A History of Lagrange's Theorem on Groups

RICHARD L. ROTH
University of Colorado
Boulder, CO 80309-0395

Introduction

In group theory, the result known as Lagrange's Theorem states that for a finite group G the order of any subgroup divides the order of G. However, group theory had not yet been invented when Lagrange first gave his result and the theorem took quite a different form. Lagrange's Theorem first appeared in 1770–71 in connection with the problem of solving the general polynomial of degree 5 or higher, and its relation to symmetric functions. It was also anticipated by some results in number theory a few years earlier.

In this article, we explain the historical setting of Lagrange's approach, and follow this train of thought into the twentieth century. Some general references on the history of group theory are Israel Kleiner's article "The Evolution of Group Theory: A Brief Survey" in this MAGAZINE [**22**] and H. Wussing's book, *The Genesis of the Abstract Group Concept* [**31**].

Preliminary Discussion

Let us first review Lagrange's Theorem and its proof, as well as some other results relevant to our discussion. Recall that the *order* of a finite group is the number of elements in the group.

THEOREM A *(Lagrange's Theorem): Let G be a group of order n and H a subgroup of G of order m. Then m is a divisor of n.*

Sketch of Proof. Suppose we list the elements of G in a rectangular array as follows: Let the top row be the list of the m elements of H: $a_1 = e, a_2, \ldots, a_m$. If b is some element of G not in H then let the second row consist of the elements $ba_1, ba_2, \ldots, ba_m$. If there is an element c not in the first two rows then the next row will be $ca_1, ca_2, \ldots ca_m$. This is continued until the elements of G are exhausted. One must then check that the elements in any row are distinct and that no two rows have an element in common. It follows that $n = km$ where k is the number of rows. ■

Early proofs of Lagrange's Theorem generally involved this "rectangular array" explicitly or implicitly. Note that the rows of the rectangular array are simply the left cosets of H in G. In fact, most current texts use the language of cosets to prove this theorem. One must show that the set of left cosets (or right cosets) forms a partition of the group G and that each coset has the same number of elements as H. Note that as an elementary consequence of Lagrange's Theorem we have that the number of cosets of H in G divides the order of G as well. This number is called the index and is denoted $[G : H]$.

Frequently in algebra textbooks, the little theorem of Fermat is proved as a corollary.

FERMAT'S LITTLE THEOREM: *If p is a prime and b is relatively prime to p, then $b^{p-1} \equiv 1 \pmod{p}$.*

Proof. The nonzero elements of $\mathbb{Z}/p\mathbb{Z}$ form a group of order $p-1$ under multiplication, called $(\mathbb{Z}/p\mathbb{Z})^*$. If the congruence class $\overline{b}$ in $(\mathbb{Z}/p\mathbb{Z})^*$ has order m, then it generates a cyclic subgroup of order m. By Lagrange's Theorem, m divides $p-1$; thus $(\overline{b})^{p-1} = (\overline{b}^m)^k = \overline{1}$, where $p-1 = mk$, and the theorem follows. ∎

Also relevant is Theorem B, given below. As we will see, it might be just as appropriate to call this Lagrange's Theorem.

THEOREM B: *Let G be a finite group acting by permutations on a finite set S. Then the size of any orbit is a divisor of the order of G.*

Proof. Let b be an element of some particular orbit and H be the subgroup of G stabilizing b. If c is another element in this orbit then for some τ in G, $\tau b = c$. If $\sigma \epsilon H$ then $(\tau\sigma)b = \tau(\sigma b) = \tau b = c$ (in our notation it is the element σ in the stabilizer of b that acts first) and it is easily seen that the coset τH consists of precisely the elements mapping b onto c. Thus the elements of the orbit are in one-to-one correspondence with the left cosets of H, hence the size of the orbit equals the index $[G : H]$, which, by our consequence of Theorem A above, must divide G. ∎

Lagrange's version of the theorem

In 1770–71, Lagrange published a landmark work on the theory of equations, "*Reflexions sur la résolution algébrique des équations*" [**23**]. (Note: Katz gives a useful discussion of this paper in section 14.2.6 of [**21**].) His concern was the question of finding an algebraic formula for the roots of the general 5th degree polynomial and more generally for the nth degree polynomial for $n > 4$. The quadratic formula had, of course, been known for a very long time and the cubic and quartic equations had been solved in the sixteenth century by algebraists of the Italian school. However, for polynomials of degree greater than four this had remained an open problem for two centuries. Lagrange observed that the solutions for the cubic and quartic equations involved solving supplementary "resolvent" polynomials of lower degree whose coefficients were rational functions of the coefficients of the original polynomial. He found that the roots of these auxiliary equations were in fact "functions" of the roots of the original equation that took on a small number of values when the original roots were permuted in the formulas for these functions.

For example, the quartic was solved using a cubic resolvent polynomial whose roots could be written as $\frac{x_1x_2+x_3x_4}{2}$, $\frac{x_1x_3+x_2x_4}{2}$ and $\frac{x_1x_4+x_2x_3}{2}$ where x_1, x_2, x_3, x_4 were the roots of the original polynomial. If the four roots x_1, x_2, x_3, x_4 are permuted in all 24 possible ways, only these three different "values" typically occur. For convenience, we will omit the denominator 2 in what follows. We list below the result of operating on the function $x_1x_2 + x_3x_4$ by the 24 different permutations of the four variables. The stabilizer of the function is a group of 8 elements and the first row shows the seven other values that are equal to the original one. Underneath each value is the corresponding permutation. The second row shows the set of eight values arising from a different way of combining the four roots, all equal to $x_1x_3 + x_2x_4$; similarly for the third row. Underneath each value is the corresponding permutation. We note that in the second and third lines we have permuted the positions of the variables in the same manner as was done in the first line. This corresponds to the step in the proof of Theorem B above where the stabilizing permutation σ must act before the permutation τ that changes the object. For example, referring to the first two equal functions in

$x_1x_2 + x_3x_4$	$x_2x_1 + x_3x_4$	$x_3x_4 + x_1x_2$	$x_1x_2 + x_4x_3$	$x_4x_3 + x_1x_2$	$x_3x_4 + x_2x_1$	$x_4x_3 + x_2x_1$	$x_2x_1 + x_4x_3$
id	(12)	(13)(24)	(34)	(1423)	(1324)	(14)(23)	(12)(34)
$x_1x_3 + x_2x_4$	$x_3x_1 + x_2x_4$	$x_2x_4 + x_1x_3$	$x_1x_3 + x_4x_2$	$x_4x_2 + x_1x_3$	$x_2x_4 + x_3x_1$	$x_4x_2 + x_3x_1$	$x_3x_1 + x_4x_2$
(23)	(23)(12) = (132)	(23)(13)(24) = (1243)	(23)(34) = (234)	(23)(1423) = (143)	(23)(1324) = (124)	(23)(14)(23) = (14)	(23)(12)(34) = (1342)
$x_1x_4 + x_2x_3$	$x_4x_1 + x_2x_3$	$x_2x_3 + x_1x_4$	$x_1x_4 + x_3x_2$	$x_3x_2 + x_1x_4$	$x_2x_3 + x_4x_1$	$x_3x_2 + x_4x_1$	$x_4x_1 + x_3x_2$
(243)	(243)(12) = (1432)	(243)(13)(24) = (123)	(243)(34) = (24)	(243)(1423) = (13)	(243)(1324) = (1234)	(243)(14)(23) = (134)	(243)(12)(34) = (142)

rows 1 and 2, we have $\sigma = (1\ 2)$, $\tau = (2\ 3)$ and $\tau\sigma = (2\ 3)(1\ 2) = (1\ 3\ 2)$. We thus have a 3 by 8 rectangular array. The 3 "values" multiplied by 8 gives $24 = 4! = |S_4|$.

The 4th degree polynomial was solvable because there was a "function" of 4 variables which took on 3 "values" when the 4 variables are permuted in all $24 = 4!$ ways. That is, these 3 "values" were the roots of a cubic polynomial (which it was known how to solve); these roots could be used to modify the original 4th degree polynomial so that it would factor into quadratic polynomials. Using the theory of symmetric functions, Lagrange proved that if a rational function of the n roots of a general polynomial of degree n takes on r "values" under the action of all $n!$ permutations, then the function will be a root of a polynomial of degree r whose coefficients are rational functions of the coefficients of the original equation.

Thus Lagrange reasoned that to solve a 5th degree polynomial, one should try to find a function in 5 variables that takes on 3 (or 4) different typical "values" when the variables are permuted in all 5! ways. This would lead to a cubic (or quartic) resolvent that might help to solve the original equation. A similar approach might apply for solving equations of degree n for n greater than 5.

Lagrange was unable to determine if such functions exist. But he did come up with, in essence, the following theorem.

THEOREM C: THEOREM OF LAGRANGE: *If a function $f(x_1, \ldots, x_n)$ of n variables is acted on by all $n!$ possible permutations of the variables and these permuted functions take on only r distinct values, then r is a divisor of $n!$.*

In fact, Lagrange stated his theorem in terms of the degree of the corresponding resolvent equation. Also, we note that if $n = 5$, then 3 and 4 are both divisors of $n!$, so the theorem of Lagrange doesn't answer the previous question as to whether or not a cubic or quartic resolvent exists for the 5th degree equation.

Lagrange's proof of Theorem C consisted essentially of discussing some special cases; it is interesting to note that the treatment that he gave for the first case was partly wrong, although it did give the correct idea for a proof. He said: let us suppose that a function satisfies $f(x', x'', x''', x^{iv} \ldots) = f(x'', x''', x', x^{iv} \ldots)$. Such a function satisfies $f(x^{iv}, x''', x', x'', \ldots) = f(x''', x', x^{iv}, x'' \ldots.)$, because we have permuted the first 3 variables in the same way; hence, he said, all the values will match up in pairs and the possible number of distinct values will be reduced to $\frac{n!}{2}$. However, the permutation involved is actually a cycle of length 3, so in fact in this example the number of values would be divided by 3, and not by 2.

Lagrange then said that if the original function remains the same under 3 or 4 or a larger number of permutations, then the other values will also have that property, and the total number of distinct values will be $\frac{n!}{3}$ or $\frac{n!}{4}$, etc.

Thus Lagrange's original Theorem C might be regarded as a special case of Theorem B, where the group G is the symmetric group S_n, the set S is the set of functions (or formulas) involving n variables formed by all permutations of the n variables and the group action is that which arises from permuting the variables in these functions.

Coincidentally, in 1771 Vandermonde also wrote a paper ([**28**]) on the theory of equations that took an approach similar to Lagrange's. The alternating function $\prod_{1\le i<j\le n}(x_i - x_j)$ takes on exactly two values when the variables are permuted. In Vandermonde's paper we find this function for the case when $n = 3$. It is used today (for arbitary n) in contemporary abstract algebra books to study even and odd permutations; the set of permutations that stabilize it forms the alternating group. As can be seen from our paper, the use of polynomial functions is historically very much a part of group theory. The alternating function $\prod_{1\le i<j\le n}(x_i - x_j)$ is actually equal to the Vandermonde determinant, and it is probable that the origin of the name comes from this reference, although Vandermonde (who elsewhere did do important work on

determinants) did not express it as a determinant. For the case of $n = 5$, Vandermonde gave a different example of a function taking on two values under permutations, and expressed the opinion that there does not exist such a function of five variables taking on either three or four values.

Some later developments related to Lagrange's work

Several decades later, Paolo Ruffini made further progress in Lagrange's approach to solving polynomial equations. His book of 1799 [**25**] included an informal proof (by example) of Lagrange's Theorem C. Further, Ruffini showed that there does not exist any function of 5 variables taking on three values or four values. Thus the "converse" to Theorem C is false. It seems appropriate to call this "Ruffini's Theorem." In modern terminology this shows that the converse to Lagrange's Theorem (Theorem A) is false because the symmetric group on 5 letters of order 120 has no subgroup of order 40 or 30.

Ruffini also claimed to have proved (as a consequence) that the 5th degree equation (and in general the nth degree for $n \geq 5$) was not solvable. His work drew much criticism and even though he published several more versions, his proof is generally regarded as incomplete. (A more satisfactory proof would be given later by Abel in the 1820s.) A friendly response came from Abbati in 1802 [**1**] who gave some suggestions to improve Ruffini's proof. Abbati's note included an extensive and thoughtful proof of Lagrange's Theorem C. According to Heinrich Burkhardt [**5**] this was the first time a complete proof was given. It resembled the proof of Theorem A given above, and presented a rectangular array of terms (as illustrated in our earlier discussion of the function $x_1x_2 + x_3x_4$). The given function was acted on by the $n!$ permutations. The first row gave the list of functions (arising from permutations) that are equal to the original function. The next row started with a value that was not in the first row and consisted of all subsequent permutations of this function. Because the modern notation used in the proof of Theorem B was not available, a lengthier discussion was used. Abbati took some care to explain that what mattered was the positions of the variables in the function and the way they were combined. Then, as in the proof of Theorem A, each row was shown to have the same number of elements, with no two rows having any element in common.

Further developments were obtained with Cauchy's important 1815 paper [**7**], whose title roughly translated into English reads "*Memoir on the number of values that a function can acquire (when one permutes in all possible ways the quantities it contains).*" This paper launched permutation group theory as an independent topic even though the notion of a group did not appear in it. The paper was not concerned directly with the theory of equations. Cauchy included a proof of Theorem C similar to Abbati's, pointing out that the permutations fixing a function are to be applied to the positions of the variables and not their indices. He went on to generalize Ruffini's theorem as follows: If the number of values of a non-symmetric function of n quantities is less than the largest prime divisor p of n then it must be 2. In the 1820s Abel cited and used this theorem from Cauchy's paper specifically for the special case that $n = 5$ in his work on the unsolvability of the quintic (see [**2**, p. 31] of Oeuvres). Although Ruffini was not explicitly mentioned in Abel's paper, Abel's proof of the unsolvability of the 5th degree polynomial thus relied indirectly on the work of Ruffini.

Galois introduced the term *group* for permutation groups in a paper on solutions of polynomials by radicals in 1831 [**14**, pp. 35–36]. He didn't explicitly mention either form of Lagrange's Theorem in any of his papers, but in the famous letter written the night before his death [**15**], he did include the suggestive equation for the coset

decomposition

$$G = H + HS + HS' + \cdots. \tag{1}$$

These works were not widely known until they were published in Liouville's Journal in 1846.

Theory of permutation groups

Almost 30 years after Cauchy's 1815 paper, Cauchy again took up the subject of permutation groups. His paper of 1844 [**8**] did not take the "values of a function" approach; rather it dealt directly with permutation groups. Keeping the older use of the word "permutation" to refer to an ordered arrangement of symbols, he used the term "substitution" to refer to a permutation, and "system of conjugate substitutions" (système de substitutions conjugées) to refer to a permutation group. This was defined here and in later works on the subject merely as a set of permutations closed under composition. This is of course sufficient to define a group: composition is associative, and since we are dealing with finite permutations, the existence of the identity and inverse operations will be necessarily implied. (It is perhaps unfortunate that in many modern books the definition of group only lists three axioms, namely associativity and the existence of the identity and inverse operations. The closure property, which was really the original property used to define permutation groups, is hidden in the term "operation" or "binary operation.") Cauchy then proved the following theorem (translated into English): "The order of a system of conjugate substitutions on n variables is always a divisor of the number N of arrangements which one can form with these variables" [**8**, p. 207]; i.e. the order of a subgroup of the symmetric group S_n is a divisor of $n!$. Thus we now had Theorem A (Lagrange's Theorem) for the case that G is the symmetric group.

Cauchy's long paper of 1844 was followed by a series of shorter papers over the next couple of years that further developed the theory of permutation groups. They included showing the connection between Theorem A and Theorem C. In [**9**], he showed that the set of permutations fixing a function forms a permutation group (i.e. a subgroup of S_n, which today we call the stabilizer of that function) and that, conversely, for any such subgroup H of S_n there is a function whose stabilizer is precisely that subgroup. Given a subgroup H of S_n, we can build a function of n variables whose stabilizer is H in the following way. Let $s = a_1x_1 + \cdots + a_nx_n$, where $a_1, \ldots, a_n$ are distinct numbers. Let $s_1, s_2, \ldots, s_t$ be the images under the t elements of the subgroup H (note we may assume $s = s_1$); then consider the product $s_1s_2 \ldots s_t$. With appropriate assumptions on the coefficients $a_1, a_2, \ldots, a_n$, this will be a function whose stabilizer is precisely H. Cauchy required these coefficients to be nonzero elements whose sum is not zero and his proof is a bit unclear, but if we require instead that the coefficients be a set of nonzero integers whose greatest common divisor is 1, then the result can be shown using the unique factorization property of polynomial rings over a unique factorization domain.

Another example of a function whose stabilizer is H, also given by Cauchy, arises by taking $s = x_1x_2^2x_3^3 \ldots x_n^n$, and defining $s_1 = s, s_2, \ldots, s_t$ to be the images of s under the subgroup H; then the sum $s_1 + \cdots + s_t$ is a function whose stabilizer is H. In a later note [**10**] he showed that if a function takes on m distinct values and the subgroup fixing the function has order M then $mM = n!$. Because every subgroup is the stabilizer of some function, this is a kind of "hybrid form" of Lagrange's Theorem in which Cauchy combined Theorem C and Theorem A for the case of $G = S_n$. In this

note he also laid the foundations for the idea of a group acting on a set (as in Theorem B above). In particular he showed that a group of permutations on the n variables $x_1, \ldots, x_n$ could also be regarded as acting by permutations on the set of functions that arise from a particular function of those variables under permutations of the variables.

The next step in the development of Lagrange's Theorem was to see that Theorem A holds for any finite permutation group G. This result may be found in Camille Jordan's thesis, published in 1861 [**19**]. This is a rather lengthy and technical paper on permutation groups. In the introduction he cited Lagrange's Theorem, in the "hybrid" form just mentioned, calling it a theorem due to Lagrange and crediting his proof to Cauchy. Forty-five pages later in the midst of a complicated counting argument, he mentioned that he would need a generalization of the Lagrange's Theorem and proved that (in modern language) the order of a subgroup of any permutation group divides the order of the group. Lagrange's Theorem then appeared in this form in the 3rd edition of Serret's important algebra text *Cours d'Algébre supérieure* published in 1866 [**27**] (in a later chapter, the theory is applied to the topic of the number of values of a function of n variables). And it is also found in Jordan's influential 1871 book *Traité des substitutions et des èquations algébriques* [**20**].

Lagrange's Theorem C is essentially Theorem A for the case where $G = S_n$, as we have seen above: each function has its stabilizing subgroup of S_n and each subgroup of S_n may be regarded as the stabilizer of an appropriate function. The extension of Theorem A to the case of G, an arbitrary permutation group, was more general and might seem to have no analogue in terms of functions. However, it is interesting to note that in Netto's book of 1882 [**24**] he did give a translation into the function approach. "Lehrsatz VI" in chapter 3 states: If φ and ψ are two functions of the same n variables and the permutations leaving φ fixed also leave ψ unchanged, then the number of values taken on by φ is a multiple of the number of values taken on by ψ.

Other directions in group theory

While the theory of permutation groups played a major role in the development of general group theory, there were a number of other important sources of group theory including geometry and number theory. One connection to number theory, as we noted in the introduction, is Fermat's Little Theorem. Euler gave several proofs of this theorem. Of interest here is his paper whose title translated into English is "Theorems on residues obtained by division of powers," written in 1758–59 and published in 1761 [**11**]. In it he gave a proof along the lines indicated in the beginning of this article. He proved Lagrange's Theorem in essentially the usual way (the rectangular array) for the case that G is $(\mathbb{Z}/p\mathbb{Z})^*$ (the multiplicative group of the integers relatively prime to p, modulo p) and H is the cyclic subgroup generated by $\overline{b}$. Thus one could argue that in some sense, Lagrange's Theorem appeared 10 years before Lagrange's work. The theorem of Fermat generalizes in two directions; if, instead of $(\mathbb{Z}/p\mathbb{Z})^*$, we consider $(\mathbb{Z}/n\mathbb{Z})^*$, the classes of integers relatively prime to n, then $b^{\phi(n)} \equiv 1 \bmod n$ (where φ is the Euler function). This was shown by Euler in a paper written 1760–61, published in 1763 [**12**]. The other direction is to regard $(\mathbb{Z}/p\mathbb{Z})^*$ as a finite field and to generalize this to Galois fields $GF(p^n)$, in which case we have $x^{p^n-1} = 1$, and this was done by Galois in 1830 [**13**]. Again, in both cases, the proofs implicitly involved a rectangular array approach to a special case of Lagrange's Theorem.

The development of the abstract approach to groups is discussed in [**22**] and [**31**]: abstract groups generalized permutation groups, various groups arising in number theory (including those arising from modular arithmetic, as we've seen in the previous paragraph) and geometrical groups. The abstract approach to groups caught on in the

1880s. Of course Lagrange's Theorem (Theorem A) and its proof were by then easily adapted to abstract groups. It is hard to pinpoint the first abstract version; but, for example, in a paper of Hölder in 1889 [**18**] on Galois theory, an abstract definition for finite groups was given. Lagrange's Theorem was proved by merely displaying the familiar rectangular array. The theorem was not identified by Lagrange's name, but the rectangular array is credited to Cauchy's 1844 paper.

In 1895–96, Weber's *Lehrbuch der Algebra* [**30**] was published in two volumes. This became the standard text for modern algebra for the next few decades. Volume 1 (1895) was the more elementary of the two volumes. As for group theory it treated only permutation groups. It included Theorem A for permutation groups. The theorem was credited to Cauchy. The proof, however, was done using the terminology of cosets ("nebengruppe"). It was shown that two cosets are either equal or disjoint, and stated that the size of any coset equals the order of the subgroup. Following the proof, the symbolic equation

$$P = Q + Q\pi_1 + Q\pi_2 + \cdots + Q\pi_{j-1} \tag{2}$$

was displayed (where P is the group and Q is the subgroup) to give more insight into the proof, and this equation was credited to Galois. Note that Galois's version (equation (1)) did not specify a finite sum, however. This may have been the first time that the theorem was proved using the language of cosets.

Weber's Volume 2 (1896) covered more advanced material than volume 1; it began with abstract groups and proved Lagrange's Theorem again, this time for abstract groups, in section 2 of chapter 1. The proof was essentially the rectangular array approach (without explicitly giving the array). The theorem and its proof were then used as motivation for introducing the concept of coset ("as in the special case of permutation groups") and a coset decomposition equation, similar to equation (2) was displayed.

Theorem C did not appear in Weber's book and the only consideration of the number of values taken on by a function under permutations was the case of symmetric functions and the alternating function (discussed in volume 1). A proof of Fermat's Little Theorem appeared in volume 1 (proved without group theory); in chapter 2 of volume 2 the generalization of the Fermat Theorem for the integers modulo n was proved using group theory and Lagrange's Theorem.

Twentieth century developments

Increasingly in the twentieth century, coset terminology was used in the proof of Lagrange's Theorem. It is not so different from the rectangular array approach, since the rows of the array are in fact the cosets. But it is a different style, and while the rectangular array is usually credited to Cauchy, the coset approach seems to have been inspired by Galois's coset decomposition equation (1). We saw that these two historical threads were brought together in Weber's book.

According to Wussing [**31**], the first monograph devoted to abstract group theory was that of DeSeguier, appearing in 1904 [**26**]. He showed that the double cosets form a partition of a group. Then as a special case, he got the coset decomposition of a group G and then simply mentioned that $[G, A][A, 1] = [G, 1]$. Here $[G, A]$ denotes the index of a subgroup A in G so $[G, 1]$ is the order of G.

Although a number of authors credited the theorem to Lagrange, many did not mention Lagrange's name and it was some time before it became widely known as "Lagrange's Theorem." Van der Waerden's *Moderne Algebra* [**29**] was one of the most influential texts in algebra. It first appeared in 1930. The coset approach was used in

proving Lagrange's Theorem. Lagrange's name did not appear; however, in a footnote it was mentioned that the coset decomposition equation (1), which is frequently found in the literature, is due to Galois. In the "second revised edition" of 1937, the fact that any two cosets have the same number of elements was explicitly stated and proved, thus filling a gap in the first edition. In the English translation (of the second revised edition), appearing in 1949, the translator added a footnote stating that this theorem is also known as Lagrange's Theorem.

We might compare two other books which also appeared in 1937. Albert's *Modern Higher Algebra* [**3**] used the language of cosets to prove the theorem; there was no mention of Lagrange or other historical references. On the other hand, in Carmichael's *Introduction to the Theory of Groups of Finite Order* [**6**], the proof was given using the rectangular array. Carmichael called the theorem "First Fundamental Theorem" (his chapter 2 contained five fundamental theorems) but there was a footnote stating "This has sometimes been called the *theorem of Lagrange*".

In 1941, Birkhoff and MacLane's *A Survey of Modern Algebra* first appeared [**4**]. This book became a model for undergraduate modern algebra textbooks and helped to attach Lagrange's name firmly to the theorem. Section 9 of chapter VI is entitled "Lagrange's Theorem." It started with two lemmas showing that each coset has the same number of elements as the subgroups and any two distinct cosets are disjoint. This led up to "Theorem 18 (Lagrange): *The order of a finite group G is a multiple of the order of every one of its subgroups.*" There were various corollaries including Fermat's Little Theorem.

Some modern books apply the general theory of equivalence relations in connection with cosets and Lagrange's Theorem (for example, see Herstein's *Abstract Algebra* [**17**]). One defines a relation on the group G by letting aRb if $ab^{-1} \epsilon H$. It is proved that this is an equivalence relation and the right cosets are the equivalence classes. The right cosets thus form a partition of the group G. One of the earliest examples of this approach is found in a book of Hasse, published 1926 (see [**16**]).

Acknowledgment. We wish to thank the referees for many helpful suggestions.

REFERENCES

1. P. Abbati, Lettera di Pietro Abbati Modenese al socio Paolo Ruffini da questo presentata il dì 16. Decembre 1802, *Mem. Mat. Fis. Soc. Ital.* **10** (part 2) 1803, 385–409.
2. N. H. Abel, Mémoire sur les équations algébriques, ou l'on démontre l'impossibilité de la résolution de l'equation génerale du cinquième degré. Christiania 1824 (*Oeuvres Complétes de Niels Henrik Abel*, Christiania 1881, Johnson Reprint 1965, vol. 1, pp. 28–33).
3. A. Albert, *Modern Higher Algebra*, University of Chicago Press, Chicago, IL, 1937.
4. G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, Macmillan, New York, 1941.
5. H. Burkhardt, Die Anfänge der Gruppentheorie, Abh. Gesch. **6** (1892), 119–159.
6. R. D. Carmichael, *Introduction to the Theory of Groups of Finite Order*, Dover Publications, New York, 1956 (originally published 1937).
7. A.-L. Cauchy, Sur Le Nombre des Valeurs qu'une Fonction peut acquérir, lorsqu'on y permute de toutes les manières posibles les quantités qu'elle renferme, Journ. Ecole Polyt. **10**, (1815), 1–28. (Oeuvres Complètes, Gauthiers Villars, Paris, 1905, Series 2, vol. 1, pp. 64–90).
8. A.-L. Cauchy, Mémoire sur les arrangements que l'on peut former avec des lettres donneés, et sur les permutations ou substitutions à l'aide desquelles on passe d'un arrangement à un autre. In: Exercises d'analyse et de physique mathématique, 3, Paris, 1844, pp. 151–252. (Oeuvres Complètes, Gauthiers Villars, Paris, 1932, Series 2, vol. 13, pp. 172–282).
9. A.-L. Cauchy, Sur le nombre des valeurs ègales ou inègales que peut acquérir une fonction de n variables indèpendantes, quand on permute ces variables entre elles d'une manière quelconque, Comptes Rendus Paris 21, Oct. 6, 1845, pp. 779–797 (Oeuvres Complètes, Series 1, vol. 9, Gauthiers Villars, Paris, 1896, Extract 303, pp. 323–341).
10. A.-L. Cauchy, Mèmoire sur un nouveau calcul qui permet de simplifier et d'étendre la théorie des permutations, Comptes rendus, Paris 22, Jan. 12, 1846, pp. 53–63 (Oeuvres Complètes, Series 1, vol. 10, Gauthiers-Villars, Paris, 1897, pp. 35–46).

11. L. Euler, Theoremata circa residua ex divisione potestatum relicta, Nov. Com. Acad. Petrop. (1758–59), Petersburg, 1761. (English translation of portion in D. Struik, *Source Book of Mathematics, 1200–1800*, Harvard Univ. Press, Cambridge, MA, 1969, Chapter 1, section 8).
12. L. Euler, Nova Methodo Demonstrata Theoremata Arithmetica, Nov. Comm. Acad. Sci. Petrop. **8** (1760–61, pp. 74–104, publ. 1763), Opera (1), 3, 535–55.
13. E. Galois, Sur La Théorie des Nombres, Bull. Sc. math. Férussac 13, June 1830, (Oeuvres Math., Gauthiers-Villars, Paris, 1897, pp. 15–23).
14. E. Galois, Mémoire dur les conditions de résolubilité des équations par radicaux (Oeuvres math., Gauthiers Villars, Paris, 1897, pp. 33–61).
15. E. Galois, Lettre a Auguste Chevalier, Revue Encyclopedique, 1832 (Oeuvres math., Gauthiers-Villars, Paris, 1897, pp. 25–32).
16. H. Hasse, *Höhere Algebra*, vol 1, Berlin, W. de Gruyter and Co., (Sammlung Göschen 931), 1926.
17. I. N. Herstein, *Abstract Algebra*, Prentice Hall, 1995 (Third edition).
18. O. Hölder, Zurückführung einer beliebigen algebraischen Gleichung auf eine Kette von Gleichungen, Math. Ann. 34 (1889) 26–56.
19. C. Jordan, Mémoire sur le nombre des valeurs des fonctions. Journ. Ecole Polyt. **22** (1861), 113–194. (Oeuvres, Gauthier-Villars, Paris, 1961, vol. 1, item 1, pp. 1–82).
20. C. Jordan, *Traité des Substitutions et des équations algébriques*, Paris, 1870.
21. V. J. Katz, *A History of Mathematics, An Introduction*, Addison Wesley, 1998, Second Edition.
22. I. Kleiner, The Evolution of Group Theory: A Brief Survey, this MAGAZINE **59** (1986) 195–215.
23. J.-L. Lagrange, Réflexions sur la résolution algébrique des équations. Nouv. Mém. Acad. Berlin, pour les anneés 1770/71, Berlin 1772/1773. (See also Oeuvres de Lagrange, edited by J.-A. Serret, Paris, 1869.)
24. E. Netto, *Die Substitutiontheorie und ihre Anwendung auf die Algebra*, Leipzig, 1882.
25. P. Ruffini: *Teoria Generale delle Equazioni, in cui si è dimostrata impossibile la soluzione algebraica delle equazioni generali di grado superiore al quarto*. Bologna, 1799 (Opere Mathematiche, vol I, Ed. E. Bortolotti, Palermo 1915).
26. J.-A. de Séguier, *Théorie des Groupes Finis. Eléments de la Théorie des Groupes Abstraits*, Gauthier Villars, Paris, 1904.
27. J.-A. Serret, *Cours de Algèbre supérieure*, 3rd edition, Paris, 1866.
28. A. Vandermonde, Mémoire sur la résolution des équations, Hist. Acad. Sc. Paris, année 1771: Paris, 1774, pp. 365–416.
29. B. L. van der Waerden, *Moderne Algebra*, Springer, Berlin, First Edition 1930, Second Revised Edition 1937. (English translation of the second revised published by Frederick Ungar, New York, 1949, 1953.)
30. H. Weber, *Lehrbuch der Algebra*, F. Vieweg & Sohn, Braunschweig vol 1, 1895, vol. 2, 1896.
31. H. Wussing, *The Genesis of the Abstract Group Concept*, M.I.T. Press, Cambridge, MA, 1984. (Translation from the German by A. Shenitzer of *Die Genesis des Abstraken Gruppenbegriffes*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1969.)

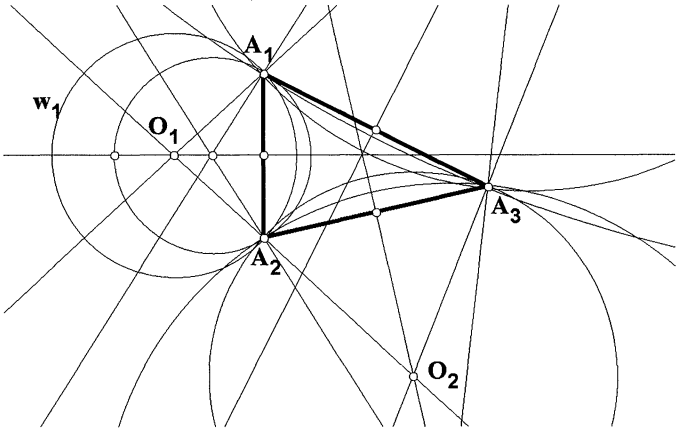

Figure 1 Problem 5 from USA Mathematical Olympiad (p. 167)