MORE ON THE SYLOW THEOREMS

1. Introduction

Several alternative proofs of the Sylow theorems are collected here. Section 2 has a proof of Sylow I by Sylow, Section 3 has a proof of Sylow I by Frobenius, and Section 4 has an extension of Sylow I and II to p-subgroups due to Sylow. Section 5 discusses some history related to the Sylow theorems and formulates (but does not prove) two extensions of Sylow III to p-subgroups, by Frobenius and Weisner.

2. Sylow I by Sylow

In modern language, here is Sylow's proof that his subgroups exist.

Pick a prime p dividing $|G|$. Let P be a p-subgroup of G that is as large as possible. We call P a maximal p-subgroup. We do not yet know its size is the biggest p-power in $|G|$. The goal is to show $[G : P] \not\equiv 0 \bmod p$, so $|P|$ is the largest p-power dividing $|G|$.

Let $N = \mathrm{N}(P)$ be the normalizer of P in G. Then all the elements of p-power order in N lie in P. Indeed, an element of N with p-power order that is not in P would give a non-identity element of p-power order in N/P. Then we could take inverse images through the projection $N \to N/P$ to find a p-subgroup inside N properly containing P, but this contradicts the maximality of P as a p-subgroup of G. (Here is a more elementary approach: if $n \in N$ is not in P then $\langle n, P\rangle = \{n^i x : i \in \mathbf{Z}, x \in P\}$ is a subgroup of N since $P \lhd N$, $\langle n, P\rangle$ strictly contains P, and $|\langle n, P\rangle| = |\langle n\rangle||P|/|\langle n\rangle P|$, so if n has p-power order then $\langle n, P\rangle$ is a p-group in N strictly larger than P, which contradicts the meaning of P being a maximal p-subgroup of G.)

Since there are no non-trivial elements of p-power order in N/P, the index $[N : P]$ is not divisible by p by Cauchy's theorem: if $[N : P]$ were divisible by p then N/P would have an element $\overline{n}$ of order p by Cauchy's theorem, and $n^p \in P$ implies n has p-power order (since P is a p-group), which makes n an element of p-power order in N that is not in P.

Now let the p-group P act on G/N by left multiplication. Since $P \subset N$, $tN = N$ for all $t \in P$, so N is a fixed point in this group action. Let's show N is the only fixed point for left multiplication of P on G/N. Suppose gN is a fixed point, so for every $t \in P$, $tgN = gN$. Thus $g^{-1}tg \in N$, so $g^{-1}Pg \subset N$. Because $g^{-1}Pg$ is a p-group and (as shown above) all elements of p-power order in N lie in P, $g^{-1}Pg \subset P$, and therefore $g^{-1}Pg = P$ since $g^{-1}Pg$ and P have the same size. Thus $g \in \mathrm{N}(P) = N$, so $gN = N$.

Thus $N \in G/N$ is the only fixed point for left multiplication of P on G/N. Every other orbit of P on G/N has size divisible by p, so by writing the order of G/N as the sum of the sizes of its P-orbits (the orbit-stabilizer formula) and reducing everything mod p, we get $[G : N] \equiv 1 \bmod p$. Therefore, since we know $|N/P|$ is not divisible by p,

$$[G : P] = [G : N][N : P] \equiv [N : P] \not\equiv 0 \bmod p,$$

which proves P is a p-Sylow subgroup of G.

3. Sylow I by Frobenius

Here is Frobenius' first proof on the existence of Sylow subgroups. It takes for granted that there are Sylow subgroups of symmetric groups; this had been shown in a paper of Cauchy before Sylow's work.

By Cayley's theorem, every finite group can be embedded in a symmetric group. Given a finite group G, suppose we have $G \subset S_n$. Pick a prime p. By Cauchy's work, S_n has a p-Sylow subgroup, say P. Consider the (G, P) *double coset* decomposition of S_n:

$$S_n = \bigcup_i G\sigma_i P.$$

Each double coset $G\sigma_i P$ has size $|G||P|/|G \cap \sigma_i^{-1} P \sigma_i|$, which is divisible by $|P|$. Therefore

$$\frac{|S_n|}{|P|} = \sum_i \frac{|G|}{|G \cap \sigma_i^{-1} P \sigma_i|}.$$

Since $|S_n|/|P| \not\equiv 0 \bmod p$, one of the terms in the sum is not divisible by p. Let it be the j-th term. Then $G \cap \sigma_j^{-1} P \sigma_j$ is a p-group (since it's a subgroup of $\sigma_j^{-1} P \sigma_j$) with maximal p-power size inside of G (since its ratio with $|G|$ is not divisible by p). Thus $G \cap \sigma_j^{-1} P \sigma_j$ is a p-Sylow subgroup of G.

4. Sylow's extension of Sylow I and II to p-power subgroups

It is natural to ask how the Sylow theorems can be extended to p-subgroups that are not p-Sylow subgroups. The first Sylow theorem generalizes as follows, and was proved by Sylow in his original paper.

Theorem 4.1. *If $p^d \mid |G|$ then there is a subgroup of G with size p^d.*

Part of the second Sylow theorem extends to non-Sylow p-subgroups, and was also proved by Sylow.

Theorem 4.2. *Let G be a finite group. If $p^d \mid |G|$ and $d > 0$ then each subgroup of G with size p^{d-1} has index p in a subgroup of G.*

Let p^k be the largest p-power dividing $|G|$. Since the trivial subgroup is a p-group, Theorem 4.2 tells us we can make a nested chain of p-subgroups of G from the identity all the way up to p-Sylow subgroup

$$\{e\} = G_0 \subset G_1 \subset G_2 \subset \cdots \subset G_k \subset G, \tag{4.1}$$

where $[G_i : G_{i-1}] = p$, so $|G_i| = p^i$. In particular, Theorem 4.2 has Theorem 4.1 as a special case. We will not use the proof of Theorem 4.1 in our proof of Theorem 4.2, so the proof of Theorem 4.1 could be disregarded; we include it just for the sake of illustrating some techniques.

Knowing we always have a chain as in (4.1) for each finite group G (and prime p), we can build one that passes through a p-subgroup H of G: starting with H use Theorem 4.2 to build up successively larger subgroups p times as big until we end at a p-Sylow subgroup of G. Then apply Theorem 4.2 to H as the top group and build up to it from the identity by a chain of p-subgroups as in (4.1).

The conjugacy property of p-Sylow subgroups in Sylow II does *not* carry over to proper p-subgroups. That is, it is *not* true in general that p-subgroups of a common non-maximal size are all conjugate. Said differently, the number of conjugacy classes of p-subgroups

with a fixed non-maximal size can be greater than 1. For instance, in S_{p^2} the number of conjugacy classes of subgroups of size p is p. Or if G is a non-abelian p-group containing an element h of order p not in the center of G, then $\langle h\rangle$ and a subgroup of order p in the center of G are nonconjugate subgroups of order p.

We now proceed to the proofs of Theorems 4.1 and 4.2, first treating Theorem 4.1.

Proof. We induct on the size of G. The case when $|G| = 1$ or prime is trivial. Now suppose $|G| > 1$ and the theorem is proved for all groups of smaller size. That is, we assume each group G' with $|G'| < |G|$ has a subgroup of size equal to an arbitrary prime power dividing $|G'|$.

Choose a prime power p^d dividing $|G|$, with $p^d > 1$. We seek a subgroup of G with size p^d. If G has a proper subgroup H such that $p^d \mid |H|$, then we're done: H has a subgroup of size p^d by induction (since $|H| < |G|$) and this subgroup is in G too.

Now we suppose every proper subgroup $H \subset G$ has size not divisible by p^d. Since $|G| = |H|[G : H]$ is divisible by p^d, we see every proper subgroup of G has index divisible by p. Consider the class equation

$$|G| = |Z(G)| + \sum_{i=1}^{r} [G : Z(g_i)],$$

where $g_1, \ldots, g_r$ represent the conjugacy classes of size greater than 1. We have $p \mid |G|$ and $p \mid [G : Z(g_i)]$ for each i since each subgroup $Z(g_i)$ of G is proper (if $Z(g_i) = G$ then g_i would be in a conjugacy class of size 1, which isn't true). Therefore $p \mid |Z(G)|$. By Cauchy's theorem, $Z(G)$ has an element of order p, say z. As $z \in Z(G)$, $\langle z\rangle \lhd G$.

We now consider the quotient group $G/\langle z\rangle$, which is a group with size less than that of G. Since $p^{d-1} \mid |G/\langle z\rangle|$, by induction $G/\langle z\rangle$ has a subgroup with size p^{d-1}. Its inverse image under $G \to G/\langle z\rangle$ is a subgroup of G with size $p \cdot p^{d-1} = p^d$. □

Remark 4.3. If we take for p^d the largest power of p dividing $|G|$, so p^{d-1} at the end of the proof is the largest power of p dividing $|G/\langle z\rangle|$, this proof shows G contains a p-Sylow subgroup by an argument that is quite similar to one of the proofs of Cauchy's theorem (but note this proof uses Cauchy's theorem).

Now we prove Theorem 4.2.

Proof. The case $d = 1$ says there is a subgroup of size p in G. This is Cauchy's theorem.

Now take $d > 1$. Let H be a subgroup of G with size p^{d-1}. We want to find a subgroup $K \subset G$ in which H has index p. Consider the left multiplication action of H (not G!) on G/H. Since H is a non-trivial p-group,

$$|G/H| \equiv |\{\text{fixed points}\}| \bmod p. \tag{4.2}$$

The left side of the congruence is $[G : H]$, which is divisible by p. Which cosets in G/H are fixed points? They are

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

Therefore the set of fixed points of H acting on G/H is $\mathrm{N}(H)/H$, which has the structure of a group since $H \lhd \mathrm{N}(H)$. By (4.2), $p \mid |\mathrm{N}(H)/H|$, so Cauchy tells us there is a subgroup

$H' \subset \mathrm{N}(H)/H$ of order p. Its inverse image under $\mathrm{N}(H) \to \mathrm{N}(H)/H$ is a subgroup of $\mathrm{N}(H)$ with size $p \cdot p^{d-1} = p^d$, and it contains H with index p. □

Corollary 4.4. *Let H be a p-subgroup of the finite group G. Then*

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

In particular, if $p \mid [G : H]$ then $H \neq \mathrm{N}(H)$.

Proof. The congruence here is (4.2). When $p \mid [G : H]$, $[\mathrm{N}(H) : H] \neq 1$, so $H \neq \mathrm{N}(H)$. □

5. Historical Remarks

Sylow's proof of his theorems appeared in [2]. While the original version of his theorems came in three parts, they do not correspond exactly to the way they are usually labelled today. Here is what he showed (of course, without using the label "Sylow subgroup").

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

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

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

Applying this third theorem to a p-Sylow subgroup of a finite group proves the existence of p-subgroups of every possible size in the group once we have a p-Sylow subgroup.

Sylow's choice of notation is amusing. Here is exactly how he phrased his first theorem (the first item on the above list):

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

An English translation really shouldn't be needed even if you haven't studied French, but here is one:

> If n^α is the largest power of the prime n that divides the size of the group G, this group contains a subgroup g of order n^α; if moreover $n^\alpha\nu$ is the size of the largest subgroup of G containing elements commuting with g, the size of G is of the form $n^\alpha\nu(np+1)$.

Notice n is the prime, while p is something else, and g denotes a subgroup. 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 $x \in G$ commutes with a subgroup H means $xH = Hx$, or in other words $x \in \mathrm{N}(H)$. So the last part of the excerpt is saying the normalizer of a Sylow subgroup has index $np+1$ (n is the prime!) for some p, which means the index is $\equiv 1 \bmod n$.

The existence of Sylow subgroups was established in a special case by Cauchy (1845) a quarter-century before Sylow's work, and as part of his proof Cauchy constructed p-Sylow subgroups of the symmetric groups.

Following Sylow's work, Frobenius set himself the task of finding alternate proofs. In 1887, he gave two new proofs of Sylow's theorems. The first proof used Cauchy's theorem

and double cosets. The second proof used the class equation. That is the context where conjugacy classes (and the class equation) were first introduced historically.

Frobenius not only reproved the Sylow theorems, but he extended part of Sylow III to p-subgroups of every fixed size, as follows.

Theorem 5.1 (Frobenius, 1895)**.** *If $p^r \mid |G|$, the number of subgroups of G with size p^r is $\equiv 1 \bmod p$.*

Forty years later, this was generalized still further.

Theorem 5.2 (Weisner, 1935)**.** *For a p-subgroup $K \subset G$, the number of intermediate p-subgroups $K \subset H \subset G$ with a fixed size is $\equiv 1 \bmod p$.*

Proof. See the handout on transitive group actions at https://kconrad.math.uconn.edu/blurbs/grouptheory/transitive.pdf.

□

Theorem 5.1 is the special case of Theorem 5.2 where K is the trivial subgroup.

References

[1] R. Gow, Sylow's proof of Sylow's theorem, *Irish Math. Soc. Bull.* (1994), 55–63.
[2] L. Sylow, Théorèmes sur les groupes de substitutions, *Mathematische Annalen* **5** (1872), 584–594.
[3] W. C. Waterhouse, The early proofs of Sylow's theorem, *Arch. Hist. Exact Sci.* **21** (1979/80), 279–290.