Available online at www.sciencedirect.com

SCIENCE DIRECT®

Historia Mathematica 31 (2004) 196–221

HISTORIA MATHEMATICA

www.elsevier.com/locate/hm

The mathematical life of Cauchy's group theorem

M. Meo

Benson Polytechnic High School, Portland, OR, 97232, USA

Abstract

Cauchy's theorem on the order of finite groups is a fixture of elementary course work in abstract algebra today: its proof is a straightforward exercise in the application of general mathematical tools. The initial proof by Cauchy, however, was unprecedented in its complex computations involving permutational group theory and contained an egregious error. A direct inspiration to Sylow's theorem, Cauchy's theorem was reworked by R. Dedekind, G.F. Frobenius, C. Jordan, and J.H. McKay in ever more natural, concise terms. Its most succinct form employs just the structure lacking in Cauchy's original proof—the wreath product.

Résumé

Aujourd'hui le théorème de Cauchy sur l'ordre des groupes finis est énoncé dans tous les manuels d'algèbre abstraite et sa démonstration se réduit à un simple exercice d'application d'outils mathématiques généraux. La démonstration originale donnée par Cauchy comportait, en revanche, des calculs complexes de groupes de permutations et contenait au fond une erreur. Ce théorème, source directe d'inspiration pour le théorème de Sylow, est énoncé en termes de plus en plus naturels par R. Dedekind, G.F. Frobenius, C. Jordan, et J.H. McKay. Sa forme plus concise emploie précisément l'outil manquant dans la démonstration originale de Cauchy—le produit en couronne.

MSC: 01-01; 01-02; 01A55; 01A70; 08-03; 20-03; 20B35

Keywords: Cauchy's theorem; Permutation groups; Sylow subgroups; History of finite group theory

Celui qui s'engage le premier dans une région inconnue mérite quelque indulgence pour les fausses routes qu'il y peut faire.
—Jean Montucla, *Histoire des Mathématiques* (1763)

E-mail address: mmeo@pps.k12.or.us.

0315-0860/$ – see front matter
doi:10.1016/S0315-0860(03)00003-X

1. Introduction

Almost all university departments of mathematics in the U.S. require abstract algebra at an entry level, and motivational sections in textbooks frequently explain how the theorems arose—Lagrange's theorem in the search for the solution to the fifth-degree algebraic equation, for example, or Cayley's theorem in the effort to give an abstract definition of the concept of groups. On the other hand, Cauchy's group theorem, that to every prime number p that divides the order of a finite group there corresponds a subgroup of order p, even though of historic importance, receives a treatment completely divorced from its original context. This paper attempts to correct and to explain that silence.

Cauchy's original proof of his theorem contained a significant logical gap. At least one knowledgeable contemporary complained of the proof's obscurity. Ludvig Sylow's extension of it, and Georg Ferdinand Frobenius' more abstract proof, both avoided the method of Cauchy's original proof, a method we now call the wreath product. In the 20th century, the wreath product was correctly characterized and served as the basis for the most recent proofs of Cauchy's theorem.

It is a well-worn commonplace of the history of science that the initial statement of a scientific finding is often partial and confused, leaving to later investigators the opportunity to clarify, generalize, and simplify. One *bon mot* has it that "a mathematician's reputation rests on the number of bad proofs he has given" [Littlewood, 1953, 41], but the problem is not limited to mathematics. Sir Isaac Newton's original demonstration that the orbital motions of the planets of the solar system are governed by a force which varies as the inverse square of distance is notorious for its obscurity, continuing to evoke controversy among scholars to our own day, while introductory textbooks dispose of it in a couple of pages.[1] Rafael Bombelli introduced the algebra of the complex numbers in the 16th century, but the subject remained obscure until the early 19th century, with Jean Robert Argand's and Carl Friedrich Gauss' independent graphical geometrical representation of complex numbers—and Cauchy's own fertile use of it in mathematical analysis during the 1820s. Cauchy's theorem in permutation groups, which constituted the major conclusion of the 101 pages[2] of "Mémoire sur les arrangements que l'on peut former avec des lettres données" [Cauchy, 1845], occupies a grand total of ten lines in McKay [1959]. *Nil mirari*, we might say. There is nothing to be wondered at here.

On the other hand, the work of which this theorem was the main result was a milestone in the development of abstract algebra. You might call it the first nontrivial result in permutation groups. Cajori [1919, 352] flatly states "Cauchy has been given the credit of being the founder of groups of finite order;" according to Kiernan [1971, 97], "Here finally was a man of stature in mathematics who thought it worthwhile to publish extensively on the question of permutations." It appeared just before the posthumous publication of Galois [1846] in Joseph Liouville's *Journal de mathématiques pures et appliquées*, and the two publications together have recently been characterized as "the two sources that introduced group theory to mathematics" [Neumann, 1989, 293].

Several recent works have disagreed on the stimulus for Cauchy's publication. The 56-year-old baron already had hundreds of contributions, in dozens of different mathematical areas. He dominated the mathematical section of the Paris Academy of Sciences, but had not discussed permutations for 30 years.

[1] The contrast between the original statement and modern proofs (although not the scholarly controversy) appears in succinct form in Pourciau [1997].

[2] That was the length of the original article in a learned journal; in the 1932 edition of Cauchy's collected works (cited here) the article occupies 111 pages. Translations in this paper not credited to other scholars are my own.

Kiernan [1971, 97] speculated that it was in order to forestall the soon-to-be-published results of Galois that Cauchy brought out his prolix, inelegant, although beautiful, results.[3]

Against that suggestion, Dahan Dalmedico [1980] attempted by means of a painstaking study of mathematical styles to distinguish clearly between the two approaches to group theory of Galois and Cauchy. She found the former to display "a suppleness and a power of articulation and analysis which remained absolutely foreign to the point of view of Cauchy" [Dahan Dalmedico, 1980, 296]. Still, she found undeniable mathematical connections. Not only were the problems Cauchy handled at several points very close to those of Galois—Galois had specifically stated Cauchy's theorem without proof in an unpublished manuscript. It also happened that Cauchy once where his method resembled Galois' gave credit instead to Charles Hermite, who was at about that time attending Liouville's lectures about Galois' unpublished manuscripts [Cauchy, 1845/1846, IX, 459]. Citing [Dahan Dalmedico, 1980] as her main source, Toti Rigatelli [1989, 46] inquired rhetorically whether Cauchy had asked Liouville to delay publication of Galois' work until he produced his own.

Such hypotheses seem to slight the fact that Cauchy gave an explicit reason for presenting his results: the 23-year-old J.L.F. Bertrand had submitted to the Academy a proof of a conjecture appearing in Cauchy [1815], that the (what we would today call) order of a (nontrivial) permutation group on n elements is at least n. Appointed to evaluate the validity and originality of the work for the Academy, Cauchy submitted 25 separate papers of his own, between 15 September 1845 and 11 April 1846, almost 300 pages in all, recording "some of the most notable propositions to which [he had] arrived," in extending his 1815 result and that of Bertrand.

It was under still highly contested circumstances, then, that Augustin-Louis Cauchy, the dominant French mathematician of his generation, produced his *Mémoire*. Let us begin with an examination of Cauchy's own reasoning, and then follow how the proof has changed with time. Cauchy's theorem was generalized, then made more abstract, well before it was simplified in its present form. For a long time it was submerged. It has had an interesting career.

2. Cauchy's proof

Cauchy [1845] begins with a presentation of the notation that had first appeared in Cauchy [1815] to denote permutations, notation that has since become standard. Cauchy distinguished between the arrangement of a set of letters and the permutations—or, synonymously, "substitutions"—which operate on them. If the three letters

$$x, y, z$$

[3] He has been echoed in this by other authorities, a recent example being Scholz [1990, 387]. Cauchy practiced a pious Catholicism and held strictly legitimist monarchial views; if he did take the idea for his theorem from Galois without attribution he would have had a political motive for it. (Perhaps a religious motive, as well; Grattan-Guiness [2000, 482–483] presents an argument for Cauchy's mathematics "emulating" his Christianity in its rigor and claim to absolute truth.) The Revolution of 1830 that established a cadet branch of the Bourbons sent Cauchy into eight years of self-imposed exile; but it was denounced by Galois because it did not abolish the monarchy altogether. The talented youth's protests led to a prison sentence. Nor was contemporary practice of mathematics without a political impact; see Mazzotti [1998] for the situation in the Kingdom of Naples at about this time.

are rearranged as

$$y, z, x,$$

the permutation is the operation

$$\begin{pmatrix} y & z & x \\ x & y & z \end{pmatrix},$$

where the arrangement on the bottom line is replaced with the arrangement on the top line. This particular permutation, since it shifts the letters while retaining the same order, is named by Cauchy *cyclic* and given the alternate representation

$$(x, y, z).$$

Permutations of a set of letters have a given order; that is, a given number of repetitions will generate identity. For example, by repeating our sample permutation three times the letters of the initial arrangement will have returned to their starting positions. This order is written aptly as an exponent. "Nothing prevents," Cauchy [1845, 185] comments, "the representation of substitutions by simple letters

$$P, Q, R, \ldots,"$$

a representation which results in

$$(x, y, z)^3 = P^3 = 1$$

denoting the situation succinctly and unambiguously. The permutation P is third-order.

Well aware of the great advantages of his clearer notation, Cauchy elsewhere compared it to the differential notation of calculus introduced by Leibniz [Cauchy, 1845/1846, 10: 35–36]. The focus of study now became permutations rather than the arrangements of letters. The permutation which changes nothing is a natural identity element, and that permutation which returns P to the identity element is its inverse, P^{-1}. All these results from Cauchy's 1815 paper on permutations were reiterated without reference to the earlier work.[4]

Cauchy's theorem concerns the permutations or substitutions (Cauchy employs both terms, but prefers the latter) on n letters, the composition of which generate a closed system by sequential performance. Writing down a composition repeats the order of their operation. That is, the permutation Q followed by the permutation P—written as PQ—is again a permutation. Applying one permutation after another is for Cauchy a "conjugation," a joining, so he is dealing with a "system of conjugate substitutions" where we speak of a group.[5]

For every prime number p that divides the order of a system of conjugate substitutions, then, Cauchy's theorem states that there exists among the elements of that system at least one regular substitution of order p. Cauchy does not use the notion of a subgroup. In contrast to notation, however, since Cauchy's day the meaning of the crucial term "regular permutation" has changed: for him a regular permutation is

[4] With respect to Galois, in contrast, both Kiernan [1971, 82–83] and van der Waerden [1985, 107] agree with Dahan Dalmedico [1980, 286] that his inconsistent terminology made his revolutionary papers difficult for his contemporaries to understand.

[5] Galois had used the word "group" to describe just this situation in a paper written in 1830 and being prepared for publication by Liouville in the period during which [Cauchy, 1845] was in preparation.

cyclic, either of a single component or a product of several disjoint components, all of the same order. One of the first theorems proved in the 1845 *Mémoire* is that it is possible to write any permutation as a product of disjoint cycles; once that is done, all permutations are taken in that form.

If a permutation is of prime order, it is "primitive;" one of the first consequences of these definitions is that all primitive permutations are regular, but not conversely. Cauchy calls two permutations on a given set of letters "similar" if they have the same number of cycles, with the same number of letters in corresponding cycles, so that each pair of corresponding cycles have the same order. He then proves the relation that, if P and Q are related by

$$RP = QR$$

by means of a third permutation R, then they are similar. In contrast, we currently define similarity by means of the equation and then use it to prove the correspondence of disjoint cycles.

Cauchy included a variety of results about groups of permutations—conjugate systems of substitutions—in his great 1845 memoir, but he begins the section devoted to the proof of his pioneering theorem with a statement of Lagrange's theorem, that (in today's terms) every subgroup divides the order of the group. On the very first page of his Memoir Cauchy had alluded to the fact that, in arranging a set of n different letters, one would be able to make a choice among n different letters for the first position, among $n-1$ for the second, and so on, with the result that

$$(n)(n-1)(n-2)\cdots(1) = n!$$

would be the total number of possible arrangements of n letters—the order of our present-day "symmetric group." The permutations composing this conjugate system of substitutions are the *operations* on an arbitrary initial arrangement which will produce each of the $n!$ final arrangements.

Without making any direct reference to Lagrange, Cauchy [1845, 207] makes use of a tabular display (for the first time in the paper) in his graphic and intuitive proof.[6] Given a conjugate system of substitutions, an arbitrary subgroup of the group of n letters, if its order M equals $n!$, we are done. If it is less than $n!$, then there exist some permutations of n letters that are not included in the system of order M. We can write the members of the system of order M one after another on one line,

$$1, P, Q, R, \ldots,$$

and form the multiples of each of them with the permutations that are not included, say $U, V, W, \ldots$. The resulting two-dimensional array contains only distinct permutations

$$\begin{array}{ccccc} 1 & P & Q & R & \ldots \\ U & UP & UQ & UR & \ldots \\ V & VP & VQ & VR & \ldots \\ W & WP & WQ & WR & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{array}$$

and therefore $n!$ is a multiple of M.

[6] The first use of a complete rectangular-array proof of Lagrange's theorem has been attributed to Pietro Abbati in 1802, in discussion of the work of Ruffini. In his 1815 paper Cauchy discussed Ruffini's work as well, employing an array similar to Abbati's. See Roth [2001, 103].

Characteristically, Cauchy is careful to demonstrate that each of the permutations in the table must indeed be distinct from any other; that the proof would also work if the order of multiplying permutations were reversed; and that the order of any single permutation will divide the order of the group. His next step was to construct a special system of conjugate permutations, one of prime-power order; that is, in today's language, to construct a Sylow subgroup of the symmetric group.

The claim is, that there exists a closed subgroup of this symmetric group, which has the order of the power of a prime number. If

> ... p is a prime number, equal to or less than n, i the greatest multiple of p contained in n, and p^f the greatest power of p which divides integrally the product $n! = 1 \cdot 2 \cdot 3 \cdots n$
>
> [then] with the number i of arbitrarily chosen variables one can always construct a system of conjugate substitutions of the order p^f.
>
> [Cauchy, 1845, 221]

We begin the construction of the special system by considering a regular permutation P of the n letters $x, y, z, \ldots$, which means that P contains a set of cyclical elements of the same order, let us say, of order a. If there are b such factors, then

$$n = ab.$$

Next, let us put the n letters into a two-dimensional array where each power of the regular permutation P occupies a given point on the horizontal. The result of the operation of each power of the permutation P is a new arrangement of the a letters, with each arrangement beginning with a different one of the a letters, and the a letters are placed one after another along the line. If instead of permutations within each cyclic factor we permute one letter in a given set with a letter in another cyclic factor, we generate a different permutation, which we can call Q. The powers of the permutation Q replace just one of the a letters with one of the letters which are not among the original set permuted by P; for each of the a letters Q can replace it with b others, distinct from the replacements generated by powers of P. Let us arrange the powers of Q in the vertical direction. Then an array of all possible derived permutations of P and Q would look like

$$\begin{array}{ccccc} 1 & P & P^2 & \ldots & P^{a-1} \\ Q & QP & QP^2 & \ldots & QP^{a-1} \\ Q^2 & Q^2P & Q^2P^2 & \ldots & Q^2P^{a-1} \\ & & & \vdots & \\ Q^{b-1} & Q^{b-1}P & Q^{b-1}P^2 & \ldots & Q^{b-1}P^{a-1} \end{array}$$

and constitute a system of conjugate substitutions of the order ab. The two permutations P and Q commute; that is,

$$PQ = QP,$$

since the first of these moves a given letter one space over and one space down, while the second moves the same letter one space down and one space over, which is the same operation.

At this point, as quite frequently in the *Mémoire*, Cauchy pauses for an example, "pour fixer les idées," as he says. With six variables we have the array

$$\begin{array}{lll} x, & y, & z \\ u, & v, & w \end{array}$$

resulting in the regular permutations

$$P = (x, y, z)(u, v, w), \qquad Q = (x, u)(y, v)(z, w)$$

where, clearly $P^3 = 1 = Q^2$. The array of all possible derived permutations is then

$$\begin{array}{lll} 1 & P & P^2 \\ Q & QP & QP^2 \end{array}$$

which is a system of order 6.

Moreover, for any number $n = ab$, where a and b are primes, each of the derived permutations $Q^h P^k$, where the powers h and k are integral numbers less than the orders b (of Q) and a (of P), is distinct from any other derived permutation.

The next step is to generalize the procedure. If the number n is the product of l prime factors

$$n = a \cdot b \cdot c \cdots,$$

recourse to the above construction will generate permutations P, composed of n/a cyclic units each of order a; R, of n/b cyclic factors each of length b; $S \ldots$ so that the system of all possible derived permutations, in the case $a = b = c = \cdots$, has the order

$$n = a^l.$$

Nor need the permutations cover all of the n letters with which we begin. If the number ha is a multiple of a less than or equal to n, we can form with the ha letters the h disjoint cyclic permutations

$$P_1, P_2, \ldots, P_h,$$

each of order a. The system of conjugate substitutions formed of these and their derived permutations will be of order a^h.

If the number h is equal to or larger than the product kb, Cauchy points out, then from the h permutations P_i of the "first kind" it will be possible to construct k permutations of the "second kind"

$$Q_1, Q_2, \ldots, Q_k,$$

each of which will consist of permutations among the cyclic factors of the P_i. That is, the permutations P may be arranged in h blocks all on a horizontal, producing ah different arrangements of the n letters, while the permutations Q_j each replace a letter in one of the distinct groups of a letters by one of the letters in another set distinct from that P_i. Such permutations Q_j will have k sets of cyclic components, each of order b. They will commute with each other and with the P_i, if they behave as the permutations on the first array do, so that the system of conjugate substitutions will be of the order

$$a^h b^k.$$

From this point it should be clear that if we suppose p to be a prime number less than n, and I the multiple of p contained in n—that is, less than or equal to n—Cauchy is able, using I variables, to construct a system of the order

$$p^f,$$

where p^f is the largest number which divides $n!$. He is careful to clarify each of the details, but we can already see the course of the proof.

"Pour fixer les idées," let us use his first example [Cauchy, 1845, 222] of this theorem in action. With a set of five letters, we use $p = 2$. The multiple of p contained in n is $2p = i = 4$, so we will be using only four of the five letters. These may be written

$$\begin{array}{ll} x, & y, \\ z, & u, \end{array}$$

which suggests the cyclic permutations

$$P_1 = (x, y), \qquad P_2 = (z, u)$$

of second order for the first kind. There are two of them (i. e., $h = 2$), and their letters will be exchanged by a permutation of the second kind,

$$Q = (x, z)(y, u),$$

of which there is only one, so $k = 1$. The system of all possible derived permutations is

$$\begin{array}{llll} 1 & P_1 & P_2 & P_1P_2 \\ Q & QP_1 & QP_2 & QP_1P_2 \end{array}$$

of order eight. But $2^{h+k} = 2^3 = 8$ is the largest power of 2 which divides $5! = 120 = (8)(5)(3)$.

So Cauchy argues. (Note that he changed the meaning of the permutation P in generalizing the initial, correct result.) Suppose we ask whether the Ps and Q above do in fact permute. We write

$$QP_1 = (x, z)(y, u)(x, y)$$

and read off the result of the composition of permutations by starting on the right-hand side, with the element y, and moving to the left: y becomes x, and then x becomes z. Thus far we have

$$(y, z;$$

the repetition of this procedure shows that the z becomes x:

$$(y, z, x;$$

eventually we have $QP_1 = (y, z, x, u)$. On the other hand,

$$P_1Q = (x, y)(x, z)(y, u) = (u, x, z, y),$$

that is, *not* QP_1.

Thus there is a serious logical gap at a crucial point in Cauchy's proof.[7] When looking for the reason for the failure, we see that the confusion arose from an inappropriate direct product.

Cauchy's argument relies upon the assertion that (in modern terms) the direct product of two subgroups of a group is a subgroup itself, but this assertion is only true if both are normal subgroups of the group

[7] The gap is passed over in silence in Dahan Dalmedico [1980, 301–303] and pointedly ignored by Waterhouse [1980, 282 note]: "The original paper has a minor slip: Cauchy was overhasty in treating the wreath product, and wrote that all elements in the group had order p. The error is put right in the version published in the *Exercises = Oeuvres*, (2) 13: 171–282 (see pp. 221–235)." This cites the very pages (221–225) containing the specific example of noncommuting Ps and Q reproduced above. In fact, two of three specific examples given by Cauchy here do not in fact commute, contrary to Cauchy's claim.

containing them both. It is not true in the generalization he presents here because the group generated by Q is not necessarily normal in the symmetric group S_n. That is, if all elements p_i of the subgroup P and all elements q_j of the subgroup Q of the containing group G enjoy the property that

$$gp_1g^{-1} = p_2$$

for all elements g of G, then the product elements pq do generate a subgroup of G, whose order is the product of the orders of P and Q.

When the question becomes, as in this case, one of finding the product of two subgroups, only one of which is known to be normal in the containing group, then the resulting subgroup, which still is of the order of the product of the orders of the two subgroups, is constructed by means of a generalization of the direct product, namely the semidirect product.

Suppose A and B two subgroups of a group G, where only B is normal in G. We take each element of the product of the two groups to be of the form a_ib_j, so that the first group element comes from the normal subgroup (since we are discussing permutations of arrangements of letters, the permutation of a normal subgroup of the symmetric group is applied first) and is then operated on by an element of the other, not necessarily normal subgroup.

We proceed to multiply two elements of the product subgroup, call them a_1b_1 and a_2b_2; we want the result to look like a_3b_3,

$$(a_1b_1)\cdot(a_2b_2) = a_1b_1a_2b_2 = a_1\left(a_2a_2^{-1}\right)b_1a_2b_2 = a_1a_2\left(a_2^{-1}b_1a_2\right)b_2,$$

where, since B is normal in G, $a_2^{-1}b_1a_2$ is an element in B, so

$$= a_3b_3.$$

In taking the semidirect product, then, the element from the normal group is conjugated—$(a_2^{-1}b_1a_2)$—with an element from the nonnormal group before it is multiplied. It is now no accident that the Q always stays on the left in the table above.

Having accomplished, with deceptive clarity, the construction of a subgroup of prime-power order, Cauchy proceeds to examine the characteristics of a system of conjugate substitutions of prime-power order. Such groups are composed of powers of cyclic permutations, and he proves that in a cyclic group of order divisible by the prime p there always exists at least one element of order p [Cauchy, 1845, 234].

When we turn to the construction of a subgroup of order p within a not necessarily commutative group, the proof is not at all so straightforward. Up to this point it seems reasonable to characterize Cauchy's presentation as extending Lagrange's theorem, even if there are claims of commutative permutations that are unjustified. The remainder of the proof employs additional structure. It begins with the lemma

> Let two systems of conjugate substitutions be formed from the n variables
>
> $$x, y, z, \ldots,$$
>
> so that the first system,
>
> $$1, P_1, P_2, \ldots, P_{a-1},$$
>
> is of order a, and the second,
>
> $$1, Q_1, Q_2, \ldots, Q_{b-1},$$

is of order b. Let I be the number of permutations R which satisfy a symbolic equation of the form

$$RP_h = Q_k R,$$

where h, k are any given pair of integers, such that

$$a - 1 \geqslant h \geqslant 1 \quad \text{and} \quad b - 1 \geqslant k \geqslant 1.$$

Then the number I, divided by ab, will have the same remainder as the number $n!$ divided by ab, so that

$$I \equiv n! \pmod{ab}.$$

[Cauchy, 1845, 274]

The proof of this is not difficult, but seems a long way from an assertion of the existence of a subgroup of prime order dividing the order of the group.

At the start of the proof, Cauchy asks us to look at those permutations which can*not* satisfy any symbolic equation of similarity between permutations. Since there are only $n!$ permutations in the entire symmetric group as a whole, there are

$$J = n! - I$$

of these. The theorem is proved if J is a multiple of ab.

Suppose U to be a permutation of which this is true. Any combination

$$Q_k U P_h$$

will be distinct from all other forms. If this were not the case, if there existed any k, h, k', and h' such that

$$Q_k U P_h = Q_{k'} U P_{h'},$$

then multiplying both sides by the respective inverses would give

$$U P_h P_{h'}^{-1} = Q_k^{-1} Q_{k'} U.$$

The $P_h P_{h'}^{-1}$ will be one of the permutations contained within the system P_i; and $Q_k^{-1} Q_{k'}$ within the system Q_j; but this contradicts the assumption that U could *not* satisfy any

$$U P_i = Q_j U \quad \text{(unless both } P \text{ and } Q \text{ are the identity)}.$$

Therefore there are no such k, h, k', h'. Since all of the $Q_k U P_h$ are distinct, there must be ab such combinations or derived permutations. Setting V as a second permutation enjoying the same property, and using the same argument to show that all $Q_{k'} V P_{h'}$ must be distinct from all $Q_k U P_h$, as well as the fact that the total of derived permutations for V must be the same as for U, Cauchy concludes that J, the total number of permutations which cannot satisfy the symbolic equation of similarity, must be a multiple of ab.

This crucial step in Cauchy's proof of his general theorem is the first appearance[8] of double cosets in mathematics. Just as we would today call the set of elements of a group $U P_i$, with i varying throughout the closed subgroup P, a coset, so the structure under discussion here, $P_i U Q_j$, is a double coset. The

[8] See Speiser [1937, 64], Waterhouse [1980, 282].

lemma states that these double cosets partition a certain set of the elements of the symmetric group, that is, divide the set into distinct subsets. We have a name for the set which is partitioned, too: we say that subgroups P and Q are conjugate if there exist i, j such that

$$U P_i U^{-1} = Q_j.$$

Were the condition on Cauchy's system of conjugate substitutions a little stronger, that is, if there existed i, j such that

$$U P_i U^{-1} = P_j$$

for all elements U in the symmetric group, P would be a normal subgroup, a concept approached but not grasped by Cauchy; it is central to, although not clearly articulated in, Galois [1846]. If Cauchy is suspected of making use of Galois, he has to be credited with following his own agenda.

From this lemma Cauchy can draw the logical corollary that if no such permutation R exists which can satisfy a symbolic equation of similarity, that is, if $I = 0$, then $J = n! - I = n!$ and $n!$ must now be divisible by ab.

The contrapositive of this first corollary is the second, that is, that if $n!$ cannot be divided by the product ab of the orders of two systems of conjugate substitutions (in our terms, subgroups),

$$P_1, P_2, \ldots, P_{a-1} \quad \text{and} \quad Q_1, Q_2, \ldots, Q_{b-1},$$

formed from subsets of n distinct letters, then one or more of the substitutions P_h is similar to one or more of the substitutions Q_k.

We are now ready for the proof of the statement that there exists a subgroup of order p for every prime which divides the order of the group. Now the "first step" with which we began asserted that it was always possible to form a system of conjugate substitutions—a subgroup of the symmetric group—of the order p^f, the highest power of the prime number p which divides $n!$, the order of the group. Let the system

$$1, Q_1, Q_2, \ldots, Q_{b-1}$$

be that special system where $b = p^f$. Let the order of the system

$$1, P_1, P_2, \ldots, P_{a-1}$$

be a multiple of the prime number p. (It is important to note here that the Q subgroup is commutative—see above, pp. 6–7—while the P subgroup has a much less restrictive definition. Its elements do not necessarily commute.) In such a case, then, the product ab of the orders of the two systems cannot divide $n!$, and at least one of the Qs will be similar to—that is will be conjugate to, in modern terms—at least one of the Ps. That two similar substitutions are the same order is so easy an inference that Cauchy alluded to it in his definition of similarity.

Further, says Cauchy at this point, the system of Qs is a cyclic group, since it is of the order of a power of a prime number. Although the whole group is of the order p^f, a previous theorem [Cauchy, 1845, 234–235] had shown that at least one permutation in any cyclic group of the order of a multiple (let alone a power) of a prime number will be of order p. Therefore, he concludes, at least one of the permutations $P_1, P_2, \ldots, P_{a-1}$ will be a cyclic permutation of order p.

It is in Cauchy's final step that Dahan Dalmedico [1980, 303] detects a logical error. Although there is one Q_i that is similar to a certain P_j, she points out, Cauchy did not establish that particular Q_i was of order p. Still, she adds, the error was not fatal:

> But this error carries with it no consequence, since this Q_i will be [of prime power order], of the order p^i, and a certain power of this Q_i will be of order $p \ldots$

and so will the same power of the P_j which was similar to Q_i, be of order p. It is this *power of* P_j, an element of order p, which can generate, in a cyclic fashion, a subgroup of order p.

Dahan Dalmedico [1980, 303] draws attention to this seemingly minor lapse:

> The indicated lack of precision may be due not only to too rapid publication, but may also testify to a certain vacillation in the use made of global commutativity. This proof is revealing of Cauchy's thought: in order to obtain a result regarding the system of conjugate substitutions of order M, he brings back the familiar symmetric group S_n, for which he now has quite a few results and appropriate techniques. It [the proof—MM] remained in use for rather a long time, being incorporated whole into Jordan's "Traité des Substitutions"

and later Dahan Dalmedico [1980, 316] concludes that Cauchy's role in the birth of group theory has been "misunderstood:" [Cauchy, 1845], stuck in the consideration of the special case of the symmetric group, never examining the nature of subgroups or of the normality of a subgroup to the group containing it, was only influential in subsequent mathematics by the use to which it was put in Jordan [1870]. Citing Dahan Dalmedico, Belhoste [1985, 207] judges that "Cauchy remained a prisoner of his calculation techniques."

3. Richard Dedekind

Thanks to a happy confluence of a high regard for the life of the mind and a metropolitan educational establishment of a number of national engineering schools, Paris in the first half of the 19th century served as the "mathematical capital of the world." By the late 1840s, however, the best mathematical research had become more widely distributed.

A student of Gauss while an undergraduate (1850–1852) at Göttingen, Richard Dedekind was pressed to broaden his mathematical skills after taking a degree. He worked closely with Dirichlet when the latter arrived in 1855, and this may have put Dedekind in touch with Liouville's *Journal*. In any case, during the winter 1857–1858 he gave a course on the Galois theory of the solubility of equations by means of finite groups.[9]

The manuscript pages of Dedekind's writings of this period, published 16 years after his death in his *Gesammelte mathematische Werke* [Dedekind, 1932, 3: 439–445], attracted the attention of both Nicholson [1993], for its pioneering enunciation of the concept of quotient group, and Waterhouse [1980], for its statement of an original, abstract, proof of Cauchy's theorem.

Dedekind marked his statement: Given a group G of order g, in which the prime number p divides g, there exists at least one element of order p in G, with a question mark.[10] He remarks that were the theorem false, it would have to be so only for $g \geqslant 6$, so he proceeds by induction, making the assumption that the theorem is true for all groups K of order $k < g$. He observed next that the order of G could be

[9] This course was published in Scharlau [1981, 59–100], with explanatory notes by Scharlau (pp. 101–108). I thank the anonymous referee for bringing this reference to my attention.

[10] Waterhouse [1980, 288] considers this and the headline "Attempt at a proof" persuasive that Dedekind was not previously acquainted with Cauchy's theorem. This would be unusual, given the immediate relevance of [Cauchy, 1845] to group theory, Cauchy's personal eminence, and Dedekind's effort to become familiar with the most recent French work.

written as

$$g = p^{\omega} n,$$

n not divisible by p, and where n cannot be set equal to one, since clearly (*einleuchtet*) a group of prime power order is going to have a subgroup of order of that prime. Recall, in comparison, that this is just the sort of thing Cauchy took the trouble to prove in detail.

This subgroup K must be of order prime to p, given the induction supposition. We next look at the set H of all elements ϕ of G such that

$$\phi^{-1} K \phi = K,$$

noting that H includes all of G if K is a "proper" subgroup, a designation which echoes Galois' term "décomposition propre," but is today a "normal" subgroup. This H is itself a group, and its order divides the order of G.[11] Its degree must also be a multiple of K, since we defined K as a normal subgroup of H. Dedekind writes

$$g = \mu h$$
$$h = \nu k$$

to denote the multiples of the orders of the groups K and H. Either there exists a normal subgroup of G (and we will choose K to be that one), or G contains no nontrivial normal subgroups.

Supposing K normal to G, $g = h = \nu k$. By the use of the fundamental concept of quotient group developed in the previous couple of pages, K partitions G into ν distinct sets,

$$G = K_0 + K_1 + \cdots + K_{\nu-1},$$

each of which has the property of combining (that is, $K_\alpha K_\beta = K_\gamma$), and so forms a group of order ν. Since $\nu = g/k = p^{\omega} \cdot n/k$, where k is prime to p, the number ν must be divisible by a power of p. We then have a group of order less than g whose order is divisible by p, and which by the induction hypothesis must have at least one element of order p. Call it K_1, which by definition is equal to ϕK, where ϕ is some element of G. K_1 raised to the pth power recovers the identity element, which in this case is K,

$$(\phi_K)^p = K,$$

whence ϕ^p is an element of K, and the order to which ϕ itself must be raised to recover the identity of G must consequently be a multiple of p. By the induction hypothesis, the cyclic subgroup generated by ϕ contains at least one element of order p.

Alternatively, we may suppose that G contains no nontrivial normal subgroup. Dedekind at this point uses the fact that the order of every element of a group divides the order of the group. K can be a cyclic subgroup, of order d, let us say, containing $\phi(d)$ generators. All of the d's in the group G must divide n, and the product of the number of cyclic groups of order d times the number of generators of order d, must equal the number of distinct elements in the group. Call $\psi(d)$ the number of cyclic subgroups of

[11] Dedekind proved neither of these trivial statements, illustrating his awareness of the theorem of Lagrange. The set H, today in general use in abstract group theory, is the "normalizer" of a group.

order d. Then, summing over all possible values of d,

$$g = \sum \psi(d)\phi(d).$$

The subgroup H formed from K is a proper subgroup of G,[12] and its order h must be prime to p. Since $g = \mu h$, Dedekind concludes that μ is always divisible by p^{ω}, no matter what the size of K. Everything done to this point is true, but then Dedekind's notation fails him. He implicitly identified the index μ of the normalizer H with the number of generators of cyclic groups of order d, $\psi(d)$. The unit element being unique, and the only cyclic group of order one, he then rewrites the above, summing over all $d' > 1$,

$$g = 1 + \sum \psi(d')\phi(d') \qquad (*)$$

or, substituting the values above,

$$p^{\omega}n = 1 + p^{\omega}m,$$

which is impossible. Thus there must be at least one nontrivial normal subgroup of G, and the proof is done.

Waterhouse [1980, 288] described this "strange but ingenious" proof as "of course never polished for publication," but it is fairer to point out that there is no connection between the true statement of the divisibility of μ by p^{ω} and its application to the right-hand side of $(*)$. The set of all elements of G which conjugate K do indeed form a partition of G, but not necessarily into so many cyclic groups. We call this partition of G by means of the normalizer H of K the "class equation," using the term introduced by Frobenius [1887a]. The similarity of this classic presentation (discussed below in the section on George Miller) to that of Dedekind is striking. On 8 February 1895 Dedekind himself wrote to Frobenius

> I have been quite taken (*sehr gespannt*) with your work on groups, especially with the simplicity of your methods, among others your proof that every group whose degree is divisible by the prime p contains an element of order p. In the first years of my own study of groups (1855–1858) I arrived at it by a much more complicated (*umständlicheren*) path [Dedekind, 1932, 2, 419].

Notice that even here Dedekind does not refer to what Frobenius had called Cauchy's theorem by that name. As shown above, Dedekind's attempt at a proof is actually far from complicated; rather, it lacks the final step contained in Frobenius' presentation.

Dedekind does not mention it, but (as pointed out by Scharlau [1988, 44]) he proceeded directly to derive one of Sylow's theorems in the same manuscript. He speaks throughout of groups in completely abstract terms, and has no need for a subgroup of the symmetric group in order to realize what a group is.

4. Ludvig Sylow

The diffusion of mathematical research in group theory from Paris into Scandinavia also intimately involved Cauchy's theorem, to some extent because so few mathematicians were doing research in this

[12] If H were not a proper subgroup of G, then K would be a normal subgroup, which was excluded.

area, and they were virtually all known to one another.[13] Of course, the logical connection was also strong. In a marginal note not published in Galois [1846], Galois himself had written

> *Theorem*. If the number of permutations of a group is divisible by p^n (p being prime), the group will have, for a divisor, a group of p^n permutations (cited in Dieudonné [1978, 1: 117])

—the first of the three generalizations of Cauchy's theorem proved in 1872 by the remarkable Norwegian high-school teacher, Ludvig Sylow [Sylow, 1872], and now named after him.

Acquaintance with Cauchy's theorem, and more generally with the Paris group of mathematicians, stimulated Sylow to his discoveries. Sylow started his mathematical career with the study of the theory of equations in the unpublished works of his deceased countryman, Niels Henrik Abel, a study which led Sylow to the closely related work of Galois. It is of note that Galois himself referred to Abel[14] in a passage in the unpublished "Preface" (to Galois [1846]) which has been interpreted by many as condemning Cauchy:

> I must tell you [the reader] how manuscripts go astray in the portfolios of the members of the Institute, although I cannot in truth conceive of such carelessness on the part of those who already have the death of Abel on their consciences. I do not want to compare myself with that illustrious mathematician but, suffice to say, I sent my memoir on the theory of equations to the Academy in February of 1830 (in a less complete form in 1829) and it has been impossible to find them or get them back.

Sylow's selection of Galois theory as a discipline within which to concentrate his efforts was confirmed by his 1861–1862 trip to Paris and Berlin. Cauchy had died four years before, but Sylow reported to the ministry which funded his trip that he had attended lectures by Liouville and had "made [himself] acquainted with newer works, particularly in the theory of equations" [Birkeland, 1996, 185]. Upon his return to Norway Sylow gave a lecture course at the University of Oslo on the central parts of Abel's and Galois' theory of equations.

The manuscript lecture notes, in Sylow's own hand, contain the inquiry, right after the statement of Cauchy's theorem, "What if g is divisible by p^n? Can the above be extended?" [Birkeland, 1996, 191] (we use letters for the variables corresponding to ones used above). While Sylow examined groups which were of mathematical significance solely as Galois groups of a certain equation, as Scharlau [1988] emphasizes, still Cauchy's result appears as an explicit stimulus. In addition, Sylow operated within a subgroup of the symmetric group—just as Cauchy had—both in stating the special-case Sylow theorems in 1867 and, in 1872, in his proof of the general ones. Tellingly, Sylow spoke not of a "group"—Galois' word—but of a "system of conjugate substitutions"—Cauchy's.

The story of how Sylow's generalizations of Cauchy's theorem came to be published at all recalls the direct influence of Paris mathematicians. Sylow had proved his theorems as early as 1870, but he

[13] In at least one case personal intervention brought notice of Cauchy [1845] to the attention of a creative contributor to Galois theory. Beginning in 1852, complaining that Liouville had not provided the commentary he had promised in his 1846 publication of Galois, Enrico Betti published a series of proofs of theorems enunciated there without proof. Betti cited and used [Cauchy, 1815] in 1852 but not [Cauchy, 1845]; only in 1855, after a visit by J.J. Sylvester to Italy, did Betti, citing "the advice of a great geometer who has honored me with his friendship," make note of the more recent work of Cauchy, so close in spirit to his own. See Toti Rigatelli [1989, 59–61].

[14] Galois refers to an 1826 manuscript of Abel's notoriously ignored by Cauchy, who also received Galois' 1829 manuscript. This English translation of the "Preface" is from Rothman [1982, 97]. Andrea Del Centina [2002] has just located the missing pages of Abel's original manuscript.

withheld them from publication for at least two years until one of Liouville's former students from the École Polytechnique, Camille Jordan, on a visit to Norway, assured Sylow that the theorems were both new and significant.[15]

Especially with respect to Sylow's work, but also with respect to Dedekind's, the question arises of why the techniques of Cauchy [1845] did not lead to the Sylow subgroups, but the work of Galois did. To the insightful comments of Scharlau [1988, 43] that Galois' project of the solubility of equations by radicals led Sylow to a wide-ranging and penetrating study of the structure of finite groups, and therefore to a predominant role in the development of 19th-century group theory, I believe it plausible to add that the logical gap in Cauchy [1845] hindered advance by that avenue. Sylow wondered about generalizing Cauchy's theorem as early as 1863; yet he did not construct Sylow subgroups as Cauchy had attempted. Sylow's hesitations to publish may have arisen from modesty, but the fact is that his worry that someone else had found a proof may easily have originated from the incorrect construction of Sylow subgroups in S_n in Cauchy [1845]. Qualified mathematicians considered the proof difficult. "As late as 1878," comments Waterhouse [1980, 281], "Netto could begin a paper with the remark that the proof of Cauchy's theorem was as recondite as that of Lagrange was simple."

5. Camille Jordan

When the Paris mathematical school in the person of Camille Jordan synthesized permutation theory and Galois' theory of equations and articulated the result in terms comprehensible to other mathematicians, the proof of Cauchy's theorem was incorporated into the new structure, but in a form that silently corrected its error. Dahan Dalmedico agrees with earlier writers (e.g., [Wussing, 1969, 1984, 141–142]) on the history of group theory on the direct line of mathematical affiliation from Cauchy [1845] and Galois [1846] to Jordan [1870]. Speaking of Jordan's 1861 thesis, she notes:

> Thus in this thesis and its appendix we find for the first time a comparative presentation of the two contributions, the two points of view, of E. Galois and of A.-L. Cauchy [Dahan Dalmedico, 1980, 314–315]

and of an 1866 textbook presentation of the theory of permutations by Serret, she concludes [Dahan Dalmedico, 1980, p. 316]:

> In fact, this edition of 1866 marks a sort of apogee for Cauchy's studies on substitutions. Beginning with 1870, they only appear as mediated by the work of Jordan. Although that scholar had at one time relied upon the mathematical techniques devised by Cauchy with considerable success, nevertheless the intellectual re-creation of the originality of Galois by Camille Jordan proved so powerful that it absorbed Cauchy and all other protagonists.

The picture resembles Euclid's *Elements* in its ability to erase the memory of previous contributions to the mathematical theory of groups.

[15] See Lutzen [1992, 442–443, 446], where in a letter to his Danish colleague Julius Petersen dated 13 September 1870 Sylow stated, "I am able to prove this using a theorem from the theory of substitutions which I have already known for a long time but I have not published."

Engaged as we have seen in a much larger enterprise than simply combining the ideas of Galois and Cauchy, Jordan does not hesitate to improve the proof of Cauchy's theorem. "This beautiful theorem," he announces "is due to Cauchy, who proved it pretty much as follows: ..." [Jordan, 1870, 26].

Given any two groups (let us call them, recalling Cauchy, P and Q) contained in a third group J, of orders M, N, and O, respectively; then Jordan's "Lemma I" states that the number of substitutions U which satisfy no relation of the form

$$P_h U = U Q_k$$

is either zero or a multiple of MN. Just as Cauchy had, so Jordan showed that any permutation U fulfilling the condition generated MN "derived" permutations of the form

$$P_h U Q_k$$

(as the indices h and k run through the M and N possibilities) is distinct from all others of that form. For if, for some h, h', k, k',

$$P_h U Q_k = P_{h'} U Q_{k'}$$

were true, then multiplication by inverses would give

$$P_h^{-1} P_{h'} U = U Q_k Q_{k'}^{-1},$$

which is of the form

$$P_i U = U Q_j$$

for some i, j; a form impossible by supposition. A second permutation V which can satisfy the same condition, if such a permutation exists and is distinct from U, will also generate a group of order MN, distinct from each other and from the group generated by U.

Thus Jordan's proof for this lemma is indeed "pretty much" the same as Cauchy's, although it is deployed in order to prove the relationship directly, without the use of modular arithmetic with which Cauchy graced his statement of the original theorem. Jordan draws the same corollary from this lemma—that if no permutation of Q is "similar" to any permutation of P, then the order O of the containing group is divisible by MN.

Jordan's second lemma also looks familiar: if p is a prime number and p^f the highest power of p which divides $n!$, then one can construct a subgroup of order p^f from the group of permutations among n letters. His proof, however, differs substantially from Cauchy's construction of permutations of "the first kind," of "the second kind," and the resulting sum resulting from reindexing. Rather, it both sidesteps the erroneous assurance of commutative permutations and introduces a semidirect product in a proof which employs the use of induction on f.

Suppose $n < p$. The highest power of p that divides $n!$ is $p^0 = 1$. This subgroup is the identity element, and its existence satisfies the lemma for $f = 0$. From now on we must have $n > p$.

Now let us suppose that the lemma is true for all powers of p less than p^f; we will show it true for p^f. Note that any whole number n which is larger than p^f yet smaller than p^{f+1} can be written as

$$n = qp + r,$$

where $q < p^f$ and $r < p$. Let us then focus on the $q \cdot p$ of the given n letters and partition them into q sets $(a, b, c, \ldots, a_1, b_1, c_1, \ldots)$ of p letters each.

Call P the permutation that circulates the letters $a, b, c, \ldots$, without displacing any; P_1 is the cyclic permutation of $a_1, b_1, c_1, \ldots$, and so on. (These correspond to the "first kind" of Cauchy.) We also have the group T $(= ta, tb, tc, \ldots)$ composed of the circular replacements among the q systems of the corresponding letters. (These are Cauchy's "second kind.") The group G derived from all possible combinations of these $P, P_1, \ldots$, and of T will have, Jordan asserts, elements of the form

$$t_\mu P^\alpha P_1^\beta \cdots.$$

Suppose this were not so, and there were an element $P^\alpha t_\mu P^\beta P_1^\gamma \cdots$, that is, an element which did not have its element of T written on the left and could be written in no other way. This would equal

$$t_\mu t_\mu^{-1} P^\alpha t_\mu P^\beta P_1^\gamma \cdots$$

with $[t_\mu^{-1} P^\alpha t_\mu]$ equivalent to some P_i^α. This fact follows from what has been proved previously, but the reader gets a concrete example "pour fixer les idées": if t_μ replaces each of the letters $a, b, c, \ldots$, with the letters $a_1, b_1, c_1, \ldots$, then (proceeding from the left, following Jordan)

t_μ^{-1}	replaces a_1 with a;
P	replaces a with b;
t_μ	replaces b with b_1;

which of course equals the action of P_1 $(a_1 \to b_1)$.

Now, how many distinct elements are there in the group G of the form $t_\mu P^\alpha P_1^\beta \cdots$? If M is the order of the group T, the order of $G = Mp^q$; and since one can construct, by the induction hypothesis, for any largest power p^σ contained in $q!$ (where we selected $q < p^f$) a group T of order p^σ, the total comes to

$$p^{\sigma+q} = p^f,$$

which was to be proved.

With these two lemmas Jordan proceeded to the proof of Cauchy's theorem, that there exists a subgroup of order p for every prime number p dividing the order of the group. He did not construct such a subgroup of order p, however; rather, he considered H, defined as that subgroup of the symmetric group of n letters which contains no permutation of order p. Nor can any be a multiple of p, for a permutation of order λp, raised to the power λ, would be of order p, excluded by hypothesis.

All of the permutations of the subgroup G, the existence of which has just been demonstrated, must by contrast divide p^f, the order of the group,[16] and so be a power of p. None of the permutations contained in G, therefore, are similar to any of the permutations in H.

By the corollary to Lemma I, the product of the orders of G and H must now divide the order of the group containing both of them. Since the symmetric group of n letters is of order $n!$, we have that the order of H divides

$$\frac{1 \cdot 2 \cdot 3 \cdots n}{p^f},$$

[16] This is true by the theorem of Lagrange, that the order of a subgroup must divide the order of the group, which immediately precedes Cauchy's theorem in Jordan [1870]. Ironically, the circumstance that Jordan titled the section "The theorems of Lagrange and Cauchy" led Kiernan [1971, 96] to state erroneously that Jordan attributed Lagrange's theorem jointly to Lagrange and Cauchy.

and so contains no factor of p.

With such a statement, Jordan's proof is complete: the statement that all groups which contain no element of order p have orders that are prime to p is the contrapositive of the statement that all groups whose order is divisible by p contain at least one element (and therefore a cyclic subgroup) of order p.

With its use of the subgroups of the symmetric group of permutations on n arbitrary letters, its deployment of double cosets and its general strategy, [Jordan, 1870] resembles Cauchy's proof but circumvents the logical gap pointed out above, since Jordan's process of correction circumvented just that section of Cauchy's proof. From the point of view of Sylow (or of Dedekind) hypothetically attempting to modify [Cauchy, 1845] just enough to obtain Sylow's result as an extension of it, it is interesting that Jordan's silent emendation of 1870 did not lead him, more than a year after publication, to tell Sylow that his first theorem was contained in Jordan's second lemma cited above. This suggests to me that the preference for Galois theory over permutation group theory is not the only reason [Cauchy, 1845] did not lead directly to Sylow's theorems; the logical tangle in the proof was substantial.

6. George A. Miller

Cauchy's theorem went into eclipse with the arrival of the abstract conception of group theory. Once Walther Dyck's work [Dyck, 1882] enunciated with precision ideas which were already widely shared, it was only two years until Georg Frobenius found a proof of Sylow's first theorem which used neither the subgroups of the symmetric group nor Cauchy's theorem nor his construction of a special case of Sylow subgroups. That "class-equation" proof of Frobenius became standard, so much so that subsequent textbooks on group theory, since Sylow's theorems have far more mathematical consequence than Cauchy's theorem, typically relegated the latter to a footnote.

Scharlau [1988, 49] warns us of the danger of an overemphasis on the mutual incomprehension between the permutation-theoretic view of groups and the abstract view; nevertheless, Wussing's picture of a reworking of permutation-theoretic results within the framework of the abstract group theory [Wussing, 1984, 243–244] goes far to answer Waterhouse's [1980, 279] question of what Frobenius was up to when in 1884 he reproved Sylow's theorem. Thomas Hawkins [1981] adds explanatory power when he displays Frobenius as a member of the "Berlin School" of mathematics, producing proofs in algebra without appeal to "generic" argument.[17] Hawkins refers to the example of Karl Weierstrass, who had acted as Frobenius' thesis advisor and in 1868 carried out an exhaustive study of the elementary divisors in quadratic and bilinear forms. He quotes Leopold Kronecker in disapproval of "generic" reasoning; Frobenius cites Kronecker prominently in both [Frobenius, 1887a, 1887b] of the papers he devoted to reproving Sylow's theorem, papers which appeared in a journal edited, in Berlin, by Kronecker and Weierstrass.

William Waterhouse's penetrating study [Waterhouse, 1980] of the early proofs of Sylow's theorem develops all the major lines of mathematical argument in Frobenius; it is of note here that those lines only

[17] As a possible example of what style of algebra was to be avoided, I propose E.E. Kummer, who retired from the chair of mathematics in Berlin only in 1884, and who, to quote Edwards [1977, 382], "was chronically optimistic. He attacked problems by means of "induction," that is, by extensive numerical computations of specific examples, from which he would then abstract the theorems to be proved; once he became convinced that a particular statement was true, he was prone to overlook deficiencies in his proof of it."

implicitly demonstrate Cauchy's theorem. Frobenius mentions Cauchy only to specify what he wants to banish from his assumptions.

The implicit inclusion was clear enough for Dedekind to compare the line of proof to his own work on Cauchy's theorem in the letter to Frobenius cited above,[18] and it was also clear enough for the American mathematician George A. Miller to elaborate into a proof specifically of Cauchy's theorem, noting "[t]he main features of this method of proof are due to Frobenius" [Miller, 1898, 323]. It is Miller's explicit application of Frobenius which appears in, among others, the textbooks of Hall [1959, 43–44], Herstein [1964, 74], and Birkhoff and Mac Lane [1967, 468], but we must note that, with Frobenius' proof, Cauchy's theorem had entered the ranks of elementary results, of interest only to beginners in group theory and those who teach them.

Miller himself inserted the proof at the beginning of a paper "On an Extension of Sylow's Theorem," with the laconic comment that "Since we shall employ Cauchy's theorem in what follows it seems desirable to give a simple proof of it."

Suppose the group G, whose order is divisible by a prime number p, is Abelian. If this group is generated by a single element of order np, then that single element $(s)^{np} = (s^n)^p = 1$, and the element s^n is of order p, as desired. An Abelian group generated by a set of elements $s_1, s_2, \ldots, s_r$ cannot have an order which is divisible by p unless some one of the commuting elements is of an order divisible by p. Some power of that element will then be of order p.

So suppose that G is non-Abelian. Then make the induction assumption that the theorem is true for all groups of order less than np, and proceed to prove it true for a group of order np.

Following Frobenius, Miller forms what we now call the normalizer of G: "the largest subgroup of G that transforms a given [element] into itself," that is, $\{g \in G \mid gag^{-1} = a\} \equiv N(a)$. The index of $N(a)$ in G, in modern notation $|G|/|N(a)|$, Miller states without proof, is the number of conjugates of this element a. In his proof, Frobenius devoted a few sentences to persuade his reader of this result, no doubt so that the proof could stand alone. The normalizer was as we have seen important in the elaboration of Galois theory, and it had been in use since Jordan's day. Almost 30 years later Miller feels that he can take its partitioning of G for granted.

The order of G is then equal to the sum of all the classes—the word was introduced by Frobenius—of the elements of G,

$$|G| = |G|/|N(e)| + |G|/|N(a)| + \cdots;$$

in Miller's notation,

$$g = g/g_1 + g/g_2 + \cdots + g/g_k. \qquad (*)$$

Miller points out that g_1, certainly, and perhaps other normalizers have an index of 1. That is, all the elements of G which commute with all the elements of G, a set which we now term the center of G, form an Abelian subgroup of G. "If the order of this subgroup is not divisible by p some $g_\beta < g$ must be divisible by p, since the second member of $(*)$ must be divisible by this number," concludes Miller. He then goes on with the first theorem of his own.

[18] Hawkins [1971, 143] credits this same letter of 8 February 1895 from Dedekind to Frobenius with stimulating the latter to develop his theory of group characters.

As with Cauchy, as with Dedekind, the last sentence of Miller's proof glides over a step. Let us look at it more closely. We have a series of ones on the right-hand side of the class equation $(*)$,

$$g = 1 + 1 + \cdots + 1 + g/g_{m+1} + \cdots + g/g_k,$$

let us say m of them, and the question is whether the sum of these ones is divisible by p. By hypothesis the orders of the normalizers

$$g_{m+1}, \ldots, g_k$$

are none of them divisible by p (else the theorem would be proved), so each of the quotients

$$g/g_{m+1}, \ldots, g/g_k$$

has a numerator divisible by p and a denominator which is not. Each quotient must therefore be divisible by p, and the sum of all the nonunit quotients must also be divisible by p. With g itself divisible by p, and all the nonunit quotients divisible by p, we must have a *center* divisible by p. It is therefore possible to find in the Abelian subgroup an element of order p, as already shown.

In contrast to those of Cauchy and of Dedekind, one senses that the minor slip in Miller's presentation has nothing to do with a struggle with inadequate notation or novel mathematical structures; the proof of Cauchy's theorem has become a preliminary to be disposed of on the way to something of more pertinent interest. A reader of Weber's influential textbook [Weber, 1894, second ed. 1898] or that of Burnside [1897, second ed. 1911] finds Sylow's theorems proved, in each case with full credit to Frobenius and his class equation, but Cauchy's theorem mentioned only in a footnote or in a corollary to the first Sylow theorem. Indeed, in Miller's own monograph on the theory of finite groups [Miller et al., 1916], where with frequent historic asides Sylow's theorem is proved twice, once with double cosets and once with a class equation, Cauchy's theorem is mentioned once in passing and its proof is never adduced.[19]

7. James H. McKay

In an unusually discursive "Preface" to an undergraduate textbook, Saunders Mac Lane and Garrett Birkhoff [1967, v–ix], both of whom have had distinguished research and pedagogical careers, discuss the history, not so much of algebra as of its instruction in U.S. universities. Axiomatic modern algebra, which unifies so many branches of higher mathematics, made its way into the graduate curriculum with B.L. van der Waerden's influential *Moderne Algebra* during the 1930s. Undergraduate instruction in the same subject followed during the '40s and '50s.

As noted above, Cauchy's theorem only reappears in the last of those decades, for the reason, I believe, that the undergraduate instruction of abstract algebra has the need to develop an appreciation for, or better the capacity to construct, a mathematical proof. The presentation of a proof for Cauchy's theorem, followed by a separate proof for the related generalization Sylow's theorem, helps to develop this ability in stages.

[19] This in spite of the fact that Cajori's evaluation of Cauchy as the founder of the theory of finite groups cites Miller as its source.

Textbook writers tend to take ideas from one another without attribution much more frequently than researchers do, and I have not yet found the route from Miller's formulation of the Frobenius class-equation proof to its use in the undergraduate textbooks of the 1950s. On the other hand, there is a clear trail for the beautifully compact proof now most commonly used [McKay, 1959]. It went from an original version which displayed a maximum of simplicity, to a graduate student version which recast it in more modern mathematical tools, to an undergraduate version. The audience is now the university student, and the form of proof changes according to the pedagogical purpose at hand.

In polished, lapidary form—nine sentences, of an average of 11 words each—McKay [1959, 119] made use of a special class equation. Given the group G of order g, divisible by the prime p; we want to show that there are kp elements of G satisfying the equation $x^p = 1$.

Form then S, the set of p-tuples of elements of G which have the property that, when all multiplied together, they equal the identity

$$S = \{(a_1, \ldots, a_p) \mid a_1 a_2 \cdots a_p = 1\}.$$

The first $p - 1$ members of this p-tuple can be any elements, but the last one is fixed; it has to be the inverse of that element which is the product of the first p^{-1} components. The set S, as a result, can have at most g^{p-1} members.

> Define an equivalence relation on S by saying two p-tuples are equivalent if one is a cyclic permutation of the other. If all components of a p-tuple are equal then its equivalence class contains only one member. Otherwise, if two components of a p-tuple are distinct, there are p members in the equivalence class.

More than two components can be different, of course; there are just p distinguishable circular permutations of p elements, at least two of which are distinct.

The p-tuples composed of all the same element are solutions to the equation $x^p = 1$. Let us suppose there are r of those. Of the p-tuples with at least two distinct elements, let us say there are s. That of course exhausts the possibilities, so, using these two equivalence classes as a way to count the number of p-tuples which are members of the set S,

$$r + sp = g^{p-1}.$$

Since p divides g, it divides the right-hand side; in order to divide the left-hand side we must have p divides r. Thus there are $kp = r$ elements of G for which $x^p = 1$.

There *is* a class equation, with the same bunch of ones on one side and a number divisible by p on the other, but we do not have to form a normalizer or quotient groups or talk about abstract congruences. No Abelian special case, no reduction of the general case to the special—this is probably the simplest known form of the proof of Cauchy's theorem.

Introductory mathematical pedagogy at the university level, however, is not characterized completely by the effort to make it easily digestible. As Thomas Kuhn [1970] emphasized, textbook writers prepare the student for autonomous problem-solving as quickly as possible. A shift of point of view at the research level requires the introduction of the new "paradigm" down to the entry level. "Recent years have seen striking developments in the conceptual organization of Mathematics," begin Mac Lane and Birkhoff [1967] in the preface previously cited, with a somewhat pretentious capital letter. They continue, after specifying the developments,

> This book proposes to present algebra for undergraduates on the basis of these new insights. In order to combine the standard material with the new, it seemed best to make a wholly new start. [Mac Lane and Birkhoff, 1967, vi]

As another distinguished mathematician and pedagogue, Anthony W. Knapp, indicated not long ago, group representations "play a critical role" in many areas of modern mathematics [Knapp, 1996, 410]. The classic work of Frobenius initiated modern study of the topic, but we express group representations today in terms of group action on a vector space. In short, in order to employ the current mathematical tools as soon as possible, McKay's proof of Cauchy's theorem has been rewritten using the technique of group action, first at the graduate level [Hungerford, 1974, 93][20] and then, with explicit citation of the filiation, at the undergraduate [Fraleigh, 1989, 190–203]. Here is how the latter does this.

We begin with the definition. If X is any set and G is a group, the group action of G on X is a map from X to X such that the group identity makes no change in any element of X and the action is associative, that is,

$$(g_1 g_2)(x) = g_1(g_2 x)$$

for all g in G and x in X. We show that the group action partitions X into equivalence classes, calling the class containing a particular x its "orbit," denoted Gx. We then rewrite the class equation in this notation. The number of elements in the set X equals the sum of the number of elements in the orbits. Some orbits, like the identity element in G, have one-element orbits. Separate the orbits into two terms, then, the number of one-element orbits X_G and the orbits with more than one element, designated Gx_i:

$$|X| = |X_G| + \sum |Gx_i|.$$

If now the prime number p divides the group G, then it will divide the number of elements contained in a single nonunit orbit, Gx_i for any particular value of i (since Gx_i runs through all of G). It will thus divide the sum of such numbers, the second term on the right-hand side. If ever we have a p which divides $|X|$, the order of the set, it will thus have to divide the number of one-element orbits as well.

The translation of McKay [1959] is now straightforward. The set X is the set of p-tuples, on which the "group action" is the cyclical permutation of the p components of each p-tuple. This group action represents a group of order p, each of which shifts the p-tuple by one place. The group identity does not shift anything at all. The machinery clanks and squeaks a little, since our *original* group, which we have always designated G and which we are assuming has an order divisible by p, is quite distinct from the group which in our proof forms the group action. Fraleigh uses the Greek letter σ to designate the cyclic permutation of order p.

All we have to do is set up the class equation in the new terms. $|X|$ is divisible by p—it is a set of p-tuples. The number of one-element orbits $|X_G|$ is thus also divisible by p. It is nonzero, since it includes the identity element. So there are elements $a \neq e$ such that $(a)^p = e$ or identity; that is, there are subgroups of order p.

Our exercise at translation is not without its merit, since the patient reader will agree how easily a similar translation into group action terms could be accomplished for the Frobenius–Miller proof as well. However, such a translation is not possible, it seems to me, for the original double-coset proof of Cauchy. That proof contains no class equation and, after all, relies upon properties of the symmetric group and an especially-constructed subgroup of order the highest power of p which will divide n factorial (just the "generic reasoning" Weierstrass opposed). The original proof of Cauchy, however, is also the first use of

[20] In a footnote Hungerford credits R.J. Nunke for the line of proof, so I believe Nunke suggested the rewriting.

the wreath product, named and popularized in 1937 by the Hungarian–American mathematician George Pólya.[21]

At the end of our transit of the mathematical life of Cauchy's theorem, then, reflecting on the many surprises therein—the sensational suspicion of plagiarism at its birth, the foreshadow of Frobenius' proof in the unpublished papers of Dedekind, the near-disappearance of the theorem and its resuscitation as a vehicle to teach undergraduates how to prove a result—perhaps the greatest one of all is the fact that our 56-year-old baron, "writing his mathematics during one week, giving his manuscript to the secretaries of the Academy at the Monday meeting the next week, and seeing his notes published the next Monday after that" [Neumann, 1989, 296], retains his unique originality of method after 150 years.

Acknowledgments

I thank Rudi Beyl, without whose assistance this paper would be of considerably less interest, and Curtis Wilson, Jesper Lutzen, two referees, and editor of *Historia Mathematica* Umberto Bottazzini for reading previous drafts and providing constructive criticism. My research would not have been possible without the help of the Multnomah County Library's interlibrary loan service. A previous version of this paper was presented to the Northwest Independent Scholars Association.

References

Belhoste, B., 1985. Cauchy, un mathématicien légitimiste au XIXe siècle. Belin, Paris (quotations from the English edition: Augustin-Louis Cauchy. A Biography. Springer-Verlag, New York, Berlin, 1991).

Birkeland, B., 1996. Ludvig Sylow's "Lectures on Algebraic Equations and Substitutions, Christiana, 1862." Hist. Math. 23, 182–199.

Burnside, W., 1897. Theory of Groups of Finite Order. Cambridge Univ. Press, Cambridge, UK.

Cajori, F., 1919. A History of Mathematics, second revised ed. Macmillan, New York.

Cauchy, A.-L., 1815. Sur le nombre des valeurs qu'une fonction peut acquérir lorsqu'on y permute de toutes les manières possibles les quantités qu'elle renferme. J. l'Ecole Polytech. 10, 1–112;
Reprinted in: Oeuvres complètes, vol. 1. Gauthier-Villars, Paris, 1932, pp. 64–169.

Cauchy, A.-L., 1845. Mémoire sur les arrangements que l'on peut former avec des lettres données et sur les permutations ou substitutions à l'aide desquelles on passe d'un arrangement à un autre. Exercises d'analyse et de physique mathématique 3, 151–252;
Reprinted in: Oeuvres complètes, vol. 13, second series. Gauthier-Villars, Paris, 1932, pp. 171–282 (citations are from the reprinted edition).

Cauchy, A.-L., 1845/1846. Comptes rendus de l'Académie des Sciences de Paris, 21(1845)–22(1846) (25 papers);
Reprinted in consecutive order: Oeuvres Complètes, vol. 9, first series. Gauthier-Villars, Paris, 1896, p. 277; vol. 10, pp. 1–68 (citations are from the reprinted edition).

[21] See Pólya [1937, 1984, 341] and the discussion in Read [1987, 275]. A detailed construction of the Sylow subgroup of order p^n as an illustration of the wreath product appears in the influential monograph [Hall, 1959, 81–83] on group theory. The algorithm used by Cauchy is so similar that Waterhouse [1980, 282] terms it "the one still familiar;" very likely it is familiar due to Hall's exposition. Note, finally, that the wreath product is present in the very structure of McKay's succinct proof of Cauchy's theorem: McKay selects just the diagonal elements of the group table of a finite group and forms their wreath product. Just what was confusing in Cauchy's original proof has been creatively employed to distill what was difficult into a model of concision.

Dahan Dalmedico, A., 1980. Les travaux de Cauchy sur les substitutions. Étude de son approche du concept de groupe. Arch. Hist. Exact Sci. 23, 279–319.
Dedekind, R., 1932. In: Fricke, R., Noether, E., Ore, O. (Eds.), Gesammelte mathematische Werke, 3 vols. Vieweg, Braunschweig;
Reprinted Chelsea, Bronx, NY, 1968.
Del Centina, A., 2002. The manuscript of Abel's Parisian mémoire found in its entirety. Hist. Math. 29, 65–69.
Dieudonné, J. (Ed.), 1978. Abrégé d'histoire des mathématiques, 2 vols. Hermann, Paris.
Dyck, W., 1882. Gruppentheoretische Studien. Math. Ann. 20, 1–23.
Edwards, H.M., 1977. Postscript to "The background of Kummer's proof. . . ." Arch. Hist. Exact Sci. 17, 381–394.
Fraleigh, J.B., 1989. A First Course in Abstract Algebra, fourth ed. Addison–Wesley, Reading, MA.
Frobenius, G., 1887a. Neuer Beweis des Sylowschen Satzes. Journal für die reine und angewandte Mathematik 100, 179–181;
Reprinted in: Gesammelte Abhandlungen, Serre, J.-P. (Ed.), vol. 2. Springer-Verlag, Berlin, pp. 301–303.
Frobenius, G., 1887b. Über die Congruenz nach einem aus zwei endlichen Gruppen gebildeten Doppelmodul. Journal für die reine und angewandte Mathematik 101, 273–299;
Reprinted in: Gesammelte Abhandlungen, Serre, J.-P. (Ed.), vol. 2. Springer-Verlag, Berlin, pp. 304–330.
Galois, E., 1846. Oeuvres mathématiques. J. Math. Pures Appl. 9, 381–444;
Reprinted in: Bourgne, R., Azra, J.P. (Eds.), Ecrits et mémoires mathématiques. Gauthier–Villars, Paris, 1962, pp. 43–71.
Grattan-Guinness, I., 2000. Christianity and mathematics: kinds of link, and the rare occurrences after 1750. Physis. Rivista Internazionale di Storia della Scienza 37, 467–500.
Hall, M., Jr., 1959. The Theory of Groups. Macmillan, New York.
Hawkins, T., 1981. The Berlin school of mathematics. In: Mehrtens, H., Bos, H., Schneider, I. (Eds.), Social History of Nineteenth Century Mathematics. Birkhäuser, Boston, pp. 233–245.
Hawkins, T., 1971. The origins of the theory of group characters. Arch. Hist. Exact Sci. 7, 142–170.
Herstein, I.N., 1964. Topics in Algebra. Blaisdell, New York.
Hungerford, T.W., 1974. Algebra. Springer-Verlag, New York.
Jordan, C., 1870. Traité des substitutions et des équations algébriques. Gauthier–Villars, Paris.
Kiernan, B.M., 1971. The development of Galois theory from Lagrange to Artin. Arch. Hist. Exact Sci. 8, 40–154.
Knapp, A.W., 1996. Group representations and harmonic analysis from Euler to Langlands. Notices Am. Math. Soc. 43, 410–415, 537–549.
Kuhn, T.S., 1970. The Structure of Scientific Revolutions, second ed. Univ. of Chicago Press, Chicago.
Littlewood, J.E., 1953. A Mathematician's Miscellany. Methuen, London;
Reprinted as: Bollobás, B. (Ed.), Littlewood's Miscellany. Cambridge Univ. Press, New York/Cambridge, 1986.
Lutzen, J., 1992. The mathematical correspondence between Julius Petersen and Ludvig Sylow. In: Demidov, S., et al. (Eds.), Amphora. Festschrift for Hans Wussing. Birkhäuser, Boston, pp. 439–467.
Mac Lane, S., Birkhoff, G., 1967. Algebra. Macmillan, New York.
Mazzotti, M., 1998. The geometers of God: mathematics and reaction in the Kingdom of Naples. Isis 89, 674–701.
McKay, J.H., 1959. Another proof of Cauchy's group theorem. Am. Math. Monthly 66, 119.
Miller, G.A., 1898. On an extension of Sylow's theorem. Bull. Amer. Math. Soc. 4, 323–327.
Miller, G.A., Blichfeldt, H.F., Dickson, L.E., 1916. Theory and Applications of Finite Groups. Wiley, New York;
reprinted Dover, New York, 1961.
Neumann, P.M., 1989. On the date of Cauchy's contributions to the founding of the theory of groups. Bull. Austral. Math. Soc. 40, 293–302.
Nicholson, J., 1993. The development and understanding of the concept of quotient group. Hist. Math. 20, 68–88.
Pólya, G., 1937. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. Acta Math. 68, 145–254;
Reprinted in: George Pólya, Collected Papers, Rota, G.C. (Ed.), vol. 4. MIT Press, Cambridge, MA, 1984, pp. 308–419.
Pourciau, B., 1997. Reading the master: Newton and the birth of celestial mechanics. Amer. Math. Monthly 104, 1–19.
Read, R.C., 1987. Pólya's theorem and its progeny. Math. Mag. 60, 275–282.
Roth, R., 2001. A history of Lagrange's theorem on groups. Math. Mag. 74, 99–108.
Rothman, T., 1982. Genius and biographers: the fictionalization of Evariste Galois. Amer. Math. Monthly 89, 84–106.
Scharlau, W. (Ed.), 1981. Richard Dedekind 183l/1981: eine Würdigung zu seinem 150. Geburtstag. Vieweg, Braunschweig.
Scharlau, W., 1988. Die Entdeckung der Sylow-Sätze. Hist. Math. 15, 40–52.

Scholz, E., 1990. Die Entstehung der Galoistheorie. In: Scholz, E. (Ed.), Geschichte der Algebra. Eine Einführung. BI Wissenschaftsverlag, Mannheim.

Speiser, A., 1937. Die Theorie der Gruppen von endlicher Ordnung, third ed. Julius Springer, Berlin.

Sylow, L., 1872. Théorèmes sur les groupes des substitutions. Math. Ann. 5, 584–594.

Toti Rigatelli, L., 1989. La mente algebrica. Storia dello sviluppo della teoria di Galois nel XIX secolo, Bramante Editrice, [n.p.].

van der Waerden, B.L., 1985. A History of Algebra from al-Khwarizmi to Emmy Noether. Springer-Verlag, New York.

Waterhouse, W.C., 1980. The early proofs of Sylow's theorem. Arch. Hist. Exact Sci. 21, 279–290.

Weber, H., 1894. Lehrbuch der Algebra, 3 vols. F. Vieweg, Braunschweig.

Wussing, H., 1969. Die Genesis des abstrakten Gruppenbegriffes. VEB Deutscher Verlag der Wissenschaften, Berlin. (Quotations from the English edition: The Genesis of the Abstract Group Concept. MIT Press, Cambridge, MA, 1984).