

Sylow Theory

1A

It seems appropriate to begin this book with a topic that underlies virtually all of finite group theory: the Sylow theorems. In this chapter, we state and prove these theorems, and we present some applications and related results. Although much of this material should be very familiar, we suspect that most readers will find that at least some of the content of this chapter is new to them.

Although the theorem that proves Sylow subgroups always exist dates back to 1872, the existence proof that we have decided to present is that of H. Wielandt, published in 1959. Wielandt's proof is slick and short, but it does have some drawbacks. It is based on a trick that seems to have no other application, and the proof is not really constructive; it gives no guidance about how, in practice, one might actually find a Sylow subgroup. But Wielandt's proof is beautiful, and that is the principal motivation for presenting it here.

Also, Wielandt's proof gives us an excuse to present a quick review of the theory of group actions, which are nearly as ubiquitous in the study of finite groups as are the Sylow theorems themselves. We devote the rest of this section to the relevant definitions and basic facts about actions, although we omit some details from the proofs.

Let G be a group, and let Ω be a nonempty set. (We will often refer to the elements of Ω as "points".) Suppose we have a rule that determines a new element of Ω , denoted $\alpha \cdot g$, whenever we are given a point $\alpha \in \Omega$ and an element $g \in G$. We say that this rule defines an **action** of G on Ω if the following two conditions hold.

- (1) $\alpha \cdot 1 = \alpha$ for all $\alpha \in \Omega$ and
- (2) $(\alpha \cdot g) \cdot h = \alpha \cdot (gh)$ for all $\alpha \in \Omega$ and all group elements $g, h \in G$.

Suppose that G acts on Ω . It is easy to see that if $g \in G$ is arbitrary, then the function $\sigma_g : \Omega \rightarrow \Omega$ defined by $(\alpha)\sigma_g = \alpha \cdot g$ has an inverse: the function $\sigma_{g^{-1}}$. Therefore, σ_g is a permutation of the set Ω , which means that σ_g is both injective and surjective, and thus σ_g lies in the symmetric group $\text{Sym}(\Omega)$ consisting of all permutations of Ω . In fact, the map $g \mapsto \sigma_g$ is easily seen to be a homomorphism from G into $\text{Sym}(\Omega)$. (A homomorphism like this, which arises from an action of a group G on some set, is called a **permutation representation** of G .) The kernel of this homomorphism is, of course, a normal subgroup of G , which is referred to as the **kernel** of the action. The kernel is exactly the set of elements $g \in G$ that act trivially on Ω , which means that $\alpha \cdot g = \alpha$ for all points $\alpha \in \Omega$.

Generally, we consider a theorem or a technique that has the power to find a normal subgroup of G to be “good”, and indeed permutation representations can be good in this sense. (See the problems at the end of this section.) But our goal in introducing group actions here is not to find normal subgroups; it is to count things. Before we proceed in that direction, however, it seems appropriate to mention a few examples.

Let G be arbitrary, and take $\Omega = G$. We can let G act on G by right multiplication, so that $x \cdot g = xg$ for $x, g \in G$. This is the **regular** action of G , and it should be clear that it is **faithful**, which means that its kernel is trivial. It follows that the corresponding permutation representation of G is an isomorphism of G into $\text{Sym}(G)$, and this proves Cayley’s theorem: every group is isomorphic to a group of permutations on some set.

We continue to take $\Omega = G$, but this time, we define $x \cdot g = g^{-1}xg$. (The standard notation for $g^{-1}xg$ is x^g .) It is trivial to check that $x^1 = x$ and that $(x^g)^h = x^{gh}$ for all $x, g, h \in G$, and thus we truly have an action, which is called the **conjugation** action of G on itself. Note that $x^g = x$ if and only if $xg = gx$, and thus the kernel of the conjugation action is the set of elements $g \in G$ that commute with all elements $x \in G$. The kernel, therefore, is the center $\mathbf{Z}(G)$.

Again let G be arbitrary. In each of the previous examples, we took $\Omega = G$, but we also get interesting actions if instead we take Ω to be the set of all subsets of G . In the conjugation action of G on Ω we let $X \cdot g = X^g = \{x^g \mid x \in X\}$ and in the right-multiplication action we define $X \cdot g = Xg = \{xg \mid x \in X\}$. Of course, in order to make these examples work, we do not really need Ω to be *all* subsets of G . For example, since a conjugate of a subgroup is always a subgroup, the conjugation action is well defined if we take Ω to be the set of all subgroups of G . Also, both right multiplication

and conjugation preserve cardinality, and so each of these actions makes sense if we take Ω to be the collection of all subsets of G of some fixed size. In fact, as we shall see, the trick in Wielandt's proof of the Sylow existence theorem is to use the right multiplication action of G on its set of subsets with a certain fixed cardinality.

We mention one other example, which is a special case of the right-multiplication action on subsets that we discussed in the previous paragraph. Let $H \subseteq G$ be a subgroup, and let $\Omega = \{Hx \mid x \in G\}$, the set of right cosets of H in G . If X is any right coset of H , it is easy to see that Xg is also a right coset of H . (Indeed, if $X = Hx$, then $Xg = H(xg)$.) Then G acts on the set Ω by right multiplication.

In general, if a group G acts on some set Ω and $\alpha \in \Omega$, we write $G_\alpha = \{g \in G \mid \alpha \cdot g = \alpha\}$. It is easy to check that G_α is a subgroup of G ; it is called the **stabilizer** of the point α . For example, in the regular action of G on itself, the stabilizer of every point (element of G) is the trivial subgroup. In the conjugation action of G on G , the stabilizer of $x \in G$ is the centralizer $\mathbf{C}_G(x)$ and in the conjugation action of G on subsets, the stabilizer of a subset X is the normalizer $\mathbf{N}_G(X)$. A useful general fact about point stabilizers is the following, which is easy to prove. In any action, if $\alpha \cdot g = \beta$, then the stabilizers G_α and G_β are conjugate in G , and in fact, $(G_\alpha)^g = G_\beta$.

Now consider the action (by right multiplication) of G on the right cosets of H , where $H \subseteq G$ is a subgroup. The stabilizer of the coset Hx is the set of all group elements g such that $Hxg = Hx$. It is easy to see that g satisfies this condition if and only if $xg \in Hx$. (This is because two cosets Hu and Hv are identical if and only if $u \in Hv$.) It follows that g stabilizes Hx if and only if $g \in x^{-1}Hx$. Since $x^{-1}Hx = H^x$, we see that the stabilizer of the point (coset) Hx is exactly the subgroup H^x , conjugate to H via x . It follows that the kernel of the action of G on the right cosets of H in G is exactly $\bigcap_{x \in G} H^x$. This subgroup is called the **core** of H in G , denoted $\text{core}_G(H)$. The core of H is normal in G because it is the kernel of an action, and, clearly, it is contained in H . In fact, if $N \triangleleft G$ is any normal subgroup that happens to be contained in H , then $N = N^x \subseteq H^x$ for all $x \in G$, and thus $N \subseteq \text{core}_G(H)$. In other words, the core of H in G is the unique largest normal subgroup of G contained in H . (It is "largest" in the strong sense that it contains all others.)

We have digressed from our goal, which is to show how to use group actions to count things. But having come this far, we may as well state the results that our discussion has essentially proved. Note that the following theorem and its corollaries can be used to prove the existence of normal subgroups, and so they might be considered to be "good" results.

1.1. Theorem. *Let $H \subseteq G$ be a subgroup, and let Ω be the set of right cosets of H in G . Then $G/\text{core}_G(H)$ is isomorphic to a subgroup of $\text{Sym}(\Omega)$. In particular, if the index $|G : H| = n$, then $G/\text{core}_G(H)$ is isomorphic to a subgroup of S_n , the symmetric group on n symbols.*

Proof. The action of G on the set Ω by right multiplication defines a homomorphism θ (the permutation representation) from G into $\text{Sym}(\Omega)$. Since $\ker(\theta) = \text{core}_G(H)$, it follows by the homomorphism theorem that $G/\text{core}_G(H) \cong \theta(G)$, which is a subgroup of $\text{Sym}(\Omega)$. The last statement follows since if $|G : H| = n$, then (by definition of the index) $|\Omega| = n$, and thus $\text{Sym}(\Omega) \cong S_n$. ■

1.2. Corollary. *Let G be a group, and suppose that $H \subseteq G$ is a subgroup with $|G : H| = n$. Then H contains a normal subgroup N of G such that $|G : N|$ divides $n!$.*

Proof. Take $N = \text{core}_G(H)$. Then G/N is isomorphic to a subgroup of the symmetric group S_n , and so by Lagrange's theorem, $|G/N|$ divides $|S_n| = n!$. ■

1.3. Corollary. *Let G be simple and contain a subgroup of index $n > 1$. Then $|G|$ divides $n!$.*

Proof. The normal subgroup N of the previous corollary is contained in H , and hence it is proper in G because $n > 1$. Since G is simple, $N = 1$, and thus $|G| = |G/N|$ divides $n!$. ■

In order to pursue our main goal, which is counting, we need to discuss the “orbits” of an action. Suppose that G acts on Ω , and let $\alpha \in \Omega$. The set $\mathcal{O}_\alpha = \{\alpha \cdot g \mid g \in G\}$ is called the **orbit** of α under the given action. It is routine to check that if $\beta \in \mathcal{O}_\alpha$, then $\mathcal{O}_\beta = \mathcal{O}_\alpha$, and it follows that distinct orbits are actually disjoint. Also, since every point is in at least one orbit, it follows that the orbits of the action of G on Ω partition Ω . In particular, if Ω is finite, we see that $|\Omega| = \sum |\mathcal{O}|$, where in this sum, \mathcal{O} runs over the full set of G -orbits on Ω .

We mention some examples of orbits and orbit decompositions. First, if $H \subseteq G$ is a subgroup, we can let H act on G by right multiplication. It is easy to see that the orbits of this action are exactly the left cosets of H in G . (We leave to the reader the problem of realizing the right cosets of H in G as the orbits of an appropriate action of H . But be careful: the rule $x \cdot h = hx$ does *not* define an action.)

Perhaps it is more interesting to consider the conjugation action of G on itself, where the orbits are exactly the conjugacy classes of G . The fact

that for a finite group, the order $|G|$ is the sum of the sizes of the classes is sometimes called the **class equation** of G .

How big is an orbit? The key result here is the following.

1.4. Theorem (The Fundamental Counting Principle). *Let G act on Ω , and suppose that \mathcal{O} is one of the orbits. Let $\alpha \in \mathcal{O}$, and write $H = G_\alpha$, the stabilizer of α . Let $\Lambda = \{Hx \mid x \in G\}$ be the set of right cosets of H in G . Then there is a bijection $\theta : \Lambda \rightarrow \mathcal{O}$ such that $\theta(Hg) = \alpha \cdot g$. In particular, $|\mathcal{O}| = |G : G_\alpha|$.*

Proof. We observe first that if $Hx = Hy$, then $\alpha \cdot x = \alpha \cdot y$. To see why this is so, observe that we can write $y = hx$ for some element $h \in H$. Then

$$\alpha \cdot y = \alpha \cdot (hx) = (\alpha \cdot h) \cdot x = \alpha \cdot x,$$

where the last equality holds because $h \in H = G_\alpha$, and so h stabilizes α .

Given a coset $Hx \in \Lambda$, the point $\alpha \cdot x$ lies in \mathcal{O} , and we know that it is determined by the coset Hx , and not just by the particular element x . It is therefore permissible to define the function $\theta : \Lambda \rightarrow \mathcal{O}$ by $\theta(Hx) = \alpha \cdot x$, and it remains to show that θ is both injective and surjective.

The surjectivity is easy, and we do that first. If $\beta \in \mathcal{O}$, then by the definition of an orbit, we have $\beta = \alpha \cdot x$ for some element $x \in G$. Then $Hx \in \Lambda$ satisfies $\theta(Hx) = \alpha \cdot x = \beta$, as required.

To prove that θ is injective, suppose that $\theta(Hx) = \theta(Hy)$. We have $\alpha \cdot x = \alpha \cdot y$, and hence

$$\alpha = \alpha \cdot 1 = (\alpha \cdot x) \cdot x^{-1} = (\alpha \cdot y) \cdot x^{-1} = \alpha \cdot (yx^{-1}).$$

Then yx^{-1} fixes α , and so it lies in $G_\alpha = H$. It follows that $y \in Hx$, and thus $Hx = Hy$. This proves that θ is injective, as required. ■

It is easy to check that the bijection θ of the previous theorem actually defines a “permutation isomorphism” between the action of G on Λ and the action of G on the orbit \mathcal{O} . Formally, this means that $\theta(X \cdot g) = \theta(X) \cdot g$ for all “points” X in Λ and group elements $g \in G$. More informally, this says that the actions of G on Λ and on \mathcal{O} are “essentially the same”. Since every action can be thought of as composed of the actions on the individual orbits, and each of these actions is permutation isomorphic to the right-multiplication action of G on the right cosets of some subgroup, we see that these actions on cosets are truly fundamental: every group action can be viewed as being composed of actions on right cosets of various subgroups.

We close this section with two familiar and useful applications of the fundamental counting principle.

1.5. Corollary. Let $x \in G$, where G is a finite group, and let K be the conjugacy class of G containing x . Then $|K| = |G : \mathbf{C}_G(x)|$.

Proof. The class of x is the orbit of x under the conjugation action of G on itself, and the stabilizer of x in this action is the centralizer $\mathbf{C}_G(x)$. Thus $|K| = |G : \mathbf{C}_G(x)|$, as required. ■

1.6. Corollary. Let $H \subseteq G$ be a subgroup, where G is finite. Then the total number of distinct conjugates of H in G , counting H itself, is $|G : \mathbf{N}_G(H)|$.

Proof. The conjugates of H form an orbit under the conjugation action of G on the set of subsets of G . The normalizer $\mathbf{N}_G(H)$ is the stabilizer of H in this action, and thus the orbit size is $|G : \mathbf{N}_G(H)|$, as wanted. ■

Problems 1A

1A.1. Let H be a subgroup of prime index p in the finite group G , and suppose that no prime smaller than p divides $|G|$. Prove that $H \triangleleft G$.

1A.2. Given subgroups $H, K \subseteq G$ and an element $g \in G$, the set $HgK = \{h g k \mid h \in H, k \in K\}$ is called an (H, K) -**double coset**. In the case where H and K are finite, show that $|HgK| = |H||K|/|K \cap H^g|$.

Hint. Observe that HgK is a union of right cosets of H , and that these cosets form an orbit under the action of K .

Note. If we take $g = 1$ in this problem, the result is the familiar formula $|HK| = |H||K|/|H \cap K|$.

1A.3. Suppose that G is finite and that $H, K \subseteq G$ are subgroups.

(a) Show that $|H : H \cap K| \leq |G : K|$, with equality if and only if $HK = G$.

(b) If $|G : H|$ and $|G : K|$ are coprime, show that $HK = G$.

Note. Proofs of these useful facts appear in the appendix, but we suggest that readers try to find their own arguments. Also, recall that the product HK of subgroups H and K is not always a subgroup. In fact, HK is a subgroup if and only if $HK = KH$. (This too is proved in the appendix.) If $HK = KH$, we say that H and K are **permutable**.

1A.4. Suppose that $G = HK$, where H and K are subgroups. Show that also $G = H^x K^y$ for all elements $x, y \in G$. Deduce that if $G = HH^x$ for a subgroup H and an element $x \in G$, then $H = G$.

1A.5. An action of a group G on a set Ω is **transitive** if Ω consists of a single orbit. Equivalently, G is transitive on Ω if for every choice of points $\alpha, \beta \in \Omega$, there exists an element $g \in G$ such that $\alpha \cdot g = \beta$. Now assume that a group G acts transitively on each of two sets Ω and Λ . Prove that the natural induced action of G on the cartesian product $\Omega \times \Lambda$ is transitive if and only if $G_\alpha G_\beta = G$ for some choice of $\alpha \in \Omega$ and $\beta \in \Lambda$.

Hint. Show that if $G_\alpha G_\beta = G$ for some $\alpha \in \Omega$ and $\beta \in \Lambda$, then in fact, this holds for all $\alpha \in \Omega$ and $\beta \in \Lambda$.

1A.6. Let G act on Ω , where both G and Ω are finite. For each element $g \in G$, write $\chi(g) = |\{\alpha \in \Omega \mid \alpha \cdot g = \alpha\}|$. The nonnegative-integer-valued function χ is called the **permutation character** associated with the action. Show that

$$\sum_{g \in G} \chi(g) = \sum_{\alpha \in \Omega} |G_\alpha| = n|G|,$$

where n is the number of orbits of G on Ω .

Note. Thus the number of orbits is

$$n = \frac{1}{|G|} \sum_{g \in G} \chi(g),$$

which is the average value of χ over the group. Although this orbit-counting formula is often attributed to W. Burnside, it should (according to P. Neumann) more properly be credited to Cauchy and Frobenius.

1A.7. Let G be a finite group, and suppose that $H < G$ is a proper subgroup. Show that the number of elements of G that do not lie in any conjugate of H is at least $|H|$.

Hint. Let χ be the permutation character associated with the right-multiplication action of G on the right cosets of H . Then $\sum \chi(g) = |G|$, where the sum runs over $g \in G$. Show that $\sum \chi(h) \geq 2|H|$, where here, the sum runs over $h \in H$. Use this information to get an estimate on the number of elements of G where χ vanishes.

1A.8. Let G be a finite group, let $n > 0$ be an integer, and let C be the additive group of the integers modulo n . Let Ω be the set of n -tuples (x_1, x_2, \dots, x_n) of elements of G such that $x_1 x_2 \cdots x_n = 1$.

(a) Show that C acts on Ω according to the formula

$$(x_1, x_2, \dots, x_n) \cdot k = (x_{1+k}, x_{2+k}, \dots, x_{n+k}),$$

where $k \in C$ and the subscripts are interpreted modulo n .

- (b) Now suppose that $n = p$ is a prime number that divides $|G|$. Show that p divides the number of C -orbits of size 1 on Ω , and deduce that the number of elements of order p in G is congruent to $-1 \pmod{p}$.

Note. In particular, if a prime p divides $|G|$, then G has at least one element of order p . This is a theorem of Cauchy, and the proof in this problem is due to J. H. McKay. Cauchy's theorem can also be derived as a corollary of Sylow's theorem. Alternatively, a proof of Sylow's theorem different from Wielandt's can be based on Cauchy's theorem. (See Problem 1B.4.)

1A.9. Suppose $|G| = pm$, where $p > m$ and p is prime. Show that G has a unique subgroup of order p .

1A.10. Let $H \subseteq G$.

- (a) Show that $|\mathbf{N}_G(H) : H|$ is equal to the number of right cosets of H in G that are invariant under right multiplication by H .
- (b) Suppose that $|H|$ is a power of the prime p and that $|G : H|$ is divisible by p . Show that $|\mathbf{N}_G(H) : H|$ is divisible by p .

1B

Fix a prime number p . A finite group whose order is a power of p is called a **p -group**. It is often convenient, however, to use this nomenclature somewhat carelessly, and to refer to a group as a " p -group" even if there is no particular prime p under consideration. For example, in proving some theorem, one might say: it suffices to check that the result holds for p -groups. What is meant here, of course, is that it suffices to show that the theorem holds for all p -groups for all primes p .

We mention that, although in this book a p -group is required to be finite, it is also possible to define infinite p -groups. The more general definition is that a (not necessarily finite) group G is a p -group if every element of G has finite p -power order. Of course, if G is finite, then by Lagrange's theorem, every element of G has order dividing $|G|$, and so if $|G|$ is a power of p , it follows that the order of every element is a power of p , and hence G is a p -group according to the more general definition. Conversely, if G is finite and has the property that the order of every element is a power of p , then clearly, G can have no element of order q for any prime q different from p . It follows by Cauchy's theorem (Problem 1A.8) that no prime $q \neq p$ can divide $|G|$, and thus $|G|$ must be a power of p , and this shows that the two definitions of " p -group" are equivalent for finite groups.

Again, fix a prime p . A subgroup S of a finite group G is said to be a **Sylow p -subgroup** of G if $|S|$ is a power of p and the index $|G : S|$ is

not divisible by p . An alternative formulation of this definition relies on the observation that every positive integer can be (uniquely) factored as a power of the given prime p times some integer not divisible by p . In particular, if we write $|G| = p^a m$, where $a \geq 0$ and p does not divide $m \geq 1$, then a subgroup S of G is a Sylow p -subgroup of G precisely when $|S| = p^a$. In other words, a Sylow p -subgroup of G is a p -subgroup S whose order is as large as is permitted by Lagrange's theorem, which requires that $|S|$ must divide $|G|$. We mention two trivial cases: if $|G|$ is not divisible by p , then the identity subgroup is a Sylow p -subgroup of G , and if G is a p -group, then G is a Sylow p -subgroup of itself. The Sylow existence theorem asserts that Sylow subgroups *always* exist.

1.7. Theorem (Sylow E). *Let G be a finite group, and let p be a prime. Then G has a Sylow p -subgroup.*

The Sylow E-theorem can be viewed as a partial converse of Lagrange's theorem. Lagrange asserts that if H is a subgroup of G and $|H| = k$, then k divides $|G|$. The converse, which in general is false, would say that if k is a positive integer that divides $|G|$, then G has a subgroup of order k . (The smallest example of the failure of this assertion is to take G to be the alternating group A_4 of order 12; this group has no subgroup of order 6.) But if k is a power of a prime, we shall see that G actually does have a subgroup of order k . If k is the largest power of p that divides $|G|$, the desired subgroup of order k is a Sylow p -subgroup; for smaller powers of p , we will prove that a Sylow p -subgroup of G necessarily has a subgroup of order k .

We are ready now to begin work toward the proof of the Sylow E-theorem. We start with a purely arithmetic fact about binomial coefficients.

1.8. Lemma. *Let p be a prime number, and let $a \geq 0$ and $m \geq 1$ be integers. Then*

$$\binom{p^a m}{p^a} \equiv m \pmod{p}.$$

Proof. Consider the polynomial $(1 + X)^p$. Since p is prime, it is easy to see that the binomial coefficients $\binom{p}{i}$ are divisible by p for $1 \leq i \leq p - 1$, and thus we can write $(1 + X)^p \equiv 1 + X^p \pmod{p}$. (The assertion that these polynomials are congruent modulo p means that the coefficients of corresponding powers of X are congruent modulo p .) Applying this fact a second time, we see that $(1 + X)^{p^2} \equiv (1 + X^p)^p \equiv 1 + X^{p^2} \pmod{p}$. Continuing like this, we deduce that $(1 + X)^{p^a} \equiv 1 + X^{p^a} \pmod{p}$, and thus

$$(1 + X)^{p^a m} \equiv (1 + X^{p^a})^m \pmod{p}.$$

Since these polynomials are congruent, the coefficients of corresponding terms are congruent modulo p , and the result follows by considering the coefficient of X^{p^a} on each side. ■

Proof of the Sylow E-theorem (Wielandt). Write $|G| = p^a m$, where $a \geq 0$ and p does not divide m . Let Ω be the set of all subsets of G having cardinality p^a , and observe that G acts by right multiplication on Ω . Because of this action, Ω is partitioned into orbits, and consequently, $|\Omega|$ is the sum of the orbit sizes. But

$$|\Omega| = \binom{p^a m}{p^a} \equiv m \not\equiv 0 \pmod{p},$$

and so $|\Omega|$ is not divisible by p , and it follows that there is some orbit \mathcal{O} such that $|\mathcal{O}|$ is not divisible by p .

Now let $X \in \mathcal{O}$, and let $H = G_X$ be the stabilizer of X in G . By the fundamental counting principle, $|\mathcal{O}| = |G|/|H|$, and since p does not divide $|\mathcal{O}|$ and p^a divides $|G|$, we conclude that p^a must divide $|H|$, and in particular $p^a \leq |H|$.

Since H stabilizes X under right multiplication, we see that if $x \in X$, then $xH \subseteq X$, and thus $|H| = |xH| \leq |X| = p^a$, where the final equality holds since $X \in \Omega$. We now have $|H| = p^a$, and since H is a subgroup, it is a Sylow subgroup of G , as wanted. ■

In Problem 1A.8, we sketched a proof of Cauchy's theorem. We can now give another proof, using the Sylow E-theorem.

1.9. Corollary (Cauchy). *Let G be a finite group, and suppose that p is a prime divisor of $|G|$. Then G has an element of order p .*

Proof. Let S be a Sylow p -subgroup of G , and note that since $|S|$ is the maximum power of p that divides $|G|$, we have $|S| > 1$. Choose a non-identity element x of S , and observe that the order $o(x)$ divides $|S|$ by Lagrange's theorem, and thus $1 < o(x)$ is a power of p . In particular, we can write $o(x) = pm$ for some integer $m \geq 1$, and we see that $o(x^m) = p$, as wanted. ■

We introduce the notation $\text{Syl}_p(G)$ to denote the set of all Sylow p -subgroups of G . The assertion of the Sylow E-theorem, therefore, is that the set $\text{Syl}_p(G)$ is nonempty for all finite groups G and all primes p . The intersection $\bigcap \text{Syl}_p(G)$ of all Sylow p -subgroups of a group G is denoted $\mathbf{O}_p(G)$, and as we shall see, this is a subgroup that plays an important role in finite group theory.

Perhaps this is a good place to digress to review some basic facts about characteristic subgroups. (Some of this material also appears in the appendix.) First, we recall the definition: a subgroup $K \subseteq G$ is **characteristic** in G if every automorphism of G maps K onto itself.

It is often difficult to find all automorphisms of a given group, and so the definition of “characteristic” can be hard to apply directly, but nevertheless, in many cases, it is easy to establish that certain subgroups are characteristic. For example, the center $\mathbf{Z}(G)$, the derived (or commutator) subgroup G' , and the intersection of all Sylow p -subgroups $\mathbf{O}_p(G)$ are characteristic in G . More generally, any subgroup that can be described unambiguously as “*the* something” is characteristic. It is essential that the description using the definite article be unambiguous, however. Given a subgroup $H \subseteq G$, for example, we cannot conclude that the normalizer $\mathbf{N}_G(H)$ or the center $\mathbf{Z}(H)$ is characteristic in G . Although these subgroups are described using “the”, the descriptions are not unambiguous because they depend on the choice of H . We can say, however, that $\mathbf{Z}(G')$ is characteristic in G because it is *the* center of *the* derived subgroup; it does not depend on any unspecified subgroups.

A good way to see why “the something” subgroups must be characteristic is to imagine two groups G_1 and G_2 , with an isomorphism $\theta : G_1 \rightarrow G_2$. Since isomorphisms preserve “group theoretic” properties, it should be clear that θ maps the center $\mathbf{Z}(G_1)$ onto $\mathbf{Z}(G_2)$, and indeed θ maps each unambiguously defined subgroup of G_1 onto the corresponding subgroup of G_2 . Now specialize to the case where G_1 and G_2 happen to be the same group G , so θ is an automorphism of G . Since in the general case, we know that $\theta(\mathbf{Z}(G_1)) = \mathbf{Z}(G_2)$, we see that when $G_1 = G = G_2$, we have $\theta(\mathbf{Z}(G)) = \mathbf{Z}(G)$, and similarly, if we consider any “the something” subgroup in place of the center.

Of course, characteristic subgroups are automatically normal. This is because the definition of normality requires only that the subgroup be mapped onto itself by *inner* automorphisms while characteristic subgroups are mapped onto themselves by all automorphisms. We have seen that some characteristic subgroups are easily recognized, and it follows that these subgroups are obviously and automatically normal. For example, the subgroup $\mathbf{O}_p(G)$ is normal in G for all primes p .

The fact that characteristic subgroups are normal remains true in an even more general context. The following, which we presume is already known to most readers of this book, is extremely useful. (This result also appears in the appendix.)

1.10. Lemma. *Let $K \subseteq N \subseteq G$, where G is a group, N is a normal subgroup of G and K is a characteristic subgroup of N . Then $K \triangleleft G$.*

Proof. Let $g \in G$. Then conjugation by g maps N onto itself, and it follows that the restriction of this conjugation map to N is an automorphism of N . (But note that it is not necessarily an inner automorphism of N .) Since K is characteristic in N , it is mapped onto itself by this automorphism of N , and thus $K^g = K$, and it follows that $K \triangleleft G$. ■

Problems 1B

1B.1. Let $S \in \text{Syl}_p(G)$, where G is a finite group.

- (a) Let $P \subseteq G$ be a p -subgroup. Show that PS is a subgroup if and only if $P \subseteq S$.
- (b) If $S \triangleleft G$, show that $\text{Syl}_p(G) = \{S\}$, and deduce that S is characteristic in G .

Note. Of course, it would be “cheating” to do problems in this section using theory that we have not yet developed. In particular, you should avoid using the Sylow C-theorem, which asserts that every two Sylow p -subgroups of G are conjugate in G .

1B.2. Show that $\mathbf{O}_p(G)$ is the unique largest normal p -subgroup of G . (This means that it is a normal p -subgroup of G that contains every other normal p -subgroup of G .)

1B.3. Let $S \in \text{Syl}_p(G)$, and write $N = \mathbf{N}_G(S)$. Show that $N = \mathbf{N}_G(N)$.

1B.4. Let $P \subseteq G$ be a p -subgroup such that $|G : P|$ is divisible by p . Using Cauchy’s theorem, but without appealing to Sylow’s theorem, show that there exists a subgroup Q of G containing P , and such that $|Q : P| = p$. Deduce that a maximal p -subgroup of G (which obviously must exist) must be a Sylow p -subgroup of G .

Hint. Use Problem 1A.10 and consider the group $\mathbf{N}_G(P)/P$.

Note. Once we know Cauchy’s theorem, this problem yields an alternative proof of the Sylow E-theorem. Of course, to avoid circularity, we appeal to Problem 1A.8 for Cauchy’s theorem, and not to Corollary 1.9.

1B.5. Let π be any set of prime numbers. We say that a finite group H is a π -**group** if every prime divisor of $|H|$ lies in π . Also, a π -subgroup $H \subseteq G$ is a **Hall** π -subgroup of G if no prime dividing the index $|G : H|$ lies in π . (So if $\pi = \{p\}$, a Hall π -subgroup is exactly a Sylow p -subgroup.)

Now let $\theta : G \rightarrow K$ be a surjective homomorphism of finite groups.

- (a) If H is a Hall π -subgroup of G , prove that $\theta(H)$ is a Hall π -subgroup of K .

- (b) Show that every Sylow p -subgroup of K has the form $\theta(H)$, where H is some Sylow p -subgroup of G .
- (c) Show that $|\text{Syl}_p(G)| \geq |\text{Syl}_p(K)|$ for every prime p .

Note. If the set π contains more than one prime number, then a Hall π -subgroup can fail to exist. But a theorem of P. Hall, after whom these subgroups are named, asserts that in the case where G is solvable, Hall π -subgroups always do exist. (See Chapter 3, Section C.) We mention also that Part (b) of this problem would not remain true if “Sylow p -subgroup” were replaced by “Hall π -subgroup”.

1B.6. Let G be a finite group, and let $K \subseteq G$ be a subgroup. Suppose that $H \subseteq G$ is a Hall π -subgroup, where π is some set of primes. Show that if HK is a subgroup, then $H \cap K$ is a Hall π -subgroup of K .

Note. In particular, K has a Hall π -subgroup if either H or K is normal in G since in that case, HK is guaranteed to be a subgroup.

1B.7. Let G be a finite group, and let π be any set of primes.

- (a) Show that G has a (necessarily unique) normal π -subgroup N such that $N \supseteq M$ whenever $M \triangleleft G$ is a π -subgroup.
- (b) Show that the subgroup N of Part (a) is contained in every Hall π -subgroup of G .
- (c) Assuming that G has a Hall π -subgroup, show that N is exactly the intersection of all of the Hall π -subgroups of G .

Note. The subgroup N of this problem is denoted $\mathbf{O}_\pi(G)$. Because of the uniqueness in (b), it follows that this subgroup is characteristic in G . Finally, we note that if p is a prime number, then, of course, $\mathbf{O}_{\{p\}}(G) = \mathbf{O}_p(G)$.

1B.8. Let G be a finite group, and let π be any set of primes.

- (a) Show that G has a (necessarily unique) normal subgroup N such that G/N is a π -group and $M \supseteq N$ whenever $M \triangleleft G$ and G/M is a π -group.
- (b) Show that the subgroup N of Part (a) is generated by the set of all elements of G that have order not divisible by any prime in π .

Note. The characteristic subgroup N of this problem is denoted $\mathbf{O}^\pi(G)$. Also, we recall that the subgroup generated by a subset of G is the (unique) smallest subgroup that contains that set.

1C

We are now ready to study in greater detail the nonempty set $\text{Syl}_p(G)$ of Sylow p -subgroups of a finite group G .

1.11. Theorem. *Let P be an arbitrary p -subgroup of a finite group G , and suppose that $S \in \text{Syl}_p(G)$. Then $P \subseteq S^g$ for some element $g \in G$.*

Proof. Let $\Omega = \{Sx \mid x \in G\}$, the set of right cosets of S in G , and note that $|\Omega| = |G : S|$ is not divisible by p since S is a Sylow p -subgroup of G . We know that G acts by right multiplication on Ω , and thus P acts too, and Ω is partitioned into P -orbits. Also, since $|\Omega|$ is not divisible by p , there must exist some P -orbit \mathcal{O} such that $|\mathcal{O}|$ is not divisible by p .

By the fundamental counting principle, $|\mathcal{O}|$ is the index in P of some subgroup. It follows that $|\mathcal{O}|$ divides $|P|$, which is a power of p . Then $|\mathcal{O}|$ is both a power of p and not divisible by p , and so the only possibility is that $|\mathcal{O}| = 1$. Recalling that all members of Ω are right cosets of S in G , we can suppose that the unique member of \mathcal{O} is the coset Sg .

Since Sg is alone in a P -orbit, it follows that it is fixed under the action of P , and thus $Sgu = Sg$ for all elements $u \in P$. Then $gu \in Sg$, and hence $u \in g^{-1}Sg = S^g$. Thus $P \subseteq S^g$, as required. ■

If S is a Sylow p -subgroup of G , and $g \in G$ is arbitrary, then the conjugate S^g is a subgroup having the same order as S . Since the only requirement on a subgroup that is needed to qualify it for membership in the set $\text{Syl}_p(G)$ is that it have the correct order, and since $S \in \text{Syl}_p(G)$ and $|S^g| = |S|$, it follows that S^g also lies in $\text{Syl}_p(G)$. In fact *every* member of $\text{Syl}_p(G)$ arises this way: as a conjugate of S . This is the essential content of the Sylow conjugacy theorem. Putting it another way: the conjugation action of G on $\text{Syl}_p(G)$ is transitive.

1.12. Theorem (Sylow C). *If S and T Sylow p -subgroups of a finite group G , then $T = S^g$ for some element $g \in G$.*

Proof. Applying Theorem 1.11 with T in place of P , we conclude that $T \subseteq S^g$ for some element $g \in G$. But since both S and T are Sylow p -subgroups, we have $|T| = |S| = |S^g|$, and so the containment of the previous sentence must actually be an equality. ■

The Sylow C-theorem yields an alternative proof of Problem 1B.1(b), which asserts that if a group G has a normal Sylow p -subgroup S , then S is the only Sylow p -subgroup of G . Indeed, by the Sylow C-theorem, if $T \in \text{Syl}_p(G)$, then we can write $T = S^g = S$, where the second equality is a consequence of the normality of S .

A frequently used application of the Sylow C-theorem is the so-called “Frattini argument”, which we are about to present. Perhaps the reason that this result is generally referred to as an “argument” rather than as a “lemma” or “theorem” is that variations on its proof are used nearly as often as its statement.

1.13. Lemma (Frattini Argument). *Let $N \triangleleft G$ where N is finite, and suppose that $P \in \text{Syl}_p(N)$. Then $G = \mathbf{N}_G(P)N$.*

Proof. Let $g \in G$, and note that $P^g \subseteq N^g = N$, and thus P^g is a subgroup of N having the same order as the Sylow p -subgroup P . It follows that $P^g \in \text{Syl}_p(N)$, and so by the Sylow C-theorem applied in N , we deduce that $(P^g)^n = P$, for some element $n \in N$. Since $P^{gn} = P$, we have $gn \in \mathbf{N}_G(P)$, and so $g \in \mathbf{N}_G(P)n^{-1} \subseteq \mathbf{N}_G(P)N$. But $g \in G$ was arbitrary, and we deduce that $G = \mathbf{N}_G(P)N$, as required. ■

By definition, a Sylow p -subgroup of a finite group G is a p -subgroup that has the largest possible order consistent with Lagrange’s theorem. By the Sylow E-theorem, we can make a stronger statement: a subgroup whose order is maximal among the orders of all p -subgroups of G is a Sylow p -subgroup. An even stronger assertion of this type is that every maximal p -subgroup of G is a Sylow p -subgroup. Here, “maximal” is to be interpreted in the sense of containment: a subgroup H of G is maximal with some property if there is no subgroup $K > H$ that has the property. The truth of this assertion is the essential content of the Sylow “development” theorem.

1.14. Theorem (Sylow D). *Let P be a p -subgroup of a finite group G . Then P is contained in some Sylow p -subgroup of G .*

Proof. Let $S \in \text{Syl}_p(G)$. Then by Theorem 1.11, we know that $P \subseteq S^g$ for some element $g \in G$. Also, since $|S^g| = |S|$, we know that S^g is a Sylow p -subgroup of G . ■

Given a finite group G , we consider next the question of how many Sylow p -subgroups G has. To facilitate this discussion, we introduce the (not quite standard) notation $n_p(G) = |\text{Syl}_p(G)|$. (Occasionally, when the group we are considering is clear from the context, we will simply write n_p instead of $n_p(G)$.)

First, by the Sylow C-theorem, we know that $\text{Syl}_p(G)$ is a single orbit under the conjugation action of G . The following is then an immediate consequence.

1.15. Corollary. *Let $S \in \text{Syl}_p(G)$, where G is a finite group. Then $n_p(G) = |G : \mathbf{N}_G(S)|$.*

The Thompson Subgroup

7A

In this chapter, we complete the proof of Thompson's theorem that Frobenius kernels are nilpotent. The missing ingredient in our partial proof in Chapter 6 is to show that if $p \neq 2$, then a sufficient condition for a group G to have normal p -complement is that for every nonidentity characteristic subgroup X of some Sylow p -subgroup P of G , the subgroup $\mathbf{N}_G(X)$ has a normal p -complement. (Since every subgroup of a group that has a normal p -complement must also have a normal p -complement, this is clearly a necessary condition too, but, of course, that fact is far less interesting.)

Thompson's first normal p -complement criterion appeared in his 1959 Ph.D. thesis, with a long and subtle proof. Then in 1964, Thompson published a much shorter (though still subtle) proof of a stronger theorem, and that is the theorem and proof we present in this chapter. (Thompson's 1964 paper appeared in the very first issue of the *Journal of Algebra*, which was an auspicious beginning for that periodical.) Thompson showed in his *Journal of Algebra* paper that if $p \neq 2$, then a group G has a normal p -complement provided that $\mathbf{N}_G(X)$ has a normal p -complement for just two specific characteristic subgroups X of P , where $P \in \text{Syl}_p(G)$. These two critical characteristic subgroups are the center $\mathbf{Z}(P)$ and the Thompson subgroup $\mathbf{J}(P)$, both of which are nontrivial if P is nontrivial. (We know, of course, that the center of a nontrivial p -group is nontrivial, and as we shall see when we present the definition, the Thompson subgroup of a nontrivial p -group is also guaranteed to be nontrivial.) Of course, if the normalizers

of *all* nontrivial characteristic subgroups of P have normal p -complements, then in particular, $\mathbf{N}_G(\mathbf{Z}(P))$ and $\mathbf{N}_G(\mathbf{J}(P))$ have normal p -complements, and so Thompson's 1964 theorem yields our Theorem 6.23, which was the key to the proof that Frobenius kernels are nilpotent in Chapter 6. In fact, Thompson's 1964 theorem is slightly stronger than we have just stated since we can replace the normalizer of $\mathbf{Z}(P)$ with the centralizer of $\mathbf{Z}(P)$. (Since $\mathbf{C}_G(\mathbf{Z}(P)) \subseteq \mathbf{N}_G(\mathbf{Z}(P))$, the requirement that the centralizer should have a normal p -complement is weaker than, and is implied by, the corresponding assumption for the normalizer.)

7.1. Theorem (Thompson). *Let $P \in \text{Syl}_p(G)$, where G is a finite group and $p \neq 2$, and assume that $\mathbf{C}_G(\mathbf{Z}(P))$ and $\mathbf{N}_G(\mathbf{J}(P))$ have normal p -complements. Then G has a normal p -complement.*

As we have explained, once we complete the proof of Theorem 7.1, we will have established that all Frobenius kernels are nilpotent. The Thompson subgroup $\mathbf{J}(P)$ has other applications too, and in particular, it provides a crucial ingredient in the group-theoretic proof of Burnside's $p^a q^b$ -theorem that we present later in this chapter. For that reason, we prove somewhat more about the Thompson subgroup than the minimum needed for the proof of Theorem 7.1.

Unfortunately, the literature contains several slightly different definitions of "the" Thompson subgroup of a p -group P , and in general, these definitions yield different subgroups, all of which have been referred to as $\mathbf{J}(P)$. In particular, the version of $\mathbf{J}(P)$ that we are about to define is not always equal to Thompson's $\mathbf{J}(P)$. Our definition is slightly easier to use than that in Thompson's paper, however, and we feel that this justifies the risk of confusion caused by competing definitions.

Given a p -group P , let $\mathcal{E}(P)$ be the set of all of those elementary abelian subgroups of P that have the maximum possible order. Then the **Thompson subgroup** $\mathbf{J}(P)$ of P is the subgroup generated by all of the members of $\mathcal{E}(P)$. For example, if P is the dihedral group of order 8, then obviously, P does not contain an elementary abelian subgroup of order 8. But P does contain an elementary abelian subgroup of order 4, and so $\mathcal{E}(P)$ is the set of all elementary abelian subgroups of order 4 in P . There are two of these, and together they generate the whole group P , and thus $\mathbf{J}(P) = P$ in this case. Of course, if P is nontrivial, then it must contain a nontrivial elementary abelian subgroup, for instance, one of order p , and thus the set $\mathcal{E}(P)$ contains at least one nontrivial subgroup. It follows that $\mathbf{J}(P) > 1$, as we mentioned previously.

In his 1964 paper, Thompson did not restrict attention to elementary abelian subgroups. Instead, he defined $\mathbf{J}(P)$ to be the subgroup generated by

all abelian subgroups of P of largest possible rank. (The **rank** of an abelian p -group A is the integer r such that the unique largest elementary abelian subgroup $\Omega_1(A)$ has order p^r . Equivalently, if the abelian p -group A is decomposed as a direct product of nontrivial cyclic factors, then the rank of A is the number of such factors.) Since an elementary abelian subgroup with maximum possible order in P is an abelian subgroup of maximum possible rank, we see that our subgroup $\mathbf{J}(P)$ is always contained in Thompson's, but it can be properly smaller. (The containment is proper, for example, if P is abelian, but not elementary abelian.) Yet another variation on the definition that has appeared in the literature is to consider the subgroup of P generated by all abelian subgroups of largest possible order.

An immediate consequence of the definition is the following.

7.2. Lemma. *Let $\mathbf{J}(P) \subseteq Q \subseteq P$, where P is a p -group. Then $\mathbf{J}(P) = \mathbf{J}(Q)$, and in particular, $\mathbf{J}(P)$ is characteristic in Q .*

Proof. Since $\langle \mathcal{E}(P) \rangle = \mathbf{J}(P) \subseteq Q$, it follows that every maximal-order elementary abelian subgroup of P is contained in Q , and because $Q \subseteq P$, these are maximal-order elementary abelian subgroups of Q . Every maximal-order elementary abelian subgroup of Q , therefore, has the same order as the members of $\mathcal{E}(P)$, and hence is a member of $\mathcal{E}(P)$. It follows that $\mathcal{E}(Q) = \mathcal{E}(P)$, and thus $\mathbf{J}(Q) = \mathbf{J}(P)$. ■

Before we can begin our proof of Theorem 7.1, we need a number of preliminary results. The first of these is a fairly technical lemma about the general linear group $GL(2, p)$, which, we recall, is the group of invertible 2×2 matrices over the field F of order p .

7.3. Lemma. *Let $G = GL(2, p)$, where $p \neq 2$ is prime, and let $P \subseteq G$ be a p -subgroup. Suppose $P \subseteq \mathbf{N}_G(L)$, for some subgroup L of G , where $|L|$ is not divisible by p , and assume further that a Sylow 2-subgroup of L is abelian. Then $P \subseteq \mathbf{C}_G(L)$.*

The condition that $p \neq 2$ in Lemma 7.3 is necessary since $GL(2, 2)$ is isomorphic to the symmetric group S_3 , and so there is a p' -group L of order 3 normalized but not centralized by a Sylow p -subgroup P of order 2. The somewhat unnatural requirement that L should have an abelian Sylow 2-subgroup also cannot be omitted, at least when $p = 3$. If $G = GL(2, 3)$, then G has a subgroup L of order 8 that is normalized but not centralized by a Sylow 3-subgroup P , which has order 3. (The product LP of order 24 is the special linear group $SL(2, 3)$, which is the set of matrices with determinant 1 in G .) This is not a counterexample to Lemma 7.3, however, since L is a nonabelian 2-group, isomorphic to the quaternion group Q_8 . In fact, it is only for $p = 3$ that such an example can occur, and so Lemma 7.3 could be

strengthened by requiring the condition on a Sylow 2-subgroup of L only if $p = 3$. When we apply the lemma, however, we will know that L is abelian, and so we have no need for a stronger result.

We digress briefly to discuss general linear groups in general, and also special linear groups and certain other related groups. (We saw some of these groups in Chapter 1, and we study them further in Chapter 8.)

If F is an arbitrary field and n is a positive integer, then $GL(n, F)$ is the group of all invertible $n \times n$ matrices over F , and the special linear group $SL(n, F)$ is the subgroup of $GL(n, F)$ consisting of those matrices that have determinant equal to 1. Since the determinant map is a homomorphism from $GL(n, F)$ onto the multiplicative group F^\times of F , it follows that $SL(n, F)$, which is the kernel of this homomorphism, must be a normal subgroup. Also, we have $GL(n, F)/SL(n, F) \cong F^\times$.

Now suppose that F is finite, so that $|F| = q$, where q is a prime power. Since F is the unique field (up to isomorphism) of order q , it is no loss to write $GL(n, q)$ in place of $GL(n, F)$, and in fact, this notation is fairly common. Similarly, we write $SL(n, q)$ for the corresponding special linear group, which is a normal subgroup of index $|F^\times| = q - 1$ in $GL(n, q)$.

We can compute $|GL(n, q)|$ by counting the number of ways we can construct an $n \times n$ matrix with linearly independent rows, where each entry comes from the field F of order q . Of course, there are a total of q^n row vectors of length n over F . Each of these other than the zero row is a possibility for the first row of our matrix, and so there are exactly $q^n - 1$ possible first rows. To keep the rows linearly independent, we must not allow the second row to be one of the q different scalar multiples of the (nonzero) first row, so after the first row is chosen, there are exactly $q^n - q$ possibilities for the second row. After the first two (linearly independent) rows are selected, we must exclude all of their q^2 different linear combinations from appearing as the third row, and this leaves $q^n - q^2$ possibilities for the third row. Continuing like this, we see that there are exactly $q^n - q^{k-1}$ possibilities for the k th row, and we deduce that

$$|GL(n, q)| = \prod_{i=0}^{n-1} (q^n - q^i) = q^N \prod_{j=1}^n (q^j - 1),$$

where $N = 1 + 2 + \cdots + (n - 1) = n(n - 1)/2$. In particular, $|GL(2, q)| = (q^2 - 1)(q^2 - q) = q(q - 1)^2(q + 1)$, and $|SL(2, q)| = q(q - 1)(q + 1)$.

Now consider the subgroup $P \subseteq GL(n, q)$ consisting of the matrices that have zeros below the diagonal, ones on the diagonal and arbitrary elements of F in all above-diagonal positions. (Note that P really is a subgroup, and that $P \subseteq SL(n, q)$.) Since row k of an $n \times n$ matrix contains exactly $k - 1$ above-diagonal positions, we see that the total number of above-diagonal

positions in an $n \times n$ matrix is $1 + 2 + \cdots + (n - 1) = N$, where N is as before. Then $|P| = q^N$, and thus $|GL(n, q)| = |P|m$, where m is a product of numbers of the form $q^j - 1$, and so m is relatively prime to q . Since q is a power of some prime, say p , we see that the subgroup P is a Sylow p -subgroup of $GL(n, q)$. Note that if $n > 1$, we can easily find a second Sylow p -subgroup, different from P , by taking lower triangular matrices in place of the upper triangular subgroup P .

The scalar matrices in $GL(n, q)$ are the scalar multiples of the identity matrix, and so there are exactly $q - 1$ of these, and it is easy to see that these matrices form the center $\mathbf{Z}(GL(n, q))$. The factor group $GL(n, q)/\mathbf{Z}(GL(n, q))$ is the projective general linear group, and is denoted $PGL(n, q)$. The scalar matrices that lie in $SL(n, q)$ are the matrices of the form $\alpha \cdot \mathbf{1}$, where $\alpha^n = 1$, and in fact, these form the center $\mathbf{Z}(SL(n, q))$. Writing $Z = \mathbf{Z}(SL(n, q))$, we see that $|Z|$ is equal to the number d of elements $\alpha \in F^\times$ such that $\alpha^n = 1$. Since F^\times is cyclic of order $q - 1$, it follows that $d = (n, q - 1)$, the greatest common divisor. By definition, the projective special linear group $PSL(n, q)$ is the factor group $SL(n, q)/Z$. For $n \geq 2$, this group is simple except when $n = 2$ and q is 2 or 3. (In fact, $PSL(n, F)$ is also simple when $n \geq 2$ and F is an infinite field.)

Taking $n = 2$, we have $|PSL(2, q)| = q(q - 1)(q + 1)/d$, where $d = 1$ if q is a power of 2 and $d = 2$ if q is odd. In particular, for prime powers $q = 4, 5, 7, 8, 9, 11$ we have $|PSL(2, q)| = 60, 60, 168, 504, 360, 660$, respectively, and these account for all of the nonabelian simple groups of order at most 1000. (We have not omitted the alternating groups A_5 and A_6 since $PSL(2, 4) \cong A_5 \cong PSL(2, 5)$ and $PSL(2, 9) \cong A_6$.) We mention also that $PSL(2, 7) \cong PSL(3, 2)$ and $A_8 \cong PSL(4, 2)$. In fact, these are the only isomorphisms among alternating groups and the simple groups of type $PSL(n, q)$. (The simple groups $PSL(3, 4)$ and A_8 have equal orders, but they are not isomorphic.)

Finally, we mention one way that general linear groups arise in abstract group theory. Consider an elementary abelian p -group E of order p^n , and view E as an additive group. Then E can be identified with a vector space of dimension n over the field of order p , and under this identification, we see that $\text{Aut}(E) \cong GL(n, p)$. This observation (in the case $n = 2$) is crucial to the proof of Theorem 7.1, and that is why Lemma 7.3 is relevant.

In order to prove Lemma 7.3, we need the following observation about the groups $SL(2, q)$, where q is odd.

7.4. Lemma. *If q is odd, then the negative of the identity matrix is the unique involution in $SL(2, q)$.*

This can be proved by a computational argument along the following lines. If t is an involution in $SL(2, q)$, then the fact that $t^2 = 1$ yields four nonlinear equations that must be satisfied by the four entries of the matrix t , and the fact that $\det(t) = 1$ yields one more equation. It is not hard to show that if the characteristic is different from 2, these five equations have just two solutions, corresponding to $t = I$ and $t = -I$, where I is the 2×2 identity matrix. We prefer the following more conceptual proof, however.

Proof of Lemma 7.4. Let $t \in SL(2, q)$ be an involution other than $-I$. Write $f(X)$ to denote the polynomial $X^2 - 1$, and observe that $f(t) = 0$. Since neither the polynomial $X - 1$ nor the polynomial $X + 1$ yields 0 when t is substituted for X , it follows that $f(X) = X^2 - 1$ is the minimal polynomial for t . The Cayley-Hamilton theorem asserts that the minimal polynomial of an arbitrary square matrix divides the characteristic polynomial, and in our situation, where t is a 2×2 matrix, we know that the characteristic polynomial of t has degree 2, and hence $f(X)$ is the characteristic polynomial of t . Then t has two distinct eigenvalues, 1 and -1 , and so the product of the eigenvalues of t is -1 . But the product of the eigenvalues (counting multiplicities) for an arbitrary square matrix is the determinant of the matrix, and it follows that $\det(t) = -1$. But $t \in SL(2, q)$, and so $\det(t) = 1$, and this contradiction completes the proof. ■

It follows by Theorem 6.11 that a Sylow 2-subgroup of $SL(2, q)$ (for odd q) is either cyclic or generalized quaternion. An elementary argument that shows that the cyclic case can never occur depends on the fact that in a finite field F of odd order q , it is always possible to find elements a and b such that $a^2 + b^2 = -1$. Assuming that a and b satisfy this condition, consider the two matrices:

$$x = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}.$$

These clearly have determinant 1, and so they lie in $SL(2, q)$. It is easy to check that $x^2 = -I = y^2$ and that $y^x = -y = y^{-1}$, and it follows that the subgroup $\langle x, y \rangle$ of $SL(2, q)$ is isomorphic to the quaternion group Q_8 . A Sylow 2-subgroup of $SL(2, q)$ cannot be cyclic, therefore, and so it must be generalized quaternion. We mention also that since $|SL(2, 3)| = 24$, a Sylow 2-subgroup of $SL(2, 3)$ is isomorphic to Q_8 .

To see why the equation $a^2 + b^2 = -1$ must have a solution in F , observe that the number of elements of F that have the form $1 + a^2$ is exactly $(q + 1)/2$. (This is so because the map $a \mapsto 1 + a^2$ is two-to-one for nonzero elements of F , and so the image of this map contains $(q - 1)/2$ elements different from 1, for a total of $(q + 1)/2$ elements.) Similarly, there are exactly $(q + 1)/2$ elements of F that have the form $-b^2$, and since the

sum of the cardinalities of these two sets exceeds q , the sets must overlap. It follows that $1 + a^2 = -b^2$ for some choice of $a, b \in F$, as wanted.

Proof of Lemma 7.3. Working by induction on $|L|$, we can assume that P centralizes every proper subgroup of L that it stabilizes. If we choose a prime q dividing $|L : \mathbf{C}_L(P)|$, we can find a P -invariant Sylow q -subgroup Q of L , and since $Q \not\subseteq \mathbf{C}_L(P)$, we must have $Q = L$, and thus L is a q -group. Also, $[L, P]$ is P -invariant, and so if $[L, P] < L$, we have $[L, P, P] = 1$, and since $[L, P, P] = [L, P]$ by Lemma 4.29, we conclude that $[L, P] = 1$, as required. We may assume, therefore, that $[L, P] = L$, and in particular, $L \subseteq G' \subseteq SL(2, p)$, where the second containment follows because $GL(2, p)/SL(2, p)$ is isomorphic to the multiplicative group of the field of order p , and so it is abelian.

If $q = 2$, then by assumption, the 2-group L is abelian, and we know by Lemma 7.4 that L contains a unique involution. It follows that L is a cyclic 2-group, and so $\text{Aut}(L)$ is also a 2-group, and the p -group P cannot act nontrivially on L . We can assume, therefore, that q is an odd prime, and since $|L|$ is a power of q that divides $|SL(2, p)| = p(p-1)(p+1)$, it follows that $|L|$ divides one of $p-1$ or $p+1$. (This, of course, is because the prime q cannot divide both $p-1$ and $p+1$.) In particular, we have $|L| \leq p+1$. Now if P acts nontrivially on L , there must be at least one P -orbit in L of size at least p , and together with the identity, this yields $|L| \geq p+1$. We conclude that $|L| = p+1$, and so $|L|$ is even. But $|L|$ is a power of the odd prime q , and this contradiction completes the proof. ■

Next, we prove what we call the “normal- P theorem”.

7.5. Theorem. *Let $P \in \text{Syl}_p(G)$, where G is a finite group, and assume the following conditions.*

- (1) G is p -solvable.
- (2) $p \neq 2$.
- (3) A Sylow 2-subgroup of G is abelian.
- (4) G acts faithfully on some p -group V .
- (5) $|V : \mathbf{C}_V(P)| \leq p$.

Then $P \triangleleft G$.

The reader may wonder how Theorem 7.5, which applies only to p -solvable groups, can possibly be relevant to the proof of Thompson’s theorem (7.1), where there certainly is no p -solvability assumption. The answer is that the proof of Theorem 7.1 proceeds via a subtle induction that will guarantee p -solvability when we need it. (Thompson’s inductive argument

will also guarantee that the Sylow 2-subgroup is abelian when that fact is needed.)

Proof of Theorem 7.5. Assume that G is a counterexample of minimum possible order. Then G does not have a normal Sylow p -subgroup, and we can choose $Q \in \text{Syl}_p(G)$ with $Q \neq P$. The subgroup $\langle Q, P \rangle$ of G has at least two Sylow p -subgroups, and it clearly satisfies all five hypotheses of the theorem. It too is a counterexample, therefore, and so by the minimality of G , we have $G = \langle P, Q \rangle$.

Now $Q = P^g$ for some element $g \in G$, and thus $\mathbf{C}_V(Q) = (\mathbf{C}_V(P))^g$ also has index at most p in V . Let $U = \mathbf{C}_V(P) \cap \mathbf{C}_V(Q)$, and observe that $|V : U| \leq |V : \mathbf{C}_V(P)| |V : \mathbf{C}_V(Q)| \leq p^2$, and $U \triangleleft V$. Also, since both P and Q act trivially on U , it follows that the action of $G = \langle P, Q \rangle$ on U is trivial, and so $[U, G] = 1$. In particular, U is G -invariant, and we have a natural action of G on V/U .

Let K be the kernel of the action of G on V/U , so that $[V, K] \subseteq U$, and thus $[V, K, K] \subseteq [U, K] = 1$. Since V is a p -group, it follows that K must also be a p -group. (If Q is a Sylow q -subgroup of K , where $q \neq p$, then $[V, Q] = [V, Q, Q] = 1$, and thus $Q = 1$ since the action of G on V is faithful, by assumption.) Since $K \subseteq \mathbf{O}_p(G)$, it follows that $K \subseteq P$ and $K \subseteq Q$, and thus the group $\overline{G} = G/K$ has distinct Sylow p -subgroups \overline{P} and \overline{Q} . Also, \overline{G} acts faithfully on V/U , and we see that \overline{P} centralizes $\mathbf{C}_V(\overline{P})/U$, which has index $|V : \mathbf{C}_V(P)| \leq p$ in V/U . It follows that the group \overline{G} satisfies all five conditions with respect to its Sylow p -subgroup \overline{P} and its action on V/U , but it does not have a normal Sylow p -subgroup. By the minimality of G , therefore, we must have $K = 1$, and so G acts faithfully on V/U . We can thus replace V by V/U and assume that $|V| \leq p^2$.

Since G acts faithfully on V , it is isomorphically embedded in $\text{Aut}(V)$. If V is cyclic, then $\text{Aut}(V)$ is abelian, and so G is abelian and $P \triangleleft G$, which is not the case. We conclude that V is noncyclic, and so it must be elementary abelian of order p^2 , and thus $\text{Aut}(V) \cong GL(2, p)$, and we can view G as a subgroup of $GL(2, p)$. In particular, $|P| \leq p$, and since P is not normal, it follows that $\mathbf{O}_p(G) = 1$. Now let $L = \mathbf{O}_{p'}(G)$, so that P normalizes the p' -group L . Since we are assuming that a Sylow 2-subgroup of G is abelian, the same is true for L , and thus by Lemma 7.3, we see that P acts trivially on L . By the Hall-Higman Lemma 1.2.3 (our Theorem 3.21) we have $P \subseteq \mathbf{C}_G(L) \subseteq L$, and thus since P is a p -group and L is a p' -group, we have $P = 1$. This is a contradiction since P is not normal. ■

Problems 7A

7A.1. Show that a nilpotent maximal subgroup of a simple group must be a 2-group.

Note. The simple group $PSL(2, 17)$ has a maximal subgroup of order 16, so nontrivial 2-groups actually can occur as maximal subgroups of simple groups.

7A.2. Let $S = SL(2, 3)$ and write $Z = \langle -I \rangle$, so that Z is the unique subgroup of order 2 in S . Show that the group S/Z of order 12 has four Sylow 3-subgroups, and deduce that it has a unique Sylow 2-subgroup. Conclude from this that S has a normal subgroup of order 8.

Hint. Use the fact mentioned in the text that if $n \geq 2$ and q is a power of the prime p , then $GL(n, q)$ has at least two Sylow p -subgroups. No further matrix computations are required for this problem.

7A.3. Let $G = GL(n, q)$, where q is a power of the prime p , and let P be the Sylow p -subgroup of G consisting of the upper triangular matrices with ones on the diagonal. Let D be the group of all diagonal matrices in G .

- (a) Show that $D \subseteq \mathbf{N}_G(P)$.
- (b) Show that DP is the group of all upper triangular matrices in G with arbitrary nonzero entries on the diagonal.
- (c) View G as acting by right multiplication on the space of n -dimensional row vectors over the field F of order q . Identify all P -invariant subspaces of V , and observe that there is exactly one such subspace of dimension k for each integer k with $1 \leq k \leq n$.
- (d) Show that all of the P -invariant subspaces of V are N -invariant, where $N = \mathbf{N}_G(P)$, and use (c) to deduce that $N = DP$.

7A.4. If q is a power of the prime p , show that $SL(2, q)$ and $PSL(2, q)$ each have exactly $q + 1$ Sylow p -subgroups.

7A.5. Let P be a p -group, and suppose that $U \triangleleft P$ is elementary abelian. Show that U normalizes some group $E \in \mathcal{E}(P)$.

Hint. Choose $E \in \mathcal{E}(P)$ such that $|U \cap E|$ is as large as possible. If U does not normalize E , write $F = E^u$, where $u \in U$ and $E \neq F$. Let $H = \langle E, F \rangle$ and $Z = E \cap F$. Show that $Z(H \cap U) \in \mathcal{E}(P)$.

7A.6. Suppose that Q is a generalized quaternion group and that $|\text{Aut}(Q)|$ is divisible by an odd prime p . Show that $p = 3$ and $|Q| = 8$, and deduce that the hypothesis that a Sylow 2-subgroup is abelian in Lemma 7.3 and Theorem 7.5 is unnecessary if $p > 3$.

7B

In this section, we prove the following “normal-J theorem”. This will be used for the proof of Theorem 7.1, and also for the group-theoretic proof of Burnside’s $p^a q^b$ -theorem.

7.6. Theorem. *Let $P \in \text{Syl}_p(G)$, where G is a finite group, and assume the following conditions.*

- (1) G is p -solvable.
- (2) $p \neq 2$.
- (3) A Sylow 2-subgroup of G is abelian.
- (4) $\mathbf{O}_{p'}(G) = 1$.
- (5) $P = \mathbf{C}_G(\mathbf{Z}(P))$.

Then $\mathbf{J}(P) \triangleleft G$.

Since we are assuming that G is p -solvable and that $\mathbf{O}_{p'}(G) = 1$, it follows (assuming that G is nontrivial) that $\mathbf{O}_p(G) > 1$. The point of the theorem, however, is not simply that G has a nontrivial normal p -subgroup; it is that a *particular* p -subgroup, namely $\mathbf{J}(P)$, is normal in G . Note that it is not until Step 7 of the following proof that we use the actual definition of the set $\mathcal{E}(P)$.

Proof of Theorem 7.6. Suppose that the theorem is false, and let G be a counterexample of minimum possible order. In particular, G is nontrivial, so $\mathbf{O}_p(G) > 1$, and we write $U = \mathbf{O}_p(G)$ and $\overline{G} = G/U$. Also, let $\overline{L} = \mathbf{O}_{p'}(\overline{G})$, where $L \supseteq U$. (This, of course, unambiguously defines the subgroup $L \triangleleft G$.) We proceed in a number of steps, the first of which consists of three closely related statements, all of which are consequences of the Hall-Higman Lemma 1.2.3.

Step 1.

- (a) $\mathbf{Z}(P) \subseteq U$.
- (b) If $U \subseteq H \subseteq G$, then $\mathbf{O}_{p'}(H) = 1$.
- (c) $\mathbf{C}_{\overline{G}}(\overline{L}) \subseteq \overline{L}$.

Proof. Since G is p -solvable and $\mathbf{O}_{p'}(G) = 1$, Lemma 1.2.3 (our Theorem 3.21) tells us that $U = \mathbf{O}_p(G) \supseteq \mathbf{C}_G(U)$. But $U \subseteq P$, and so $\mathbf{Z}(P) \subseteq \mathbf{C}_G(U)$ and (a) follows. Also, since $U = \mathbf{O}_p(G)$, we see that $\mathbf{O}_p(\overline{G}) = 1$, and Lemma 1.2.3 yields (c). For (b), let $U \subseteq H \subseteq G$ and write $M = \mathbf{O}_{p'}(H)$. Then M and U are normal in H and $M \cap U = 1$ since U is a p -group and M is a p' -group. Thus $M \subseteq \mathbf{C}_G(U) \subseteq U$, so $M = M \cap U = 1$, proving (b).

Step 2. There exists $A \in \mathcal{E}(P)$ such that $A \not\subseteq U$.

Proof. Otherwise, all members of $\mathcal{E}(P)$ are contained in U , and thus $\mathbf{J}(P) \subseteq U$. It follows by Lemma 7.2 that $\mathbf{J}(P) = \mathbf{J}(U)$ is characteristic in U , and since $U \triangleleft G$, we have $\mathbf{J}(P) \triangleleft G$. This is a contradiction since we are assuming that G is a counterexample.

Step 3. Let $UA \subseteq H < G$, and suppose that $H \cap P \in \text{Syl}_p(H)$. Then \overline{A} centralizes $\overline{H \cap L}$.

Proof. Observe that H satisfies the first four hypotheses of the theorem. First, H is p -solvable since it is a subgroup of the p -solvable group G . Also, p is still different from 2, so the second condition holds. A Sylow 2-subgroup of H is contained in a Sylow 2-subgroup of G , which is abelian, and so H satisfies condition (3), and finally, $\mathbf{O}_{p'}(H) = 1$ by Step 1(b).

Now let $S = H \cap P \in \text{Syl}_p(H)$. Since $\mathbf{Z}(P) \subseteq U \subseteq S \subseteq P$ by Step 1(a), we have $\mathbf{Z}(P) \subseteq \mathbf{Z}(S)$, and thus $\mathbf{C}_H(\mathbf{Z}(S)) \subseteq \mathbf{C}_G(\mathbf{Z}(P)) = P$. Thus $\mathbf{C}_H(\mathbf{Z}(S))$ is a p -subgroup of H containing the Sylow p -subgroup S , and so $\mathbf{C}_H(\mathbf{Z}(S)) = S$ and H satisfies the fifth hypothesis too.

Since $H < G$, the theorem holds for H , and hence $\mathbf{J}(S) \triangleleft H$. Also, since $A \in \mathcal{E}(P)$ and $A \subseteq S \subseteq P$, it follows that $A \in \mathcal{E}(S)$, and so $A \subseteq \mathbf{J}(S)$. Then

$$[H \cap L, A] \subseteq [H \cap L, \mathbf{J}(S)] \subseteq (H \cap L) \cap \mathbf{J}(S) = L \cap \mathbf{J}(S) \subseteq U,$$

where the second containment holds because both $H \cap L$ and $\mathbf{J}(S)$ are normal in H , and the final containment holds because U is the unique Sylow p -subgroup of L , and thus it contains every p -subgroup of L . We now have $1 = \overline{[H \cap L, A]} = \overline{[H \cap L, \overline{A}]}$, as wanted.

Step 4. $G = LA$ and $P = UA$.

Proof. Write $H = LA$, and observe that UA is a p -subgroup of H and that $|H : UA| = |L(UA) : UA| = |L : L \cap UA|$, which divides the p' -number $|L : U|$. It follows that $UA \in \text{Syl}_p(H)$, and thus $UA = H \cap P$. If $H < G$, then since $L \subseteq H$, Step 3 yields $\overline{A} \subseteq \mathbf{C}_{\overline{G}}(\overline{L}) \subseteq \overline{L}$, where the second containment holds by Step 1(c). Since \overline{A} is a p -group and \overline{L} is a p' -group, we have $\overline{A} = 1$, and so $A \subseteq U$. This contradicts the choice of A , however, and we conclude that $H = G$, as wanted. Finally, we have $UA = H \cap P = G \cap P = P$.

Step 5. $|\overline{A}| = p$.

Proof. First, \overline{A} is nontrivial since $A \not\subseteq U$. Also, \overline{A} is elementary abelian, and so it suffices to show that it is cyclic. Now \overline{A} acts coprimely on \overline{L} , and this action is faithful since $\mathbf{C}_{\overline{G}}(\overline{L}) \subseteq \overline{L}$ and $\overline{L} \cap \overline{A} = 1$. By Lemma 6.20,

therefore, it suffices to show that \bar{A} acts trivially on every \bar{A} -invariant proper subgroup of \bar{L} .

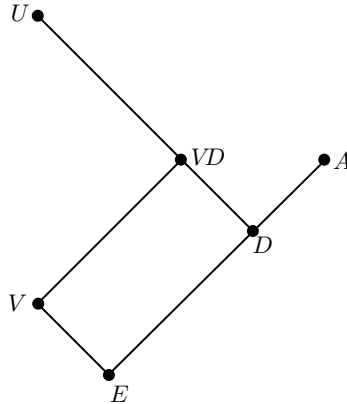
Suppose that \bar{M} is \bar{A} -invariant, where $\bar{M} < \bar{L}$, and where we assume (as we may) that $M \supseteq U$. Then $A \subseteq \mathbf{N}_G(M)$, and so MA is a group, and $MA \supseteq UA = P$. Also, since A is a p -group, the p' -part of $|MA|$ is equal to the p' -part of $|M|$, and this is strictly less than the p' -part of $|L|$ since $|L : M|$ is a p' -number exceeding 1. It follows that $MA < G$, and we can apply Step 3 to deduce that \bar{A} centralizes $\overline{MA \cap L} \supseteq \bar{M}$, proving Step 5.

Now let $V = \{z \in \mathbf{Z}(U) \mid z^p = 1\}$, so that V is an elementary abelian normal subgroup of G . Then G acts by conjugation on V , and since $V \subseteq \mathbf{Z}(U)$, we see that U acts trivially. This yields an action of $\bar{G} = G/U$ on V .

Step 6. The action of \bar{G} on V is faithful.

Proof. Let $K = \mathbf{C}_G(V)$, so that \bar{K} is the kernel of the action of \bar{G} on V . We argue that K is a p -group by considering a Sylow q -subgroup Q of K , where q is any prime different from p . Then Q acts coprimely on the abelian group $\mathbf{Z}(U)$, and Q fixes all elements of order p in $\mathbf{Z}(U)$ since these elements all lie in V and $Q \subseteq K = \mathbf{C}_G(V)$. It follows by Corollary 4.35 that Q acts trivially on $\mathbf{Z}(U)$. But $\mathbf{Z}(P) \subseteq U$, and so $\mathbf{Z}(P) \subseteq \mathbf{Z}(U)$, and thus $Q \subseteq \mathbf{C}_G(\mathbf{Z}(P)) = P$. We conclude that $Q = 1$, and thus K is a p -group, as claimed. But $K \triangleleft G$, so $K \subseteq \mathbf{O}_p(G) = U$, and thus $\bar{K} = 1$, as wanted.

Step 7. $|V : V \cap A| \leq p$.



Proof. Write $D = U \cap A$ and $E = V \cap A$ as in the diagram, and observe that $|V : E| = |V : V \cap D| = |VD : D|$. Now D is an elementary abelian subgroup of U , and V is a central elementary abelian subgroup of U , and thus VD is elementary abelian. Since $A \in \mathcal{E}(P)$, it follows that $|VD| \leq |A|$, and thus $|VD : D| \leq |A : D| = |\overline{A}| = p$. We now have $|V : E| = |VD : D| \leq p$, as wanted.

Step 8. We have a contradiction.

We want to apply the “normal- P theorem” (Theorem 7.5) to the faithful action of \overline{G} on V , and so we need to check that $|V : \mathbf{C}_V(\overline{P})| \leq p$. But $P = UA$, so $\overline{P} = \overline{A}$, and of course, \overline{A} centralizes $V \cap A$ since A is abelian. Thus $|V : \mathbf{C}_V(\overline{P})| \leq |V : V \cap A| \leq p$, as wanted, and we conclude that $\overline{P} \triangleleft \overline{G}$. Then $P \triangleleft G$ and $A \subseteq P \subseteq \mathbf{O}_p(G) = U$, which is not the case. ■

7C

In this section, we prove Theorem 7.1, thereby completing the proof that Frobenius kernels are nilpotent. We begin with an extension of Lemma 2.17.

7.7. Lemma. *Write $\overline{G} = G/N$, where N is a normal p' -subgroup of a finite group G . If P is a p -subgroup of G , we have*

- (a) $\mathbf{N}_{\overline{G}}(\overline{P}) = \overline{\mathbf{N}_G(P)}$ and
- (b) $\mathbf{C}_{\overline{G}}(\overline{P}) = \overline{\mathbf{C}_G(P)}$.

Proof. Statement (a) is essentially Lemma 2.17. Also, since overbar is a homomorphism, it is clear that $\overline{\mathbf{C}_G(P)} \subseteq \mathbf{C}_{\overline{G}}(\overline{P})$, and so to establish (b), it suffices to prove the reverse containment. By (a), overbar defines a surjective homomorphism from $\mathbf{N}_G(P)$ to $\mathbf{N}_{\overline{G}}(\overline{P})$, and so by the correspondence theorem, the subgroup $\mathbf{C}_{\overline{G}}(\overline{P}) \subseteq \mathbf{N}_{\overline{G}}(\overline{P})$ is the image of some subgroup $X \subseteq \mathbf{N}_G(P)$. In other words, $\overline{X} = \mathbf{C}_{\overline{G}}(\overline{P})$, and we have $1 = [\overline{P}, \overline{X}] = \overline{[P, X]}$, and thus $[P, X] \subseteq N$. But $X \subseteq \mathbf{N}_G(P)$, and so we also have $[P, X] \subseteq P$. Thus $\overline{[P, X]} \subseteq N \cap P = 1$, and $X \subseteq \mathbf{C}_G(P)$. We conclude that $\mathbf{C}_{\overline{G}}(\overline{P}) = \overline{X} \subseteq \overline{\mathbf{C}_G(P)}$, and the proof is complete. ■

We are now ready to prove Theorem 7.1, which, we recall, asserts that G has a normal p -complement if $p \neq 2$ and both $\mathbf{N}_G(\mathbf{J}(P))$ and $\mathbf{C}_G(\mathbf{Z}(P))$ have normal p -complements, where P is a Sylow p -subgroup of G . (Of course, if these conditions hold for P , they will also hold for every Sylow p -subgroup of G .) We repeatedly use the fact that the property of having a normal p -complement is inherited by subgroups and homomorphic images.

Observe the subtle way that the subgroup U is chosen in the following proof. This is one of the keys to this very clever argument. (This trick appears both in Thompson's thesis and in his 1964 paper.)

Proof of Theorem 7.1. Suppose that G is a counterexample of minimal order. Then G fails to have a normal p -complement, and so by Frobenius' theorem (Theorem 5.26), there must be some nonidentity p -subgroup U of G such that $\mathbf{N}_G(U)$ fails to have a normal p -complement. Among all such "bad" subgroups, choose U such that $|\mathbf{N}_G(U)|_p$ is as large as possible. (In other words, U is chosen so that a Sylow p -subgroup of $\mathbf{N}_G(U)$ has order as large as possible.) Subject to that condition, choose U to have order as large as possible. We proceed in a number of steps.

Step 1. $U = \mathbf{O}_p(G)$.

Proof. Let $S \in \text{Syl}_p(N)$, where $N = \mathbf{N}_G(U)$, and suppose that $N < G$. Then N is not a counterexample to the theorem, and since N fails to have a normal p -complement, at least one of $\mathbf{N}_N(\mathbf{J}(S))$ or $\mathbf{C}_N(\mathbf{Z}(S))$ must fail to have a normal p -complement. Then one of $\mathbf{N}_G(\mathbf{J}(S))$ or $\mathbf{C}_G(\mathbf{Z}(S))$ fails to have a normal p -complement, and so by hypothesis, S cannot be a full Sylow p -subgroup of G . Then S is properly contained in a Sylow p -subgroup, and hence we can choose a p -subgroup T such that $S \triangleleft T$ and $S < T$. (We are, of course, using the fact that "normalizers grow" in p -groups.)

Now $\mathbf{N}_G(X)$ fails to have a normal p -complement, where X is one of $\mathbf{J}(S)$ or $\mathbf{Z}(S)$. Also $U > 1$, so $S > 1$, and thus $X > 1$, and therefore X is a member of the set of bad subgroups from which we selected U . Also, X is characteristic in S , and so $X \triangleleft T$ and $T \subseteq \mathbf{N}_G(X)$. We thus have $|\mathbf{N}_G(X)|_p \geq |T| > |S| = |\mathbf{N}_G(U)|_p$, and since this contradicts the choice of U , we conclude that $N = G$, and so $U \triangleleft G$ and $U \subseteq \mathbf{O}_p(G)$. Now $\mathbf{O}_p(G)$ is a member of the set of nonidentity p -subgroups whose normalizers fail to have normal p -complements, and since $|\mathbf{N}_G(\mathbf{O}_p(G))|_p = |P| = |\mathbf{N}_G(U)|_p$, the choice of U guarantees that $|U| \geq |\mathbf{O}_p(G)|$, and thus $U = \mathbf{O}_p(G)$, as required.

Step 2. G/U has a normal p -complement, and so G is p -solvable.

Proof. Write $\bar{G} = G/U$, and observe that $|\bar{G}| < |G|$. Then \bar{G} is not a counterexample to the theorem, and so since $\bar{P} \in \text{Syl}_p(\bar{G})$, it suffices to show that each of $\mathbf{N}_{\bar{G}}(\mathbf{J}(\bar{P}))$ and $\mathbf{C}_{\bar{G}}(\mathbf{Z}(\bar{P}))$ has a normal p -complement. There is nothing to prove if p does not divide $|\bar{G}|$, and so we can assume that \bar{P} is nontrivial, and thus $\mathbf{J}(\bar{P})$ and $\mathbf{Z}(\bar{P})$ are nontrivial. Let $U \subseteq X \subseteq P$, where \bar{X} is either $\mathbf{J}(\bar{P})$ or $\mathbf{Z}(\bar{P})$, and observe that $X > U$. Also, $X \triangleleft P$, and so $P \subseteq \mathbf{N}_G(X)$ and $|\mathbf{N}_G(X)|_p = |P| = |\mathbf{N}_G(U)|_p$. But $|X| > |U|$, and so by the choice of U , we see that X cannot be one of the bad subgroups whose

normalizers fail to have normal p -complements. Since X contains the kernel U of overbar, we have $\mathbf{N}_{\overline{G}}(\overline{X}) = \overline{\mathbf{N}_G(X)}$ by the correspondence theorem, and thus $\mathbf{N}_{\overline{G}}(\overline{X})$ has a normal p -complement. In particular, both $\mathbf{N}_{\overline{G}}(\mathbf{J}(\overline{P}))$ and $\mathbf{N}_{\overline{G}}(\mathbf{Z}(\overline{P}))$ have normal p -complements, and therefore, \overline{G} has a normal p -complement, as required.

Step 3. $\mathbf{O}_{p'}(G) = 1$.

Proof. Let $K = \mathbf{O}_{p'}(G)$, and write $\overline{G} = G/K$. Now \overline{P} is a Sylow p -subgroup of \overline{G} , and since K is a p' -group, overbar defines an isomorphism from P onto \overline{P} . It follows that $\overline{\mathbf{J}(P)} = \mathbf{J}(\overline{P})$ and $\overline{\mathbf{Z}(P)} = \mathbf{Z}(\overline{P})$. Then

$$\begin{aligned}\mathbf{N}_{\overline{G}}(\mathbf{J}(\overline{P})) &= \mathbf{N}_{\overline{G}}(\overline{\mathbf{J}(P)}) = \overline{\mathbf{N}_G(\mathbf{J}(P))} \quad \text{and} \\ \mathbf{C}_{\overline{G}}(\mathbf{Z}(\overline{P})) &= \mathbf{C}_{\overline{G}}(\overline{\mathbf{Z}(P)}) = \overline{\mathbf{C}_G(\mathbf{Z}(P))},\end{aligned}$$

where the second equality in each of these equations follows by Lemma 7.7. We conclude that $\mathbf{N}_{\overline{G}}(\mathbf{J}(\overline{P}))$ and $\mathbf{C}_{\overline{G}}(\mathbf{Z}(\overline{P}))$ have normal p -complements, and so if $|\overline{G}| < |G|$, it would follow that \overline{G} has a normal p -complement. This is not the case, however, since $K = \mathbf{O}_{p'}(G)$ and G fails to have a normal p -complement. Then $|\overline{G}|$ is not smaller than $|G|$, and so $K = 1$, as required.

Step 4. P is a maximal subgroup of G .

Proof. Suppose that $P \subseteq H < G$. Then $\mathbf{N}_H(\mathbf{J}(P))$ and $\mathbf{C}_H(\mathbf{Z}(P))$ have normal p -complements, and since H is not a counterexample, it follows that H has a normal p -complement K . Then $K \triangleleft H$ and $\mathbf{O}_p(G) \triangleleft H$, and since $K \cap \mathbf{O}_p(G) = 1$, we have $K \subseteq \mathbf{C}_G(\mathbf{O}_p(G)) \subseteq \mathbf{O}_p(G)$. (We are using Lemma 1.2.3, which applies because G is p -solvable by Step 2 and $\mathbf{O}_{p'}(G) = 1$ by Step 3.) Then $K = 1$ and H is a p -group, and hence $H = P$.

Step 5. $\mathbf{C}_G(\mathbf{Z}(P)) = P$.

Proof. Certainly, $\mathbf{C}_G(\mathbf{Z}(P)) \supseteq P$. But $\mathbf{C}_G(\mathbf{Z}(P)) < G$ since G fails to have a normal p -complement, and thus $\mathbf{C}_G(\mathbf{Z}(P)) = P$ by Step 4.

Step 6. The normal p -complement of G/U is abelian.

Proof. First, recall that G/U has a normal p -complement L/U by Step 2. Suppose that $U \subseteq X \subseteq L$, where X is normalized by P . Then PX is a group, and hence by Step 4, either $PX = P$ or $PX = G$. If $PX = P$, then $|X|$ is a power of p , and so $X = U$. If, on the other hand, $PX = G$, then $|G : X|$ is a power of p , and hence $X = L$. It follows that no nonidentity proper subgroup of $\overline{L} = L/U$ is P -invariant.