align*}
\]

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: $\langle (\frac{1}{1} \frac{1}{1}) \rangle, \langle (\frac{1}{1} \frac{0}{2}) \rangle, \langle (\frac{0}{2} \frac{1}{1}) \rangle, \text{ and } \langle (\frac{0}{2} \frac{2}{1}) \rangle$.

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 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). 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

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

**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 any $p$-Sylow subgroup and $N(P)$ is its normalizer.

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.

<table>
<thead>
<tr>
<th>Example</th>
<th>Group</th>
<th>Size</th>
<th>$p$</th>
<th>$H$</th>
<th>$K$</th>
<th>$g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3</td>
<td>$A_4$</td>
<td>12</td>
<td>3</td>
<td>$(123)$</td>
<td>$(124)$</td>
<td>$(243)$</td>
</tr>
<tr>
<td>1.4</td>
<td>$D_6$</td>
<td>12</td>
<td>2</td>
<td>$\langle r^3, s \rangle$</td>
<td>$\langle r^3, rs \rangle$</td>
<td>$r^2$</td>
</tr>
<tr>
<td>1.5</td>
<td>$S_4$</td>
<td>24</td>
<td>2</td>
<td>$\langle (1234), (13) \rangle$</td>
<td>$\langle (1243), (14) \rangle$</td>
<td>$(34)$</td>
</tr>
<tr>
<td>1.6</td>
<td>$\text{SL}_2(\mathbb{Z}/(3))$</td>
<td>24</td>
<td>3</td>
<td>$\langle \left(\begin{smallmatrix} 1 &amp; 1 \ 0 &amp; 1 \end{smallmatrix}\right) \rangle$</td>
<td>$\langle \left(\begin{smallmatrix} 1 &amp; 0 \ 0 &amp; 1 \end{smallmatrix}\right) \rangle$</td>
<td>$(\frac{0}{2})$</td>
</tr>
</tbody>
</table>

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 $\langle (123) \rangle = \langle (1), (123), (132) \rangle$ and $\langle (124) \rangle = \langle (1), (124), (142) \rangle$ 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 $(124)$. The 3-cycles $(123)$ and $(124)$ are conjugate in $S_4$, but not in $A_4$. Similarly, $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ and $\langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ are conjugate in $\text{GL}_2(\mathbb{Z}/(3))$ but not in $\text{SL}_2(\mathbb{Z}/(3))$, so when Sylow II says the subgroups $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ and $\langle \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ are conjugate in $\text{SL}_2(\mathbb{Z}/(3))$ a conjugating matrix must send $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)^2 = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$.

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 $n_2 \equiv 1 \mod 2$ and $n_3 \equiv 1 \mod 3$. The divisibility conditions imply $n_2$ is 1 or 3 and $n_3$ is 1, 2, or 4. The congruence $n_2 \equiv 1 \mod 2$ tells us nothing new (1 and 3 are both odd), but the congruence $n_3 \equiv 1 \mod 3$ 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 $n_3 \equiv 1 \mod 3$ 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.

<table>
<thead>
<tr>
<th>Group</th>
<th>Size</th>
<th>$n_2$</th>
<th>$n_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{Z}/(12)$</td>
<td>12</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$A_4$</td>
<td>12</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>$D_6$</td>
<td>12</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>$S_4$</td>
<td>24</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$\text{SL}_2(\mathbb{Z}/(3))$</td>
<td>24</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

2. **Proof of the Sylow Theorems**

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

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: \(|X| \equiv |\text{Fix}_G(X)| \text{ mod } p\), where \( X \) is a finite set being acted on by a finite \( p \)-group \( G \).

**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 \( p \) divides \(|G|\).

Our strategy for proving Sylow I is to prove a stronger result: there is a subgroup of order \( p^i \) for \( 0 \leq i \leq k \). More specifically, 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, we can repeat this process with \( H' \) in place of \( H \) to create a rising tower of 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 \( G \).

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 for actions of nontrivial \( p \)-groups,

\[
|G/H| \equiv |\text{Fix}_H(G/H)| \text{ mod } p.
\]  

(2.1)

Let’s unravel what it means for a coset \( gH \) in \( G/H \) to be a fixed point by the group \( H \) under left multiplication:

\[
hgH = gH \text{ for all } h \in H \iff hg \in gH \text{ for all } h \in H
\]

\[
\iff g^{-1}hg \in H \text{ for all } h \in H
\]

\[
\iff g^{-1}Hg \subset H
\]

\[
\iff g^{-1}Hg = H \text{ because } |g^{-1}Hg| = |H|
\]

\[
\iff g \in N(H).
\]

Thus \( \text{Fix}_H(G/H) = \{gH : g \in N(H)\} = N(H)/H \), so (2.1) becomes

\[
|G : H| \equiv |N(H) : H| \text{ mod } p.
\]  

(2.2)

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

When \(|H| = p^i\) and \( i < k \), the index \(|G : H|\) is divisible by \( p \), so the congruence (2.2) 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 \(|H'| = p|H| = p^{i+1}\).

**Proof of Sylow II:** Pick two \( p \)-Sylow subgroups \( P \) and \( Q \). We want to show they are conjugate.
Consider the action of $Q$ on $G/P$ by left multiplication. Since $Q$ is a finite $p$-group,
\[ |G/P| \equiv |\text{Fix}_Q(G/P)| \mod p. \]

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.

**Proof of Sylow III:** We will prove $n_p \equiv 1 \mod p$ and then $n_p \mid m$.

To show $n_p \equiv 1 \mod p$, consider the action of $P$ on the set $\text{Syl}_p(G)$ by conjugation. The size of $\text{Syl}_p(G)$ is $n_p$. Since $P$ is a finite $p$-group,
\[ n_p \equiv |\text{fixed points}| \mod p. \]

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 any 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 $Q \subset N(Q)$, 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 $n_p \equiv 1 \mod p$.

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 $n_p \equiv 1 \mod p$, the number $n_p$ is relatively prime to $p$, so $n_p \mid m$.

**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 $\text{Stab}_P$ is
\[ \text{Stab}_P = \{g : gPg^{-1} = P\} = N(P). \]

Thus $n_p = [G : N(P)]$.

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, $[G : N(P)] \equiv 1 \mod p$ for any $p$-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) Any 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):\(^1\)

---

\(^1\)We modify some of his notation: he wrote the subgroup as $g$, not $H$, and the prime as $n$, not $p$. 
Si $p^\alpha$ 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 $p^\alpha$; 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 $p^\alpha$ is the largest power of the prime $p$ which divides the size of the group $G$, this group contains a subgroup $H$ of order $p^\alpha$; 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 $\equiv 1 \mod p$.

4. Analogues of the Sylow Theorems

There are analogues of the Sylow theorems 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 any 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 Sylow subgroup. In 1928 Philip Hall proved in every solvable group of order $n$ that there is a Hall subgroup of each order $d$ dividing $n$ where $(d, n/d) = 1$ and any two Hall subgroups with the same order are conjugate. Conversely, Hall proved that a finite group of order $n$ containing a Hall subgroup of order $d$ for each $d$ dividing $n$ such that $(d, n/d) = 1$ has to be a solvable group.

(2) In a compact connected Lie group, the maximal tori (maximal connected abelian subgroups) satisfy properties analogous to Sylow subgroups: they exist, every torus is contained in a maximal torus, and all maximal tori are conjugate. Of course, unlike Sylow subgroups, maximal tori are always abelian.

(3) In a connected linear algebraic group, the maximal unipotent subgroups are like Sylow subgroups: they exist, every unipotent subgroup is contained in a maximal unipotent subgroup, and all maximal unipotent subgroups are conjugate. The normalizer of a maximal unipotent subgroup is called a Borel subgroup, and like the normalizers of Sylow subgroups all Borel subgroups equal their own normalizer. For the group $GL_n(\mathbb{Z}/(p))$, its subgroup of upper triangular matrices with 1’s along the main diagonal

$$
\begin{pmatrix}
1 & * & \cdots & * \\
0 & 1 & \cdots & * \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{pmatrix}
$$

is both a $p$-Sylow subgroup and a maximal unipotent subgroup.

References