Applications of Group Actions: Cauchy's Theorem and Sylow's Theorems

R. C. Daileda

The group-theoretic result known as Cauchy's theorem posits the existence of elements of all possible prime orders in a finite group. It can be viewed as a partial converse to Lagrange's theorem, and is the first step in the direction of Sylow theory, which studies maximal prime-power order subgroups. The goal of this note is to present proofs of Cauchy's theorem and Sylow's theorems based almost entirely on the application of group actions and the class equation (a.k.a. the orbit-stabilizer theorem). These proofs demonstrate the flexibility and utility of group actions in general. As we will see, the simplicity of the class equation, in particular, often reduces apparently difficult counting problems to simple statements about modular arithmetic.

1 Cauchy's Theorem

Here we present a simple proof of Cauchy's theorem that makes use of the cyclic permutation action of $\mathbb{Z}/n\mathbb{Z}$ on n-tuples. Although not the original proof, it is perhaps the most widely known; it is certainly the author's favorite.

Theorem 1 (Cauchy)**.** *Let G be a finite group and let p be a prime number. If p divides $|G|$, then G has an element of order p.*

Proof. Let

$$S = \{(a_1, a_2, \dots, a_p) \,|\, a_i \in G, a_0 a_1 \cdots a_{p-1} = e\},$$

the set of all ordered p-tuples of elements of G whose product is the identity. In any such p-tuple, we are free to choose a_1, a_2, …, a_{p-1}, and then a_p is completely determined by

$$a_p = (a_1 a_2 \cdots a_{p-1})^{-1}.$$

It follows that

$$|S| = |G|^{p-1}.$$

Similarly, notice that if $(a_1, a_2, \dots, a_p) \in S$, and $2 \leq r \leq p$, then we have

$$\begin{aligned} (a_1 a_2 \cdots a_{r-1})(a_r \cdots a_p) = e &\Rightarrow (a_1 a_2 \cdots a_{r-1}) = (a_r \cdots a_p)^{-1} \\ &\Rightarrow (a_r \cdots a_p)(a_1 a_2 \cdots a_{r-1}) = e \\ &\Rightarrow (a_r, \dots, a_p, a_1, a_2, \dots, a_{r-1}) \in S, \end{aligned} \tag{1}$$

where in the second line we have used the fact that inverses in a group are always two-sided. This says we can cyclically shift the entries of tuples in S, which is essential in what follows.

The idea of the proof is to let $\overline{k} \in \mathbb{Z}/p\mathbb{Z}$ act on S by cyclicly permuting indices: $a_i \mapsto a_{i+k \pmod p}$. Although it eventually works out, checking that this completely natural operation is actually an action is quite tedious. So we opt for an equivalent, permutation-theoretic approach.

We first define an action of S_p on the larger set G^p, and then restrict it to a p-cyclic action on S. For $\sigma \in S_p$ and $\mathbf{a} = (a_1, \ldots, a_p) \in G^p$, set

$$\mathbf{a}^\sigma = (a_{\sigma(1)}, \ldots, a_{\sigma(p)}),$$

so that $\mathbf{a}^\sigma$ has $a_{\sigma(i)}$ in its ith entry. It is clear that Id acts trivially on any p-tuple. Given $\tau \in S_p$, to compute $(\mathbf{a}^\sigma)^\tau$, write

$$\mathbf{a}^\sigma = (a_{\sigma(1)}, \ldots, a_{\sigma(p)}) = (b_1, \ldots, b_p).$$

Then $b_i = a_{\sigma(i)}$ for all i so that

$$(\mathbf{a}^\sigma)^\tau = (b_1, \ldots, b_p)^\tau = (b_{\tau(1)}, \ldots, b_{\tau(p)}) = (a_{\sigma(\tau(1))}, \ldots, a_{\sigma(\tau(p))}) = (a_{\sigma\tau(1)}, \ldots, a_{\sigma\tau(p)}) = \mathbf{a}^{\sigma\tau}.$$

In other words, $(\sigma, \mathbf{a}) \mapsto \mathbf{a}^\sigma$ is a *right action* of S_p on G^p.[1]

Now let $\sigma = (1\,2\,3\cdots p)$ and $C = \langle\sigma\rangle$. Notice that if $(a_1, a_2, \ldots, a_p) \in S$, then

$$(a_1, a_2, \ldots, a_p)^{\sigma^k} = (a_{k+1}, a_{k+2} \ldots, a_p, a_1, \ldots, a_k) \in S,$$

according to (1). This shows that $S \subset G^p$ is invariant under C, or that C acts on S. Because $|C| = p$, the stabilizer $C_\mathbf{a}$ of any $\mathbf{a} \in S$ is either trivial or equal to C. In other words, for any $\mathbf{a} \in S$, $[C : C_\mathbf{a}] = p$, or $\mathbf{a}$ is a fixed point of C.

Because $|S| = |G|^{p-1} \equiv 0 \pmod{p}$, the class equation

$$|S| = |\mathcal{F}| + \sum_{\mathbf{a} \in \mathcal{O}^*} [C : C_\mathbf{a}]. \tag{2}$$

therefore implies that

$$|\mathcal{F}| \equiv 0 \pmod{p}. \tag{3}$$

But

$$\mathcal{F} = \{(a, a, \ldots, a) \,|\, a^p = e\} = \{(e, e, \ldots, e)\} \sqcup \{(a, a, \ldots, a) \,|\, a \in G, |a| = p\},$$

so that $|\mathcal{F}| = 1 + N_p$, where N_p is the number of elements of G of order p. The congruence (3) then becomes

$$N_p \equiv -1 \pmod{p},$$

which proves that $N_p \neq 0$ and completes the proof.

□

We state as a corollary the following stronger statement that was obtained during the course of the proof of Cauchy's theorem.

Corollary 1. *Let G be a finite group and let p be a prime number. If p divides $|G|$, and N_p denotes the number of elements of G of order p, then*

$$N_p \equiv -1 \pmod{p}.$$

Let G be a finite group whose order is divisible by a prime p. The elements of G of order p can be partitioned according to the subgroup that they generate. Suppose there are M_p subgroups of G of order p. Then G contains exactly $M_p(p-1)$ elements of order p. Comparing this with the statement of the corollary we find that

$$M_p(p-1) \equiv -M_p \equiv -1 \pmod{p} \;\Rightarrow\; M_p \equiv 1 \pmod{p}.$$

[1] Another way to verify the action axiom $(\mathbf{a}^\sigma)^\tau = \mathbf{a}^{\sigma\tau}$ is to view each tuple $\mathbf{a} \in G^p$ as a function $\mathbf{a} : I(p) \to G$. Our definition then becomes $\mathbf{a}^\sigma = \mathbf{a} \circ \sigma$, so that $\mathbf{a}^{\sigma\tau} = \mathbf{a} \circ (\sigma\tau) = (\mathbf{a} \circ \sigma) \circ \tau = (\mathbf{a}^\sigma)^\tau$.

2 Sylow's Theorems

Let p be a prime. A group G is called a *p-group* provided the order of every element of G is a power of p. For finite groups, this is equivalent to the statement that $|G|$ is a power of p. For if G is a finite p-group, and q is a prime dividing $|G|$, then Cauchy's theorem implies that G must have an element of order q. But q cannot be a power of p unless $q = p$. Therefore the only prime dividing $|G|$ is p. The converse of this result is an immediate consequence of Lagrange's theorem.

Given a finite group G and a prime p dividing $|G|$, we say that $P < G$ is a *p-Sylow subgroup* of G if $|P| = p^m$ and $p^{m+1} \nmid |G|$. Equivalently, P is a p-Sylow subgroup of G provided P is a p-group and $p \nmid [G : P]$. Although Cauchy's theorem only asserts the existence of p-subgroups of G of order p, it is actually equivalent to Sylow's first theorem on the existence of p-Sylow subgroups.

Theorem 2. *[Sylow's First Theorem] Let G be a finite group and p a prime dividing $|G|$. Then G has a p-Sylow subgroup.*

Proof. Write $|G| = pk$ with $k \in \mathbb{N}$. We induct on k, the case $k = 1$ being trivial. So assume $k \geq 2$ and that the result holds for all groups of order $p\ell$ with $\ell < k$. Let G act on itself by conjugation and consider the class equation

$$|G| = |Z(G)| + \sum_{x \in \mathcal{C}^*} [G : C_G(x)], \tag{4}$$

where $\mathcal{C}^*$ contains one element from every nontrivial conjugacy class. If there is an $x \in \mathcal{C}^*$ so that $p \nmid [G : C_G(x)]$, then $|C_G(x)| = p\ell$, with $\ell < k$. Apply the inductive hypothesis to obtain a p-Sylow subgroup P of $C_G(x)$. Since $p \nmid [C_G(x) : P]$, the assumption that $p \nmid [G : C_G(x)]$ implies $p \nmid [G : P]$. Since P is a p-group, it follows that P is a p-Sylow subgroup of G.

In the remaining case, we must have $p|[G : C_G(x)]$ for all $x \in \mathcal{C}^*$. The class equation (4) then implies that

$$|Z(G)| \equiv 0 \pmod{p}.$$

By Cauchy's theorem, $Z(G)$ therefore contains an element x of order p. Let $H = \langle x \rangle$. if $p \nmid [G : H]$, then H is a p-Sylow subgroup of G. Otherwise, being central, H is normal in G, and G/H is a group of order $p\ell$, with $\ell < k$. By the inductive hypothesis, G/H has a p-Sylow subgroup P_0, and by the correspondence principle $P_0 = P/H$ for some subgroup P of G containing H. Notice that

$$|P| = [P : H]\,|H| = p\,|P_0|,$$

so that P is a p-group. Since $p \nmid [G/H : P_0] = [G/H : P/H] = [G : P]$, we find that P is a p-Sylow subgroup of G. This completes the induction and the proof.

□

Notice that in the proof, Cauchy's theorem was necessary in the second case to ensure that the pullback of P_0 through the quotient map was indeed a p-group. It's tempting to simply take $H = Z(G)$, but because we can't control the other prime factors of $|Z(G)|$, we would be unable to ensure that the subgroup P is genuinely a p-group. We might try to work around this possibility by inductively taking a Sylow p-subgroup of P instead, but it's entirely possible that $P = G$, which would preclude the application of the inductive hypothesis.

Prior to the statement of Sylow's first theorem, we mentioned that it is, in fact, equivalent to Cauchy's theorem. We have just seen that the former is certainly a consequence of the latter. As for the converse, assume Theorem 2 and let G be a finite group whose order is divisible by a prime p. Let $P < G$ be a p-Sylow subgroup of G. Because P is a p-group, $Z(P)$ is nontrivial. The classification of abelian p-groups then implies that $Z(P)$, and hence G, has an element of order p, proving Cauchy's theorem.

An important result in the theory of p-Sylow subgroups is that they are maximal (relative to containment) among the p-subgroups of a finite group. Important in establishing this fact, as well as a number of others, is the following lemma.

Lemma 1. *Let G be a finite group and p a prime dividing $|G|$. If H is a p-subgroup of G and P is a p-Sylow subgroup of G, then H normalizes P if and only if $H < P$.*

Proof. If H normalizes P, then H and P are subgroups of $N_G(P)$, the latter being normal. By the second isomorphism theorem,

$$HP/P \cong H/H \cap P.$$

Thus $|HP| = [HP:P]|P| = [H:H\cap P]|P|$ is a power of p dividing $|G|$. Because $|P|$ is the largest such power $|G|$, this implies that $[H:H\cap P] = 1$ and hence $H = H \cap P < P$. □

Let G be a finite group, p a prime dividing $|G|$, and P an arbitrary p-Sylow subgroup of G. The set S of G-conjugates of P has size

$$|S| = [G:N_G(P)],$$

which divides $[G:P]$ and is therefore relatively prime to p. Let $H < G$ be any p-subgroup. If we let H act on S by conjugation, the fixed points $\mathcal{F}$ are the conjugates of P normalized by H. By Lemma 1, these are precisely the conjugates of P containing H. Because H is a p-group, if we consider the class equation mod p we obtain

$$|S| \equiv |\mathcal{F}| \pmod{p}.$$

Since $p \nmid |S|$, this implies that $\mathcal{F} \neq \varnothing$. That is, there is a conjugate of P containing H. This proves the following result.

Theorem 3. *Let G be a finite group, p a prime dividing $|G|$ and $P < G$ a p-Sylow subgroup. If $H < G$ is a p-subgroup, then there is a conjugate of P containing H, and the number of such conjugates is congruent to $[G:N_G(P)] \pmod{p}$. In particular, every p-subgroup of G is contained in a p-Sylow subgroup of G. The p-Sylow subgroups are therefore the maximal p-subgroups of G.*

Theorem 4 (Sylow's Second and Third Theorems)**.** *Let G be a finite group, p a prime dividing $|G|$ and P a p-Sylow subgroup. Write $|G| = p^m k$ with $p \nmid k$. Then:*

1. *All p-Sylow subgroups are conjugate. In particular, a p-Sylow subgroup is normal in G if and only if it is unique.*
2. *The number of p-Sylow subgroups of G is $[G:N_G(P)]$, which divides k and is $\equiv 1 \pmod{p}$.*

Proof. In Theorem 3, take H to be any p-Sylow subgroup of G. Order considerations show that any conjugate of P containing H must, in fact, equal H. This proves part 1. It follows that:

- $[G:N_G(P)]$, which counts the conjugates of P, actually counts all of the p-Sylow subgroups of G;
- $[G:N_G(P)] \equiv 1 \pmod{p}$, since there can be no more that one conjugate of P equal to H.

To complete the proof, it only remains to observe that $[G:N_G(P)]$ divides $[G:P] = k$

□

Example. A beautiful application of Sylow's theorems is to groups of order pq, where $p < q$ are primes. We claim that if $q \not\equiv 1 \pmod{p}$, then any group of order pq is cyclic.

To see this, let G have order pq with p, q as above. Let Q be a q-Sylow subgroup of G. Since $[G:Q] = p$ is the smallest prime divisor of $|G|$, we have $Q \lhd G$. Now let P be a p-Sylow subgroup of G. If P is not normal in G, then the number of p-Sylow subgroups of G must be q. But $q \not\equiv 1 \pmod{p}$, so this is impossible. Hence

$P \lhd G$ as well. Since $(|P|, |Q|) = (p, q) = 1$, $P \cap Q = \{e\}$. This means that P and Q commute element-wise, and so $G = PQ \cong P \times Q$. This establishes the claim.

In the interest of completeness, there's one more thing to be said about the p-subgroups of a finite group G. It can be viewed as a generalization of Cauchy's theorem to prime power subgroups. We first state a special case, then apply it to the general setting.

Lemma 2. *Let G be a finite p-group of order p^n. There exists a chain of subgroups*

$$H_1 < H_2 < \cdots < H_n = G,$$

with $|H_i| = p^i$ for all i.

Proof. We induct on n; when $n = 1$ there is nothing to prove. So assume that G is a p-group of order p^n, $n > 1$, and that we have proven the result for all finite p-groups of order less than p^n. We have proven elsewhere (using the class equation!) that because G is a p-group, its center $Z(G)$ is nontrivial. As in the proof of the first Sylow theorem, let $x \in Z(G)$ have order p, so that $H = \langle x \rangle$ is a normal subgroup of G of order p.

The inductive hypothesis applies to G/H, yielding subgroups

$$\overline{H_2} < \overline{H_3} < \cdots < \overline{H_n} = G/H,$$

with $|\overline{H_j}| = p^{j-1}$ for all j. By the correspondence principle, for each j there is a unique subgroup $H_j < G$ containing H so that

$$\overline{H_j} = H_j/H.$$

We then have

$$H = H_1 < H_2 < H_3 < \cdots < H_n = G,$$

and

$$|H_j| = [H_j : H]\,|H| = |\overline{H_j}|\,p = p^j$$

for all $j \geq 1$, which prove the result when $|G| = p^n$. By induction, it holds for all finite p-groups. □

The use of a cyclic subgroup of $Z(G)$ in this proof was more for convenience than out of necessity. We could simply have taken $H = Z(G)$ and attempted to induct on $G/Z(G)$ and $Z(G)$. The drawback is that if G is abelian, then $G = Z(G)$, and induction doesn't apply. This situation can be dealt with directly, but rather than break the proof into multiple cases we opted to use a smaller subgroup which ensured that the quotient G/H wasn't trivial.

Our final result about p-subgroups of finite groups is a simple consequence of Lemma 2.

Theorem 5. *Let G be a finite group and p a prime dividing $|G|$. If P is a p-Sylow subgroup of G, of order p^n, then there exists a chain*

$$H_1 < H_2 < \cdots < H_n = P,$$

where $|H_i| = p^i$ for all $i \geq 1$.

Proof. Apply Lemma 2 to P. □