8. The Sylow Theorems

In the sequel, G is always a finite group.

Definition. For $a \in G$, the set $C(a) := \{x \in G : xax^{-1} = a\}$ is called the **centralizer** of a in G.

Note that $x \in C(a)$ *iff* $xa = ax$, and that for any $a \in G$ we have $a \in C(a)$.

Fact 8.1. For any $a \in G$, $C(a) \leqslant G$.

Proof. We have to verify the axioms (A0), (A1) and (A2).

(A0) For $x, y \in C(a)$ we have

$$(xy)a = x(ya) \underset{\substack{\uparrow \\ y \in C(a)}}{=} x(ay) = (xa)y \underset{\substack{\uparrow \\ x \in C(a)}}{=} (ax)y = a(xy)\,,$$

hence, $xy \in C(a)$.

(A1) $ea = ae$, thus, $e \in C(a)$.

(A2) If $x \in C(a)$, then

$$x^{-1}a = x^{-1}a(xx^{-1}) = x^{-1}(ax)x^{-1} \underset{\substack{\uparrow \\ x \in C(a)}}{=} x^{-1}(xa)x^{-1} = (x^{-1}x)ax^{-1} = ax^{-1}\,,$$

hence, $x^{-1} \in C(a)$. ⊣

Definition. For $a \in G$, the set $\operatorname{orbit}(a) := \{xax^{-1} : x \in G\}$ is called the **orbit** of a.

Fact 8.2. For $a, a' \in G$ we either have $\operatorname{orbit}(a) = \operatorname{orbit}(a')$ or $\operatorname{orbit}(a) \cap \operatorname{orbit}(a') = \emptyset$. Further, $|\operatorname{orbit}(a)| = 1$ *iff* $a \in Z(G)$.

Proof. If $\operatorname{orbit}(a) \cap \operatorname{orbit}(a') \neq \emptyset$, then $xax^{-1} = ya'y^{-1}$ (for some $x, y \in G$). Thus, $a' = y^{-1}xax^{-1}y = y^{-1}xa(y^{-1}x)^{-1} \in \operatorname{orbit}(a)$ and $a = x^{-1}ya'y^{-1}x = x^{-1}ya'(x^{-1}y)^{-1} \in \operatorname{orbit}(a')$, which implies that $\operatorname{orbit}(a) = \operatorname{orbit}(a')$.

If $|\operatorname{orbit}(a)| = 1$, then for all $x \in G$ we have $xax^{-1} = a$, thus, for all $x \in G$ we have $xa = ax$, which implies $Z(G)$. On the other hand, if $a \in Z(G)$, then $xax^{-1} = a$ (for all $x \in G$), thus, $|\operatorname{orbit}(a)| = 1$. ⊣

Lemma 8.3. For every $a \in G$ we have

$$|\operatorname{orbit}(a)| = |G : C(a)|\,.$$

Proof. $|G : C(a)| = |G/C(a)| = \big|\{xC(a) : x \in G\}\big|$. Further, we have

$$xC(a) = yC(a) \iff x^{-1}y \in C(a) \iff (x^{-1}y)a(y^{-1}x = a \iff yay^{-1} = xax^{-1}\,,$$

which implies that $\big|\{xax^{-1} : x \in G\}\big| = \big|\{xC(a) : x \in G\}\big|$. ⊣

As a consequence of Fact 8.2 and Lemma 8.3 we get

Corollary 8.4. Let $a_1, \ldots, a_n$ be representatives for the n orbits which have size larger than 1. Then

$$|G| = |Z(G)| + \sum_{i=1}^{n} |\operatorname{orbit}(a_i)| = |Z(G)| + \sum_{i=1}^{n} |G : C(a_i)|\,.$$

Proposition 8.5. If G is a group of order p^2, where p is prime, then G is abelian.

Proof. Assume that G is not abelian, then, by Corollary 8.4, we can choose some $a_1, \dots, a_n \in G$ such that $|\operatorname{orbit}(a_i)| > 1$ (for all $a_i \in \{a_1, \dots, a_n\}$) and $p^2 = |G| = |Z(G)| + \sum_{i=1}^{n} |G : C(a_i)|$. By Lemma 8.3, for each $a_i \in \{a_1, \dots, a_n\}$ we get $1 < |\operatorname{orbit}(a_i)| = |G : C(a_i)|$, so, $p \mid |C(a_i)|$, and therefore $p \mid Z(G)$, which implies that $|Z(G)| \geq p$. If we assume that G is not abelian, then $Z(G) < G$, thus, $|Z(G)| = p$. Choose some $x \in G \setminus Z(G)$, then $Z(G) \leqslant C(x)$, and since $x \in C(x)$ we get $|C(x)| \geq p + 1$. Now, since $C(x) \leqslant G$, $|C(x)| \mid |G| = p^2$, and because $|C(x)| \geq p + 1$ we get $C(x) = G$, thus $x \in Z(G)$, which is absurd. Hence, we must have $Z(G) = G$, which shows that G is abelian. ⊣

THEOREM 8.6 (Cauchy). Suppose that $p \mid |G|$ for some prime number p. Then there is an element $g \in G$ of order p.

Proof. The proof is by induction on $|G|$. If $|G| = 1$, then the result is vacuously true. Now, let us assume that $|G| > 1$ and that for every proper subgroup $H < G$ we have $p \nmid |H|$, (in other words, $p \mid |G : H|$), else we are home by induction. By Corollary 8.4 and by our assumption we get $p \mid |Z(G)|$, so, $G = Z(G)$ which implies that G is abelian. A proper subgroup $H < G$ is called maximal if $H \leqslant H' \leqslant G$ implies $H' = H$ or $H' = G$. If H, K are distinct maximal proper subgroups of G, then $HK \leqslant G$ (since G is abelian) and by maximality of H and K we get $HK = G$ (since $H, K \leqslant HK$). Now, $|G| = |HK| = \frac{|H| \cdot |K|}{|H \cap K|}$, but because $p \nmid |H|$ and $p \nmid |K|$, this implies $p \nmid |G|$, which is a contradiction. Therefore, G has a unique maximal proper subgroup, say M. Since M is the only maximal proper subgroup of G, all proper subgroups $H < G$ are subgroups of M. Choose $g \in G$ with $g \notin M$, then $\langle g \rangle = G$, (since otherwise, $\leq g\langle \leqslant M)$, and hence, G is cyclic. The order of g is $|G|$, and if we put $n = \frac{|G|}{p}$, then $\langle g^n \rangle$ is a subgroup of G of order p, which completes the proof. ⊣

DEFINITION. Let $H \leqslant G$, then the set $N(H) := \{x \in G : xHx^{-1} = H\}$ is called the **normalizer** of H in G, and $\operatorname{orbit}(H) := \{xHx^{-1} : x \in G\}$ is called the **orbit** of H.

FACT 8.7. For every $H \leqslant G$, $N(H) \leqslant G$ and $|\operatorname{orbit}(H)| = |G : N(H)|$.

Proof. Just follow the proofs of Fact 8.1 and Lemma 8.3. ⊣

FACT 8.8. For every $H \leqslant G$, $H \trianglelefteq N(H)$.

Proof. By definition, for every $x \in N(H)$ we have $xHx^{-1} = H$, thus, $H \trianglelefteq N(H)$. ⊣

LEMMA 8.9. Let G be such that $|G| = p^m n$, where p is prime, $m, n > 0$ and $p \nmid n$, and let $P, Q \leqslant G$ be such that $|P| = |Q| = p^m$. Then $Q \leqslant N(P)$ if and only if $Q = P$.

Proof. Of course, $Q = P$ implies $Q \leqslant N(P)$. On the other hand, if $Q \leqslant N(P)$, then, since $P \trianglelefteq N(P)$ (by Fact 8.8), $PQ \leqslant N(P) \leqslant G$. Thus,

$$|PQ| = \frac{|P| \cdot |Q|}{|P \cap Q|} = \frac{p^m \cdot p^m}{|P \cap Q|}$$

must divide $|G| = p^m n$, which implies $|P \cap Q| = p^m$, hence, $Q = P$. ⊣

Definition. Let G be a finite group of order $p^m n$, where p is prime and does not divide n. Then any subgroup of G of order p^m is called a **Sylow p-subgroup** of G, and the set of all such subgroups of G is denoted $\mathrm{Syl}_p(G)$.

In order to state Sylow's Theorem, we need one more definition.

Definition. Two subgroups H_1 and H_2 of a group G are called conjugate in G if $H_1 = xH_2x^{-1}$ for some $x \in G$.

Theorem 8.10 (Sylow). Let G be a finite group of order $p^m n$, where p is prime and does not divide n.

(i) There is a Sylow p-subgroup P of G.
(ii) All elements of $\mathrm{Syl}_p(G)$ are conjugate in G.
(iii) $|\mathrm{Syl}_p(G)| \equiv 1 \mod p$.
(iv) $|\mathrm{Syl}_p(G)| \mid n$.

Proof. We prove (i) by induction on $|G|$. If $|G| = 1$, then the result is vacuously true, and therefore we may assume that $|G| > 1$. By Corollary 8.4 we have $|G| = |Z(G)| + \sum_{j=1}^{s} |G : C(x_j)|$, where the x_j are a collection of representatives for those orbits which are not singletons. Thus, each $C(x_j)$ is a proper subgroup of G. If $p \mid |G : C(x_j)|$ for every $1 \leq j \leq s$, then $p \mid |Z(G)| \neq 1$. Thanks to Cauchy's Theorem 8.6 we can choose $z \in Z(G)$ of order p, so, since $z \in Z(G)$, $\langle z \rangle \trianglelefteq G$. Let $\pi : G \to G/\langle z \rangle$ be the natural projection. By induction, there is a Sylow p-subgroup P_1 of $G/\langle z \rangle$. This group has order p^{m-1}, since $|G/\langle z \rangle| = p^{m-1}n$. The preimage of P_1 under π is $P \leqslant G$, where $P/\langle z \rangle$ has order $p^{m-1} = \frac{|P|}{p}$. Thus, $|P| = p^m$ and we have found a Sylow p-subgroup of G. The other possibility is that there is some x_j with $p \nmid |G : C(x_j)|$, so, $|G : C(x_j)| = p^m k$ with $k < n$ and $p \nmid k$. By induction, $C(x_j)$ has a Sylow p-subgroup P of order p^m, and since $P \leqslant G$, P is a Sylow p-subgroup of G.

For part (ii) and (iii), let P be a Sylow p-subgroup of G. Let $\Omega = \{xPx^{-1} : x \in G\}$ denote the set of all G-conjugates of P. Now, by Fact 8.7 we have $|\Omega| = |G : N(P)|$. Further, for $P_i \in \Omega$, let $\Omega_i = \{yP_iy^{-1} : y \in P\}$, then Ω is the disjoint union of some Ω_i's, so, $|\Omega| = \sum_i |\Omega_i|$. Again by Fact 8.7 we get $|\Omega_i| = |P : N(P_i) \cap P|$, which tells us that the orbits Ω_i have size divisible by p, unless $P \leqslant N(P_i)$, in which case $|\Omega_i| = 1$ and $P = P_i$ (by Lemma 8.9). Hence, of the orbits Ω_i there is exactly one of length 1 and all the others have size divisible by p, thus, $|\Omega| = \sum_i |\Omega_i| \equiv 1 \pmod p$. If we can show that $\Omega = \mathrm{Syl}_p(G)$, then we are done. So, assume towards a contradiction that $\Omega \neq \mathrm{Syl}_p(G)$, which means that there is a Sylow p-subgroup Q which is not a conjugate of P. Now, all Q-orbits $\Omega_i = \{yP_iy^{-1} : y \in Q\}$, where $P_i \in \Omega$ have size divisible by p, since otherwise, $Q \leqslant N(P_i)$ (for some i) and therefore $Q = P_i$ (by Lemma 8.9), which implies that Q is a conjugate of P. Since Ω is a disjoint union of sets – namely the Ω_i's – of size divisible by p we deduce that $|\Omega| \equiv 0 \pmod p$. However, we already know that $|\Omega| \equiv 1 \pmod p$ so this is absurd. Thus, $\Omega = \mathrm{Syl}_p(G)$, which implies that all Sylow p-subgroups of G are conjugate and $|\mathrm{Syl}_p(G)| \equiv 1 \pmod p$.

To verify (iv), let $P \in \mathrm{Syl}_p(G)$. Then, by (ii), $\mathrm{Syl}_p(G) = \{xPx^{-1} : x \in G\}$, and by Fact 8.7 we get $|\mathrm{Syl}_p(G)| = |G : N(P)|$. Since $P \leqslant N(P)$ it follows that $p^m \mid |N(P)|$, and so $|G : N(P)|$ must divide n. ⊣

As a consequence of Theorem 8.10 (ii) we get

Corollary 8.11. Let G be a finite group of order $p^m n$, where $n, m > 0$ and p is prime and does not divide n. Then $|\operatorname{Syl}_p(G)| = 1$ if and only if the unique Sylow p-subgroup is a normal subgroup of G. In particular, $|\operatorname{Syl}_p(G)| = 1$ implies that G is not simple.