Another Look at Sylow's Third Theorem

EUGENE SPIEGEL
University of Connecticut
Storrs, Connecticut 06269
spiegel@math.uconn.edu

Among the results that Sylow showed in his famous 1872 paper [**12**] is what is now usually called Sylow's third theorem.

If G is a finite group of order $|G| = p^n m$ where p is a prime, n is a positive integer, and p and m are relatively prime, then the number, N_p, of subgroups of G of order p^n satisfies $N_p \equiv 1$ (modulo p).

This result is among the arsenal of tools that every first year algebra student obtains. A group where the order of every element is a power of p is called a p-group; a p-Sylow subgroup of G is a p-subgroup of G of maximal order p^n. The idea of Sylow's proof, which was originally stated in terms of permutation groups, is to look at the size of the equivalence classes obtained when all p-Sylow subgroups of G are conjugated by the elements of a fixed p-Sylow subgroup of G. The existence of a p-Sylow subgroup was needed for the proof of the third Sylow theorem, although the conclusion of the theorem certainly implies that there are p-Sylow subgroups. Using the Sylow results, Frobenius, in 1895 [**1**], proved a generalization: The number of subgroups of G of order p^s is congruent to 1 modulo p whenever $1 \leq s \leq n$.

Most current texts show the existence of a p-Sylow subgroup and prove Sylow's third theorem using arguments that involve a group acting on a set. This method of proof for the existence of a p-Sylow subgroup was due to Miller between 1910 and 1915 [**8, 9**], but, according to Jacobson [**9**, p. 83], was "forgotten until it was rediscovered" in 1959 by Wielandt [**14**]. Krull [**7**] showed how Wielandt's method could be used to obtain Frobenius' generalization, and Gallagher, in 1967 [**2**], simplified the argument to one that depends upon the order of G rather than the group itself. Illustrating the combinatorics of finite group actions is part of the motivation to use this method of proof both to demonstrate the existence of p-Sylow subgroups in a finite group and to determine their number.

In this note we offer another method to prove these results. Our combinatorial tool will be Möbius inversion on the lattice of subgroups of a finite group. We will see that an application of this method will easily lead to Frobenius' theorem, in fact, a generalization of it. Of course, part of the reason for presenting this proof is to highlight the method.

Möbius inversion In this section, which can be skimmed by those conversant with Möbius inversion, we present all needed facts about incidence algebras and Möbius inversion. Suppose X is a finite partially ordered set. We use standard interval notation in X, for example, $(x, y] = \{z \in X \mid x < z \leq y\}$. If X has a minimum or maximum element, it will be denoted by $\widehat{0}$ or $\widehat{1}$ respectively.

The incidence algebra, $I(X, \mathbb{C})$, of X over $\mathbb{C}$, where $\mathbb{C}$ is the complex field, is defined as $I(X, \mathbb{C}) = \{f : X \times X \to \mathbb{C} \mid f(x, y) = 0 \text{ if } x \not\leq y\}$ with the operations of addition and scalar multiplication defined in the usual way, while multiplication is defined by convolution, $fg(x, y) = \sum_{z \in [x,y]} f(x, z)g(z, y)$, for $f, g \in I(X, \mathbb{C})$. It is straightforward to check that $I(X, \mathbb{C})$ is both a vector space over $\mathbb{C}$ and a ring. This

makes it a $\mathbb{C}$-algebra with identity δ, where $\delta(x, x) = 1$ for $x \in X$, and $\delta(x, y) = 0$, if $x \neq y$. If $\phi \in I(X, \mathbb{C})$ is such that for each $x \in X$, $\phi(x, x)$ is a unit in $\mathbb{C}$, then ϕ is invertible. Its inverse will be seen to be $\phi' \in I(X, \mathbb{C})$, having the following properties: For $x \in X$, $\phi'(x, x) = (\phi(x, x))^{-1}$, while if $\phi'(u, v)$ has been given for any $|[u, v]| < |[x, y]|$, then $\phi'(x, y)$ satisfies

$$\phi'(x, y) = -\left(\sum_{z: x<z\leq y} \phi(x, z)\phi'(z, y)\right)\phi(x, x)^{-1}.$$

It is easy to verify that ϕ' is a right inverse for ϕ. Similarly, ϕ has a left inverse. It follows that an element, $\phi \in I(X, \mathbb{C})$, has an inverse if and only if $\phi(x, x)$ is a unit for each $x \in X$. In particular, the element $\zeta \in I(X, \mathbb{C})$ given by

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise} \end{cases}$$

is a unit whose inverse, μ, is called the Möbius function of the partially ordered set. When we need to specify which partially ordered set we are considering, we will let μ_X denote the Möbius function for the partially ordered set X.

The importance of the Möbius function is seen in the following result, which is known as the Möbius inversion theorem. A more general Möbius inversion theorem appears in the seminal paper of Rota [**10**]. Earlier versions of this theorem can be found, for example, in Weisner [**13**] or Hall [**4**].

THEOREM 1. *Let X be a finite partially ordered set and f a function from X to $\mathbb{C}$. If, for $x \in X$,*

$$g(x) = \sum_{y:\, x\leq y} f(y)$$

then

$$f(x) = \sum_{y:\, x\leq y} \mu(x, y)g(y).$$

Proof.

$$\begin{aligned}
\sum_{y:\, x\leq y} \mu(x, y)g(y) &= \sum_{y:\, x\leq y} \mu(x, y)\Big(\sum_{z:\, y\leq z} f(z)\Big) \\
&= \sum_{y:\, x\leq y} \mu(x, y)\Big(\sum_{z:\, y\leq z} \zeta(y, z)f(z)\Big) \\
&= \sum_{y:\, x\leq y} \sum_{z:\, y\leq z} \mu(x, y)\zeta(y, z)f(z) \\
&= \sum_{z:\, x\leq z} f(z) \sum_{y:\, x\leq y\leq z} \mu(x, y)\zeta(y, z) \\
&= \sum_{z:\, x\leq z} \delta(x, z)f(z) \\
&= f(x).
\end{aligned}$$

■

The following proposition presents three properties of the Möbius function, which are useful in its computation. The first of these, together with the condition $\mu(x, x) = 1$ for every $x \in X$, could be used as the definition of the Möbius function.

The third of these properties is a result of Weisner [**14**]. It applies only when X is a *lattice*, which is a partially ordered set in which every pair of elements, a, b, has a least upper bound, denoted $a \vee b$, and a greatest lower bound. We are particularly interested in the set of all subgroups of a finite group G, partially ordered by inclusion. This set, $\mathcal{L}(G)$, is a lattice with least element $\widehat{0} = \{e\}$ and greatest element $\widehat{1} = G$. If $A, B \in \mathcal{L}(G)$, then $A \vee B$ is the subgroup generated by A and B.

PROPOSITION 1. *Suppose X is a finite partially ordered set with Möbius function μ.*

(i) If $x < y \in X$, then $\sum_{z \in [x,y]} \mu(x, z) = 0$.

(ii) If $\phi : X \longrightarrow X'$ is an isomorphism of partially ordered sets, then $\mu_X(x, y) = \mu_{X'}(\phi(x), \phi(y))$, for any $x, y \in X$.

(iii) If X is a finite lattice and $a \neq \widehat{0} \in X$, then

$$\sum_{x:\ x \vee a = \widehat{1}} \mu(\widehat{0}, x) = 0.$$

Proof.

(i) $0 = \delta(x, y) = \mu\zeta(x, y) = \sum_{z \in [x,y]} \mu(x, z)$.

(ii) One can extend ϕ to a $\mathbb{C}$-isomorphism $\phi' : I(X, \mathbb{C}) \longrightarrow I(X', \mathbb{C})$. Since $\phi'(\zeta_X) = \zeta_{X'}$ then $\phi'(\mu_X) = \mu_{X'}$.

We check (iii) by induction on $|X|$. If $|X| = 2$, then $a = \widehat{1}$, and as $x \vee \widehat{1} = \widehat{1}$ for any $x \in X$, the result follows from (i). Now assume the result for lattices of smaller cardinality than that of X. We have

$$0 = \sum_{x \in X} \mu(\widehat{0}, x) = \sum_{x \in [\widehat{0}, a]} \mu(\widehat{0}, x) + \sum_{b:\ a < b < \widehat{1}} \left(\sum_{x:\ x \vee a = b} \mu(\widehat{0}, x) \right) + \sum_{x:\ x \vee a = \widehat{1}} \mu(\widehat{0}, x).$$

Looking at the right-hand side of this equation, we note that the first sum is 0 by (i) and that each summand in the second sum is 0 by the inductive assumption. The result then follows. ∎

We now use these properties to calculate the Möbius function of the lattice, $\mathcal{L}(G)$, of a finite p-group G. The answer was obtained by Kratzer and Thévenez, [**6**], using a different approach. The computation of $\mu_{\mathcal{L}(G)}(\{e\}, G) = \mu_{\mathcal{L}(G)}(\widehat{0}, \widehat{1})$ is a result of P. Hall [**3**]. For notation, if A is a subgroup of G, let $N(A) = N_G(A)$ denote its normalizer and write Z_p^k for the direct sum of k copies of a cyclic group of order p.

THEOREM 2. *Let G be a finite p-group, $\mathcal{L}(G)$ the lattice of subgroups of G, and μ its Möbius function. If $A, B \in \mathcal{L}(G)$, then*

$$\mu(A, B) = \begin{cases} (-1)^k p^{\binom{k}{2}} & \text{if } A \subseteq B \subseteq N(A) \text{ and } B/A \simeq Z_p^k \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $|G| = p$ the result is easily verified. Continue by induction on the order of G, and assume the result for groups of smaller order than that of G. Suppose $|G| = p^n$ and let $A \subseteq B$ be subgroups of G. Since the collection of subgroups between A and B is both an interval in $\mathcal{L}(B)$ and in $\mathcal{L}(G)$, by the induction assumption it is sufficient to verify the result when $B = G$. Suppose, first, that A is not a normal subgroup of G. Then $A \subset N(A) \subset G$ and

$$0 = \sum_{X \in [A,G]} \mu(A, X)$$

$$= \sum_{X \in [A,N(A)]} \mu(A, X) + \sum_{X \in [A,G] \setminus [A,N(A)]} \mu(A, X).$$

The first summation in the last line is zero by Proposition 1(i), and each summand with $X \neq G$ in the second summation is zero by our inductive assumption. Hence $\mu(A, G) = 0$, verifying the result in this case.

Suppose now that A is a normal subgroup of G. If $\{e\} \neq A$, by Proposition 1(ii), $\mu_{\mathcal{L}(G)}(A, G) = \mu_{\mathcal{L}(G/A)}(\widehat{0}, \widehat{1})$ and, as $|G/A| < |G|$, the result again follows by the inductive assumption. Hence, to complete the proof, all that remains is to check the result in the case $\mu(\widehat{0}, \widehat{1})$. Let H be a normal subgroup of G of order p. If X is any proper subgroup of G with $G = H \vee X$, then $G = HX$ and $X \simeq G/H$. We conclude that all such subgroups, X, are isomorphic, $G \simeq H \oplus X$, and, from Proposition 1(iii), we have $\mu_{\mathcal{L}(G)}(\widehat{0}, \widehat{1}) = -\lambda \mu_{\mathcal{L}(G/H)}(\widehat{0}, \widehat{1})$, where λ is the number of proper subgroups, X, of G with $H \vee X = G$.

Suppose $\mu_{\mathcal{L}(G)}(\widehat{0}, \widehat{1}) \neq 0$. It follows that $\mu_{\mathcal{L}(G/H)}(\widehat{0}, \widehat{1}) \neq 0$ and, by the inductive assumption, $G/H \simeq Z_p^{n-1}$ and $\mu_{\mathcal{L}(G/H)}(\widehat{0}, \widehat{1}) = (-1)^{n-1} p^{\binom{n-1}{2}}$. Further $G \simeq Z_p^n$. To calculate λ, it is convenient to consider G as an n-dimensional vector space over a field of p elements. We are then looking for subspaces, X, of dimension $n - 1$ which do not contain a fixed 1-dimensional subspace, H. The total number of ordered bases for all such X is $(p^n - p)(p^n - p^2) \cdots (p^n - p^{n-1})$, where, for example, the first factor comes from choosing the first basis vector of X to be any vector of G other than those in H. By a similar argument, each $n - 1$ dimensional space has $(p^{n-1} - 1)$ $(p^{n-1} - p) \cdots (p^{n-1} - p^{n-2})$ ordered bases. Hence $\lambda = p^{n-1}$, it being the quotient of these two numbers. We conclude $\mu_{\mathcal{L}(G)}(\widehat{0}, \widehat{1}) = -p^{n-1}(-1)^{n-1} p^{\binom{n-1}{2}} = (-1)^n p^{\binom{n}{2}}$. This is the desired value. ∎

The Sylow theorem We now obtain the result we sought by applying Möbius inversion. Before we give the main result, it is convenient to handle one special case separately in a lemma. The lemma, like the theorem following it, only requires Cauchy's theorem.

LEMMA 1. *Let G be a finite group and p a prime dividing the order of G. The number, $\mathbf{K}$, of subgroups of G of order p is congruent to 1 modulo p.*

Proof. Suppose $\mathcal{C}(p)$ is the collection of all subgroups of order p in G and P is a p-subgroup of G of maximum order. We let P act on $\mathcal{C}(p)$ by conjugation. For $A \in C(p)$, $\mathrm{Orb}(A) = \{gAg^{-1} \mid g \in P\}$ is the orbit of A. It is not difficult to check that $|\mathrm{Orb}(A)|$ is equal to the index of $\mathrm{Stab}(A) = \{g \in P \mid gAg^{-1} = A\}$ in P. This index, which is a power of p, is greater than 1 unless $P \subseteq N(A)$. Hence

$$|\mathcal{C}(p)| = \mathbf{K} \equiv \overline{\mathbf{K}} \text{ (modulo } p),$$

where $\overline{\mathbf{K}} = |\{A \in \mathcal{C}(p) \mid P \subseteq N(A)\}|$. When $A \in \mathcal{C}(p)$ with $P \subseteq N(A)$, then AP is a p-subgroup of $N(A)$ that contains P. By maximality, $A \subseteq P$. We conclude that $\overline{\mathbf{K}}$ is the number of normal subgroups of order p in P. By having the elements of P act by conjugation on themselves, one can check that any nontrivial normal subgroup of a p-group, in this case P, has a nontrivial intersection with its center, $Z(P)$. In particular, any normal subgroup of P of order p is in $Z(P)$, and in fact in $\mathrm{Soc}_p(Z(P)) = \{x \in Z(P) \mid x^p = 1\} \simeq Z_p^k$, for some integer $k \geq 1$. Conversely, each nonidentity element of $\mathrm{Soc}_p(Z(P))$ generates a normal subgroup of P of order p. Since any subgroup

of order p has $p-1$ nonidentity elements, we have $\mathbf{K} \equiv \overline{\mathbf{K}} = (p^k - 1)/(p-1) = 1 + p + \cdots + p^{k-1} \equiv 1$ (modulo p), establishing the lemma. ∎

We can now give a generalized Sylow's third theorem as an application of Möbius inversion.

THEOREM 3. *Let G be a finite group of order exactly divisible by p^n, where p is a prime and n a positive integer. Suppose m, t are nonnegative integers with $m \leq t \leq n$ and H is a subgroup of G of order p^m. Then the number of subgroups of G of order p^t, each of which contains H, is congruent to 1 modulo p.*

Proof. We prove the result by induction on $s = n - m$. If $s = 0$ the result is immediate. Suppose that we know the result for values smaller than $s = n - m$. Of course, if $t = m$ the result is clear. We can thus assume that $m < t \leq n$. Let $\mathcal{C}_H(t)$ be the collection of subgroups of G of order p^t, each of which contains H. Define the function f on $\mathcal{L}(G)$ by

$$f(S) = \begin{cases} 1 & \text{if } S \in \mathcal{C}_H(t) \\ 0 & \text{otherwise,} \end{cases}$$

and the function g on $\mathcal{L}(G)$ by

$$g(T) = \sum_{S \in \mathcal{L}(G): S \supseteq T} f(S).$$

Then $g(T)$ is the number of elements of $\mathcal{C}_H(t)$ that contain T. We wish to compute $g(H) = |\mathcal{C}_H(t)|$. Using Möbius inversion in $\mathcal{L}(G)$ we obtain

$$f(T) = \sum_{S:\ S \supseteq T} \mu(T, S) g(S).$$

In particular,

$$f(H) = \sum_{S:\ S \supseteq H} \mu(H, S) g(S).$$

Of course, $g(S) = 0$ unless S is a p-group. Furthermore, by Theorem 2, when S is a p-group containing H, $\mu(H, S)$ is divisible by p whenever the index of H in S is divisible by p^2. Hence,

$$0 \equiv g(H) - \sum_{S \subseteq N(H):\ S/H \simeq Z_p} g(S) \text{ (modulo } p).$$

That is

$$|\mathcal{C}_H(t)| \equiv \sum_{S \subseteq N(H):\ S/H \simeq Z_p} g(S) \text{ (modulo } p).$$

Each S in this summand is of order p^{m+1}. As $n - (m+1) < s$, by the induction assumption, $g(S) \equiv 1$ (modulo p). Further, the number of such S coincides with the number of subgroups of order p in $N(H)/H$. By the lemma, this number is congruent to 1 modulo p. We conclude that $|\mathcal{C}_H(t)| \equiv 1$ (modulo p), which establishes the result. ∎

If in the previous theorem we let $m = 0$ and $t = n$ then we obtain the first and third Sylow theorems. The Frobenius theorem is obtained by letting $m = 0$.

Note added in proof: The author thanks Keith Conrad for bringing to my attention another paper of L. Weisner, "Some properties of prime-power groups," *Trans. AMS* **38** (1835), 485–492, in which Weisner proves Theorem 3 in the case that G is a p-group.

REFERENCES

1. G. Frobenius, Verallgemeinerung des Sylow'schen satzes, *Berliner Sitzungsberichte* (1895), 981–993.
2. P. X. Gallagher, On the p-subgroups of a finite group, *Arch. der Math.* **18** (1967), 469.
3. P. Hall, A contribution to the theory of groups of prime-power order, *Proc. London Math. Soc.* **36** (1933), 29–95.
4. ———, The Eulerian functions of a group, *Quart. J. Math.* **7** (1936), 134–151.
5. N. Jacobson, *Basic Algebra I*, W. H. Freeman and Co., New York, 1985.
6. C. Kratzer and J. Thévenez, Fonction de Möbius d'un groupe fini et anneau de Burnside, *Commen. Math. Helvetici* **59** (1984), 425–438.
7. W. Krull, Über die p-untergruppen endlicher gruppen, *Arch. Math.* **12** (1961), 1–6.
8. G. A. Miller, Extensions of two theorems due to Cauchy, *Bull. AMS* **16** (1910), 510–513.
9. ———, A new proof of Sylow's theorem, *Ann. of Math.* **16** (1915), 169–171.
10. G.-C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, *Wahr Scheinlichkeits Thoerie und Verw. Gebiete* **2** (1964), 340–368.
11. E. Spiegel and C. O'Donnell. *Incidence Algebras*, Marcel Dekker, Inc., New York, 1997.
12. L. Sylow, Théorèmes sur les groupes de substitutions, *Math. Ann.* **5** (1872), 584–594.
13. L. Weisner, Abstract theory of inversion of finite sets, *Trans. AMS* **38** (1935), 474–484.
14. H. Wielandt, Ein beweiss für die existenz der Sylowgruppen, *Arch. Math.* **10** (1959), 401–402.

Continued from page 246

In terms of total score (out of a maximum of 252), the highest ranking of the 82 participating teams were as follows:

Bulgaria	227	Turkey	133
China	211	Japan	131
USA	188	Hungary	128
Vietnam	172	United Kingdom	128
Russia	167	Canada	119
Korea	157	Kazakhstan	119
Romania	143		

The 2003 USAMO was prepared by Titu Andreescu (Chair), Zuming Feng, Kiran Kedlaya, and Richard Stong. The Team Selection Test was prepared by Titu Andreescu and Zuming Feng. The MOSP was held at the University of Nebraska-Lincoln. Zuming Feng (Academic Director), Gregory Galperin, and Melanie Wood served as instructors, assisted by Po-Shen Loh and Reid Barton as junior instructors, and Ian Le and Ricky Liu as graders. Kiran Kedlaya served as guest instructor.

For more information about the USAMO or the MOSP, contact Steven Dunbar at sdunbar@math.unl.edu.