Irish Math. Soc. Bulletin
Number 88, Winter 2021, 57–68
ISSN 0791-5578

Four Group-theoretic Proofs of Wedderburn's Little Theorem

ALAN ROCHE

Abstract. Wedderburn proved in 1905 that a finite division ring is always a field. His result has intrigued generations of mathematicians, spurring generalizations and alternative proofs. The shortest, most elegant proof is surely Witt's from 1931, now the standard textbook treatment. Following a strategy of Zassenhaus, we present four overlapping group-theoretic proofs. The first uses a counting argument; the others hinge on properties of special classes of finite groups.

1. Introduction

Wedderburn's Little Theorem says that a finite division ring is a field. That is, if each nonzero element of a finite ring with identity has a multiplicative inverse then the ring is commutative. In the words of I. N. Herstein, the theorem "has caught the imagination of mathematicians because it is so unexpected, interrelating two seemingly unrelated things, the number of elements in a certain algebraic system and the multiplication of that system" [9]. In the words of Emil Artin, "this result of Wedderburn has fascinated most algebraists to a very high degree" [2]. As often in mathematics, the first proofs were somewhat clumsy. Quoting Artin again, Wedderburn [17] makes use of "divisibility properties which are hard to establish" and "several attempts were made to simplify the proofs." In 1931, Witt gave a short, elegant proof. It uses only elementary group theory in the form of the class equation (which Wedderburn also used) and easy divisibility properties. It is indeed a proof from THE BOOK (see [1, Chap. 6]).

Witt's argument, however, was far from the last word. There have been many other proofs—ones like Witt's that use essentially elementary methods and ones that stem from a broader perspective. As an example of the latter, note that Wedderburn's Big Theorem [18] puts the Little Theorem in a wider setting—the theory of central simple algebras and the Brauer group. This is a profound theory concerned, in its classical incarnation, with describing division algebras that are finite dimensional over their centres. Various results in the overall theory yield proofs of the Little Theorem. For instance, by properties of what are called cyclic algebras the theorem reduces to an easy calculation—checking surjectivity of the norm map for a finite extension of finite fields (see, for example, [10, Section 8.4]).

Our starting point is a technical lemma due to Zassenhaus [20]. It says that if $\mathbb{E}$ is a subfield of a finite division ring $\mathbb{D}$ then the normalizer and centralizer of $\mathbb{E}^\times$ in $\mathbb{D}^\times$ coincide. Zassenhaus goes on to deduce the Little Theorem by a purely group-theoretic argument (see Remark 1 below). We follow the same path: with Zassenhaus's lemma in hand, we apply tools from finite group theory to obtain the Little Theorem—four times. Out of a perverse sense of symmetry, we also include four proofs of Zassenhaus's lemma.

2020 *Mathematics Subject Classification.* 20D99, 16K20.
Key words and phrases. division ring, finite solvable group, Frobenius group.
Received on 24-7-2021; revised 19-11-2021.
DOI:10.33232/BIMS.0088.57.68.

The first proof of the Little Theorem is direct. Beyond the technical lemma, it uses only elementary group theory and simple counting. The other proofs rely on more advanced material: Burnside's Normal Complement Theorem, Carter subgroups, a property of Frobenius groups. For anyone new to these topics, we've tried to fill in enough background so that the bulk of the paper still makes sense. More precisely, a reader on friendly terms with (a) finite groups up to the Sylow theorems plus the notion of a nilpotent group and (b) fields and basic Galois theory should be able to follow the main thread of argument.

Three of our proofs share a common strategy: they proceed by showing that the maximal fields in a noncommutative finite division ring $\mathbb{D}$ (with $\mathbb{D}$ sometimes of minimal order) form a single conjugacy class (under the action of $\mathbb{D}^\times$). Since each element of $\mathbb{D}$ is contained in a maximal field, it follows that if $\mathbb{E}$ is a fixed maximal field in $\mathbb{D}$ then the group $\mathbb{D}^\times$ is a union of conjugates of $\mathbb{E}^\times$. A finite group, however, is never a union of conjugates of a proper subgroup (see Lemma 2 below), and so a noncommutative finite division ring cannot exist. The same strategy underlies the classic proof of the Little Theorem in van der Waerden's *Moderne Algebra*, the influential text based on lectures by Artin and Noether, first published in 1930–31. Van der Waerden's argument, which is due to Noether, uses early results in the theory of central simple algebras. It is probably the most reproduced proof after Witt's of the Little Theorem.[1]

Except for the first, our proofs are more artificial than Zassenhaus's. We could be accused (we have been) of using a series of sledgehammers to crack the Little Theorem. But they are such beautiful sledgehammers. The actual arguments, once the necessary background is in place, are short, requiring just a tap from any sledgehammers we wield. The proofs are meant as mathematical entertainments—an amusing application, we hope, of parts of finite group theory.

Acknowledgements. The treatment of Zassenhaus's lemma in Section 7 owes much to comments provided by a reviewer of a previous (pre-Bulletin) version of the paper. In addition, I've adopted several suggestions made by the referee of the current (Bulletin) version.

I would also like to note an older, more personal debt to two master expositors of mathematics, T. J. Laffey and D. L. McQuillan, under whose stimulating guidance I first learned of the world containing Wedderburn's theorem.

Notation

Throughout $\mathbb{D}$ denotes a finite divison ring. The centre of $\mathbb{D}$ is a finite field which we write always as $\mathbb{F}$. For any ring R (with identity), $R^\times$ denotes the group of units of R.

We write q for the cardinality of $\mathbb{F}$ and q^N for the cardinality of $\mathbb{D}$, so that N is the dimension of $\mathbb{D}$ as an $\mathbb{F}$-vector space.

For H a subgroup of a group G, the normalizer and centralizer of H in G are written as $N(H)$ and $C(H)$. The ambient group G should be clear from context. In fact, we work mostly with a subfield $\mathbb{E}$ of $\mathbb{D}$. Then $N(\mathbb{E}^\times)$ and $C(\mathbb{E}^\times)$ always mean the normalizer and centralizer in $\mathbb{D}^\times$ (that is, the ambient group is invariably $\mathbb{D}^\times$).

[1]Van der Waerden, Witt and Zassenhaus were part of the remarkable flowering of abstract algebra in Germany in the 1920s and early 1930s in which Artin at Hamburg and Noether at Göttingen were leading figures. Witt was Noether's doctoral student and Zassenhaus was Artin's. Van der Waerden studied at Göttingen and was in the group (or ring) of young mathematicians centred around Noether; he also spent a year at Hamburg where he worked closely with Artin [16]. And then the cataclysm. In March 1933 the Nazis took power. In April a decree dismissed "non-Aryans" from government institutions including universities. Within weeks, a vibrant mathematical culture was destroyed [12]. Hilbert, when asked in 1934 by the new minister of education about mathematics at Göttingen "now that it was freed of the Jewish influence," replied: "Mathematics at Göttingen? There is really none anymore" [13, p. 205].

For S a finite set, $|S|$ denotes the number of elements in S. Given a subset T of S, we write $S \setminus T$ for the difference of S and T, that is, the set of elements of S that do not belong to T.

2. Zassenhaus's Lemma

The following technical result, due to Zassenhaus ([20, p. 59]), underpins each of our approaches to Wedderburn's theorem.

Lemma 1. *If $\mathbb{E}$ is a subfield of a finite division ring $\mathbb{D}$, then the normalizer and centralizer of $\mathbb{E}^\times$ in $\mathbb{D}^\times$ coincide.*

We only need the statement for now and so defer the task of proving the lemma to Section 7.

Remark 1. The formulation in [20] is superficially different: it says that if A is an abelian subgroup of $\mathbb{D}^\times$ then $N(A) = C(A)$. To obtain this version for a given abelian subgroup A of $\mathbb{D}^\times$, simply apply Lemma 1 to $\mathbb{F}(A)$, the subfield of $\mathbb{D}$ generated over $\mathbb{F}$ by A. Zassenhaus goes on to prove the Little Theorem by establishing a pleasant result: if a finite group G has the property that $N(A) = C(A)$ for every abelian subgroup A then G is abelian ([20, Theorem 7]).

Remark 2. Let $\mathbb{E}$ be a subfield of $\mathbb{D}$ that strictly contains the centre $\mathbb{F}$, so that the automorphism group $\mathrm{Aut}(\mathbb{E}/\mathbb{F})$ is nontrivial. A foundational result in the theory of central simple algebras—the Skolem–Noether Theorem—implies that each element of $\mathrm{Aut}(\mathbb{E}/\mathbb{F})$ is implemented by conjugation by some element of $\mathbb{D}^\times$ (see, for example, [5, Theorem 3.14]). The action of $N(\mathbb{E}^\times)$ on $\mathbb{E}$ by conjugation therefore induces an isomorphism $N(\mathbb{E}^\times)/C(\mathbb{E}^\times) \simeq \mathrm{Aut}(\mathbb{E}/\mathbb{F})$, and so $N(\mathbb{E}^\times) \neq C(\mathbb{E}^\times)$. Thus, if we were willing to make use of the theory of algebras, Lemma 1 yields a one-line (or one-paragraph) proof of the Little Theorem.

3. First Proof: Counting Maximal Fields

We assume that $\mathbb{D}$ has minimal order among noncommutative finite division rings and hope to find a contradiction.

Let $\mathbb{E}$ and $\mathbb{E}'$ be distinct maximal fields in $\mathbb{D}$. Then the subring $\langle \mathbb{E}, \mathbb{E}' \rangle$ generated by $\mathbb{E}$ and $\mathbb{E}'$ is noncommutative. It's also a division ring: in fact, each nonzero subring of $\mathbb{D}$ is a division ring (since each element of $\mathbb{D}^\times$ has finite order). By minimality of $|\mathbb{D}|$, it follows that $\langle \mathbb{E}, \mathbb{E}' \rangle = \mathbb{D}$. Thus $\mathbb{E} \cap \mathbb{E}'$ lies in the centre of $\mathbb{D}$, and so

$$\mathbb{E} \cap \mathbb{E}' = \mathbb{F}. \tag{1}$$

Recall our notation: $|\mathbb{F}| = q$ and $|\mathbb{D}| = q^N$. Let $\mathbb{E}_1, \ldots, \mathbb{E}_r$ be representatives of the distinct conjugacy classes of maximal fields in $\mathbb{D}$. We write $|\mathbb{E}_i| = q^{m_i}$ for integers $m_i \geq 2$ (for $i = 1, \ldots, r$). We use some counting and a simple estimate to show that $r = 1$, that is, the maximal subfields of $\mathbb{D}$ form a single conjugacy class.

Each element of $\mathbb{D}$ is contained in a maximal field. Using (1), we see that $\mathbb{D} \setminus \mathbb{F}$ is the disjoint union of the sets $\mathbb{E} \setminus \mathbb{F}$ as $\mathbb{E}$ varies through the maximal fields in $\mathbb{D}$. The number of distinct fields $x\mathbb{E}_i x^{-1}$ as x varies through $\mathbb{D}^\times$ is $[\mathbb{D}^\times : N(\mathbb{E}_i^\times)]$ (for $i = 1, \ldots, r$). Thus, by Lemma 1, the field $\mathbb{E}_i$ has $\dfrac{q^N - 1}{q^{m_i} - 1}$ conjugates, and so the conjugates of $\mathbb{E}_i \setminus \mathbb{F}$ account for $\dfrac{q^N - 1}{q^{m_i} - 1}\,(q^{m_i} - q)$ elements in $\mathbb{D} \setminus \mathbb{F}$ (for $i = 1, \ldots, r$). Therefore

$$q^N - q = \frac{q^N - 1}{q^{m_1} - 1}\,(q^{m_1} - q) + \cdots + \frac{q^N - 1}{q^{m_r} - 1}\,(q^{m_r} - q).$$

Rearranging, we obtain

$$\frac{q^N - q}{q^N - 1} = \frac{q^{m_1} - q}{q^{m_1} - 1} + \cdots + \frac{q^{m_r} - q}{q^{m_r} - 1}. \tag{2}$$

The left side is less than 1. Further, each term on the right side is greater than $\frac{1}{2}$. Indeed, for $m \geq 2$,

$$\begin{aligned} \frac{q^m - q}{q^m - 1} > \frac{1}{2} &\iff 2(q^m - q) > q^m - 1 \\ &\iff q^m - 2q + 1 > 0, \end{aligned}$$

and $q^m - 2q + 1 \geq q^2 - 2q + 1 = (q-1)^2 > 0$. Hence

$$\begin{aligned} 1 &> \frac{1}{2} + \cdots + \frac{1}{2} \quad (r \text{ terms}) \\ &= \frac{r}{2}, \end{aligned}$$

so $2 > r$ and $r = 1$.

From here, there are two ways to complete the argument.

Method 1. Since $r = 1$, the equality (2) now says that there is a positive integer $m < N$ such that

$$\frac{q^N - q}{q^N - 1} = \frac{q^m - q}{q^m - 1}. \tag{3}$$

Clearing denominators, we have

$$(q^m - 1)(q^N - q) = (q^m - q)(q^N - 1).$$

Expanding each side, cancelling common terms and rearranging then gives

$$q^{N+1} - q^N = q^{m+1} - q^m.$$

Thus $q^N(q-1) = q^m(q-1)$ and $N = m$—a contradiction.

Alternatively, the real function

$$\frac{q^x - q}{q^x - 1} = 1 - \frac{q-1}{q^x - 1}$$

is increasing (since q^x is increasing). Again, we see from (3) that $N = m$ which is absurd.

Method 2. We've proved that if $\mathbb{E}$ is a maximal subfield of $\mathbb{D}$ then

$$\mathbb{D}^\times = \bigcup_{x \in \mathbb{D}^\times} x\, \mathbb{E}^\times x^{-1}.$$

This is impossible: a finite group is never a union of conjugates of a proper subgroup.

Lemma 2. *Let H be a proper subgroup of a finite group G. Then*

$$\bigcup_{x \in G} xHx^{-1} \neq G.$$

Proof. The subgroup H has at most $[G : H]$ distinct conjugates in G. Indeed, for $x \in G$, the conjugate xHx^{-1} depends only on the left coset xH of H in G. Moreover, each

conjugate of H contains the identity element. Hence

$$\begin{aligned} |\bigcup_{x\in G} xHx^{-1}| &\leq \frac{|G|}{|H|}(|H|-1)+1 \\ &= |G| - \frac{|G|}{|H|} + 1 \\ &< |G|, \end{aligned}$$

which gives the result. □

Remark 3. Lemma 2 features in several proofs of the Little Theorem. As noted in the introduction, it provides the final step in the classic proof recounted by van der Waerden. Further, the group-theoretic principle that underlies Zassenhaus's proof (recalled in Remark 1) rests ultimately on Lemma 2. The lemma also plays a role in our third and fourth proofs.

4. Second Proof: Cyclic Sylow Subgroups

Next we rework an argument from 1964 due to T. J. Kaczynski [11][2]. The key observation is the following which we derive from Lemma 1 (the first several paragraphs of [11] give another route).

Proposition 1. *Every Sylow subgroup of $\mathbb{D}^\times$ is cyclic.*

Proof. We use induction on $|\mathbb{D}|$. The base case $|\mathbb{D}| = 2$ is trivial. Assuming that each Sylow subgroup of the multiplicative group of a division ring of order less than $|\mathbb{D}|$ is cyclic, we wish to show that $\mathbb{D}$ has the same property.

Let S be a Sylow l-subgroup of $\mathbb{D}^\times$ for some prime l. If $S \subset \mathbb{F}^\times$ then S is cyclic (by cyclicity of $\mathbb{F}^\times$). If S is not contained in $\mathbb{F}^\times$, then $S\mathbb{F}^\times/\mathbb{F}^\times$ is a nontrivial Sylow l-subgroup of $\mathbb{D}^\times/\mathbb{F}^\times$. In this case, we choose a nontrivial element $\beta\mathbb{F}^\times$ in the center of $S\mathbb{F}^\times/\mathbb{F}^\times$. Then, for each $\alpha \in S$, we have $\beta\alpha\beta^{-1} = \lambda\alpha$ for some $\lambda \in \mathbb{F}^\times$. Thus conjugation by β takes the field $\mathbb{F}(\alpha)$ to itself. Using Lemma 1, it follows that α and β commute. Hence $S \subset C(\beta)$, the centralizer of β in $\mathbb{D}$. Observe that $C(\beta) \neq \mathbb{D}$ as $\beta \notin \mathbb{F}$. Using our inductive hypothesis, we conclude that S is cyclic. □

For use in this section and the next, we list some properties of finite groups in which all Sylow subgroups are cyclic. More can be said: in fact, a finite group has cyclic Sylow subgroups if and only if it's a semidirect product of cyclic groups whose orders are relatively prime ([8, Theorem 5.16]).

Proposition 2. *Let G be a finite group in which every Sylow subgroup is cyclic and write l for the largest prime divisor of $|G|$. Then:*

(a) *a Sylow l-subgroup of G is normal;*

(b) *the group G is solvable;*

(c) *if G is nonabelian, it contains a normal abelian subgroup that is strictly larger than the centre.*

The proposition is a consequence of a famous result of Burnside ([8, Theorem 5.13]).

Burnside's Normal Complement Theorem. *Suppose a finite group G admits a Sylow subgroup S such that $N(S) = C(S)$. Then S has a normal complement in G. That is, there is a normal subgroup N of G such that $G = NS$ and $N \cap S = \{1\}$.*

[2]The author of the argument is better known today for activities outside mathematics.

The result is also known as Burnside's Transfer Theorem as it follows from properties of the transfer map—a natural homomorphism from a group G to the commutator group H/H' of a subgroup H of finite index . Isaacs' book [8] contains a cogent account of this map and several of its applications including Burnside's result.

Proof of Proposition 2. If G has prime-power order then (a) is obvious and (b) and (c) are well known. For the remainder of the proof, we assume that the order of G is divisible by at least two primes.

To establish part (a), we argue by induction on $|G|$. Write l_1 for the least prime divisor of G and let S_1 be a Sylow l_1-subgroup of G. First, we note that Burnside's Theorem implies that S_1 has a normal complement in G. To this end, observe that $S_1 \subset C(S_1)$, so that $[N(S_1) : C(S_1)]$ is not divisible by l_1. The action of $N(S_1)$ on S_1 by conjugation induces an embedding from $N(S_1)/C(S_1)$ into $\mathrm{Aut}(S_1)$. We have $|S_1| = l_1^e$ for some positive integer e. Since S_1 is cyclic, $\mathrm{Aut}(S_1)$ has order $\phi(l_1^e) = l_1^{e-1}(l_1 - 1)$. Thus $[N(S_1) : C(S_1)]$ divides $l_1 - 1$, and so $N(S_1) = C(S_1)$ (as l_1 is the least prime divisor of $|G|$). Hence, by Burnside's Theorem, $G = NS_1$ for a normal subgroup N of G such that $N \cap S_1 = \{1\}$.

Every Sylow subgroup of N is again cyclic. Our inductive hypothesis therefore implies that a Sylow l-subgroup S of N is normal in N, and thus is the unique Sylow l-subgroup of N. Now S is also a Sylow l-subgroup of G. Further, for $g \in G$, the group gSg^{-1} is a Sylow l-subgroup of N (by normality of N). Using uniqueness of S, we see that S is normal in G.

For part (b), we also use induction on $|G|$. With notation as in the proof of part (a), the quotient G/S inherits from G the property that each of its Sylow subgroups is cyclic. Thus, by our inductive hypothesis, G/S is solvable. As S is certainly solvable, it follows that G is solvable.

For part (c), suppose G is nonabelian and write Z for the centre of G. Then the nontrivial group G/Z has cyclic Sylow subgroups. Using part (a), we see that G/Z admits a nontrivial cyclic normal subgroup, say H/Z. Note that H is abelian (since H/Z is cyclic), normal in G and strictly contains Z. We've proved part (c). □

We now use Proposition 1 and Proposition 2 (c) to prove Wedderburn's theorem.

Proof. We assume that $\mathbb{D}$ is noncommutative and derive a contradiction.

By Proposition 1 and Proposition 2 (c), the group $\mathbb{D}^\times$ contains a normal abelian subgroup S that is strictly larger than $\mathbb{F}^\times$. We set $\mathbb{E} = \mathbb{F}(S)$, the subfield of $\mathbb{D}$ generated over $\mathbb{F}$ by S. Since S is normal in $\mathbb{D}^\times$, we see that $\mathbb{E}^\times$ is also normal in $\mathbb{D}^\times$.

As in the first proof, we write down two ways to complete the argument (see the last two paragraphs of [11] for yet another way).

Method 1. Since $N(\mathbb{E}^\times) = \mathbb{D}^\times$, Lemma 1 says that $C(\mathbb{E}^\times) = \mathbb{D}^\times$, and so $\mathbb{E}$ is contained in the centre $\mathbb{F}$ of $\mathbb{D}$—a contradiction.

Method 2. Recall that $|\mathbb{F}| = q$ and $|\mathbb{D}| = q^N$. As $\mathbb{E}$ strictly contains $\mathbb{F}$, we have $C(\mathbb{E}) \neq \mathbb{D}$. Thus

$$|\mathbb{E}| = q^m \text{ and } |C(\mathbb{E})| = q^n$$

for integers m and n with $m \leq n < N$. Note that n divides N: in fact, $N = ne$ where e is the dimension of $\mathbb{D}$ as a $C(\mathbb{E})$-vector space.

The action of $\mathbb{D}^\times$ on $\mathbb{E}$ by conjugation induces an embedding of groups

$$\mathbb{D}^\times / C(\mathbb{E})^\times \hookrightarrow \mathrm{Aut}(\mathbb{E}/\mathbb{F}).$$

In particular, $[\mathbb{D}^\times : C(\mathbb{E}^\times)]$ is at most $|\mathrm{Aut}(\mathbb{E}/\mathbb{F})| = m$, that is,

$$\frac{q^{ne} - 1}{q^n - 1} \leq m.$$

The left side is $q^{n(e-1)} + \cdots + q^n + 1$. Certainly $n < q^n$, and thus *a fortiori*

$$n \;<\; q^{n(e-1)} + \cdots + q^n + 1 \;\leq\; m,$$

in contradiction to $m \leq n$. □

5. Third Proof: Carter Subgroups

In 1961, R. W. Carter published a striking result [4].

Carter's Theorem. *Let G be a finite solvable group. Then G contains a nilpotent subgroup C such that $N(C) = C$. The subgroup C is unique up to conjugacy in G.*

There is an analogous statement in the theory of Lie algebras which prompted Carter's discovery. A Lie subalgebra of a Lie algebra that is nilpotent and coincides with its normalizer is called a Cartan subalgebra. Under suitable hypotheses, Cartan subalgebras exist and are unique up to a natural notion of conjugacy.[3]

Carter's Theorem, once formulated, is not difficult to prove. Indeed, it's an exercise without hints in Isaacs' text [8] (see p. 91) though the author does acknowledge that the problem may be "a bit harder than most of the problems in this book."

A nilpotent subgroup C of a group G such that $N(C) = C$ is now called a Carter subgroup of G. Non-solvable finite groups need not have Carter subgroups: for example, A_5, the smallest non-solvable group, has none. However, a Carter subgroup of an arbitrary finite group—when it exists—*is* unique up to conjugacy. This was put forward as a conjecture in 1976 and stood for around thirty years (see [15] for some of the history). It was finally proved by E. P. Vdovin as a culmination of a series of reductions and calculations based around the classification of finite simple groups—in particular, detailed properties of finite groups of Lie type [19].

Using Lemma 1 and the discussion in Section 4, we can quickly deduce the Little Theorem from Carter's result.

Proof. Let $\mathbb{D}$ be a noncommutative finite divsion ring. By Proposition 1 and Proposition 2 (b), the group $\mathbb{D}^\times$ is solvable. Let $\mathbb{E}$ be a maximal field in $\mathbb{D}$, so that $C(\mathbb{E}^\times) = \mathbb{E}^\times$. Then $N(\mathbb{E}^\times) = \mathbb{E}^\times$ by Lemma 1, and so $\mathbb{E}^\times$ is a Carter subgroup of $\mathbb{D}^\times$. Carter's theorem therefore says that the maximal subfields of $\mathbb{D}$ form a single conjugacy class (under the action of $\mathbb{D}^\times$). Since each element of $\mathbb{D}$ lies in a maximal field, it follows that $\mathbb{D}^\times$ is a union of conjugates of the proper subgroup $\mathbb{E}^\times$. Lemma 2, however, tells us that this is impossible. Thus a finite division ring is a field. □

Remark 4. We only used solvability of $\mathbb{D}^\times$ to ensure that the Carter subgroups of $\mathbb{D}^\times$ are conjugate. Can we instead appeal to conjugacy of Carter subgroups in a general finite group [19]? If so, this would yield an outrageous proof of the Little Theorem given that Vdovin's paper and its antecedents rest on the classification of finite simple groups. This approach, however, is not only wildly inefficient—it's circular: Wedderburn's theorem is a step, a tiny one, in the classification results that Vdovin uses.

6. Fourth Proof: Frobenius Groups

The final proof uses a uniqueness property of Frobenius groups. It relies on a powerful result of J. G. Thompson—we are indeed wielding a sledgehammer. For a more elementary approach to the Little Theorem via properties of Frobenius groups, see [6, 3].

Definition. A finite group G is a *Frobenius group* if it admits a nontrivial proper subgroup H such that

$$gHg^{-1} \cap H = \{1\} \text{ for all } g \in G \setminus H. \tag{4}$$

[3]Thanks to T. J. Laffey for confirming the influence on Carter of the theory of Cartan subalgebras.

Following [14], we say in this case that (G, H) is a *Frobenius pair*.

To draw out the definition, let's rephrase it in terms of the action of G by left multiplication on the space G/H of left cosets of H in G. The defining property (4) says exactly that each element of $G \setminus \{1\}$ fixes at most one point in G/H. Moreover, since H is nontrivial, some non-identity element of G fixes some point in G/H—each element of $H \setminus \{1\}$, for example, fixes the coset H. Conversely, suppose a finite group G acts transitively on a set X with $|X| > 1$ so that these properties hold, that is,

(a) no element of $G \setminus \{1\}$ fixes more than one element of X,

(b) some element of $G \setminus \{1\}$ fixes some element of X.

Let $x \in X$ and write H for the stabilizer in G of x. Then H satisfies (4) by (a), is nontrivial by (b), and is proper since $|X| > 1$. In all, (G, H) is a Frobenius pair.

In other words, a Frobenius group is a transitive permutation group on a finite set such that each nonidentity element fixes at most one point and some point *is* fixed by some nonidentity element (so that the action is not just the regular action).

There is a fascinating structure theory of Frobenius groups. For compelling accounts of part of this theory, see [8, Chap. 6] and [14, Chap. 6]. The first main result is due to Frobenius in 1901. It says that in any Frobenius pair (G, H) the elements of G that act without fixed points on G/H plus the identity element form a normal complement N to H in G, that is, $G = H \ltimes N$. The proof was an application and early triumph of the emerging theory of characters of finite groups, itself initiated by Frobenius in 1896. The group N is now called a *Frobenius kernel*.

Special properties of Frobenius groups gave rise to a conjecture that Frobenius kernels are always nilpotent. Thompson proved this conjecture in his celebrated 1959 doctoral thesis. The proof drew as much attention as the result: it hinged on a new, powerful criterion for an odd Sylow subgroup to admit a normal complement. Isaacs' book [8] includes a thorough discussion of Thompson's criterion (a slight modification of a simpler but still involved version that Thompson arrived at after his thesis) and its connection with Frobenius kernels (see Theorem 6.24 and Chapter 7). Another source, a more condensed one, is Feit's classic book [7] (see 22.2 and 25.10).

Nilpotence of Frobenius kernels leads to the following uniqueness property of Frobenius groups—the key to our final proof of the Little Theorem.

Uniqueness of Frobenius Structures. *Let (G, H_1) and (G, H_2) be Frobenius pairs in the same group G. Then H_1 and H_2 are conjugate in G.*

This is Exercise 6.6.6 in Serre's book [14] (see p. 86). Despite its label, the exercise is not in the least diabolical. In fact, Serre gives a generous hint that's effectively a solution.

We've sketched (more than) enough background to write down the fourth proof.

Proof. Let $\mathbb{D}$ be a noncommutative division ring of minimal order and let $\mathbb{E}$ be a maximal subfield of $\mathbb{D}$, so that $C(\mathbb{E}^\times) = \mathbb{E}^\times$. Then, by Lemma 1, $x\mathbb{E}^\times x^{-1} \neq \mathbb{E}^\times$ for $x \in \mathbb{D}^\times \setminus \mathbb{E}^\times$. Using minimality of $|\mathbb{D}|$, we noted at the start of Section 3 that a pair of maximal fields in $\mathbb{D}$ must intersect in the centre $\mathbb{F}$ (see equation (1) and the preceding discussion). In particular,

$$x\mathbb{E}^\times x^{-1} \cap \mathbb{E}^\times = \mathbb{F}^\times, \quad \text{for all } x \in \mathbb{D}^\times \setminus \mathbb{E}^\times.$$

Working in the quotient group $\mathbb{D}^\times/\mathbb{F}^\times$, it follows that

$$\begin{aligned} x\mathbb{F}^\times \cdot \left(\mathbb{E}^\times/\mathbb{F}^\times\right) \cdot (x\mathbb{F}^\times)^{-1} \cap \mathbb{E}^\times/\mathbb{F}^\times &= (x\mathbb{E}^\times x^{-1} \cap \mathbb{E}^\times)/\mathbb{F}^\times \\ &= \{1\}, \quad \text{for all } x \in \mathbb{D}^\times \setminus \mathbb{E}^\times. \end{aligned}$$

That is, $(\mathbb{D}^\times/\mathbb{F}^\times, \mathbb{E}^\times/\mathbb{F}^\times)$ is a Frobenius pair. Using uniqueness of Frobenius structures, we see that the maximal fields in $\mathbb{D}$ form a single conjugacy class.

To finish, we argue as in Method 1 or Method 2 of our first proof. □

7. Four Proofs of Zassenhaus's Lemma

It remains to prove Zassenhaus's lemma. As with the Little Theorem, we record four overlapping proofs. All share the same first step which is borrowed from [20]. It makes crucial use of surjectivity of the norm map for finite fields. From the theory of cyclic algebras, the Little Theorem is equivalent to this surjectivity property, so it's not a surprise that it lies at the base of the arguments below.

As noted earlier, this section owes much to the suggestions of a reviewer of a pre-Bulletin version of the paper. In particular, the fourth proof is a variant of an argument provided by this reviewer. The third proof is close to Zassenhaus's original argument.

For convenience, we recall the statement of the lemma.

Lemma 1. *If $\mathbb{E}$ is a subfield of a finite division ring $\mathbb{D}$, then the normalizer and centralizer of $\mathbb{E}^\times$ in $\mathbb{D}^\times$ coincide.*

Proof. Suppose $N(\mathbb{E}^\times) \neq C(\mathbb{E}^\times)$ for some subfield $\mathbb{E}$ of the finite division ring $\mathbb{D}$. Choose $\beta \in N(\mathbb{E}^\times)$ with $\beta \notin C(\mathbb{E}^\times)$.

Let l be the least positive integer such that $\beta^l \in C(\mathbb{E}^\times)$ and set $\mathbb{K} = \langle \mathbb{E}, \beta^l \rangle$, the subring of $\mathbb{D}$ generated by $\mathbb{E}$ and β^l. Note that $\mathbb{K}$ is a field since $\beta^l \in C(\mathbb{E}^\times)$. We write σ for the automorphism of $\mathbb{K}$ given by conjugation by β, that is,

$$\sigma(\lambda) = \beta\lambda\beta^{-1}, \ \ \lambda \in \mathbb{K}.$$

Observe that σ has order l.

Let $\mathbb{D}_1 = \langle \mathbb{K}, \beta \rangle$, the subring of $\mathbb{D}$ generated by $\mathbb{K}$ and β. For later use, note that

$$\beta^i\lambda = \sigma^i(\lambda)\beta^i, \quad i \in \mathbb{Z},\, \lambda \in \mathbb{K}. \tag{5}$$

We put $\mathbb{K}_1 = \mathbb{K}^\sigma$, the fixed field of σ. Thus the automorphism group of the extension $\mathbb{K}/\mathbb{K}_1$ is generated by σ and $[\mathbb{K} : \mathbb{K}_1] = l$. Let N denote the norm map from $\mathbb{K}$ to $\mathbb{K}_1$, so that

$$N(\alpha) = \alpha\sigma(\alpha)\sigma^2(\alpha)\cdots\sigma^{l-1}(\alpha), \quad \alpha \in \mathbb{K}.$$

Now, for $\alpha \in \mathbb{K}$,

$$\begin{aligned}(\alpha\beta)^l &= \alpha\beta\cdot\alpha\beta\cdots\alpha\beta \ \ (l \text{ terms})\\ &= \alpha\cdot\beta\alpha\beta^{-1}\cdot\beta^2\alpha\beta^{-2}\cdots\beta^{l-1}\alpha\beta^{-(l-1)}\cdot\beta^l\\ &= \alpha\sigma(\alpha)\sigma^2(\alpha)\cdots\sigma^{l-1}(\alpha)\beta^l\\ &= N(\alpha)\beta^l.\end{aligned}$$

The element $\beta^l \in \mathbb{K}$ is visibly fixed under conjugation by β, that is, $\beta^l \in \mathbb{K}_1$. Since norm maps are surjective for finite fields, we can choose $\alpha \in \mathbb{K}$ such that $N(\alpha) = \beta^{-l}$, equivalently $(\alpha\beta)^l = 1$. Thus we can and *do* adjust the element β so that $\beta^l = 1$. We still have $\mathbb{D}_1 = \langle \mathbb{K}, \beta \rangle$ and the automorphism σ of $\mathbb{K}$ is still given by conjugation by β.

We rewrite $\beta^l = 1$ as $(1-\beta)(1+\beta+\cdots+\beta^{l-1}) = 0$. Since $\beta \neq 1$, we have

$$1 + \beta + \cdots + \beta^{l-1} = 0, \tag{6}$$

and so

$$\lambda + \beta\lambda + \cdots + \beta^{l-1}\lambda = 0, \quad \text{for all } \lambda \in \mathbb{K}.$$

Equivalently, via (5),

$$\lambda + \sigma(\lambda)\beta + \cdots + \sigma^{l-1}(\lambda)\beta^{l-1} = 0, \quad \text{for all } \lambda \in \mathbb{K}. \tag{7}$$

From here, we write down four ways to complete the proof.

Method 1. Fix $\zeta \in \mathbb{K}$ such that $\mathbb{K} = \mathbb{K}_1(\zeta)$. Substituting $\zeta^0 = 1, \zeta, \ldots, \zeta^{l-1}$ in (7) gives

$$\zeta^i + \sigma(\zeta)^i\beta + \cdots + \sigma^{l-1}(\zeta)^i\beta^{l-1} = 0, \quad \text{for } i = 0, 1, \ldots, l-1. \tag{8}$$

We can rewrite these l equations as a single matrix equation

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \zeta & \sigma(\zeta) & \cdots & \sigma^{l-1}(\zeta) \\ \vdots & \vdots & \ddots & \vdots \\ \zeta^{l-1} & \sigma(\zeta)^{l-1} & \cdots & \sigma^{l-1}(\zeta)^{l-1} \end{bmatrix} \begin{bmatrix} 1 \\ \beta \\ \vdots \\ \beta^{l-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{9}$$

Thus the square matrix, say M, is singular. On the other hand, it's of Vandermonde form, so

$$\det M = \prod_{0 \le i < j < l} \big(\sigma^j(\zeta) - \sigma^i(\zeta)\big)$$

is nonzero. This contradiction proves the lemma.

Method 2. We derive a contradiction from (7) by a slightly different matrix argument.

Write $\mathrm{tr} : \mathbb{K} \to \mathbb{K}_1$ for the trace map of the extension $\mathbb{K}/\mathbb{K}_1$, so that

$$\mathrm{tr}(\lambda) = \lambda + \sigma(\lambda) + \cdots + \sigma^{l-1}(\lambda), \quad \lambda \in \mathbb{K}.$$

Since $\mathbb{K}/\mathbb{K}_1$ is separable, the symmetric bilinear form

$$(\lambda, \mu) \mapsto \mathrm{tr}(\lambda\mu) : \mathbb{K} \times \mathbb{K} \to \mathbb{K}_1$$

is nondegenerate. Fixing a $\mathbb{K}_1$-basis $\lambda_1, \ldots, \lambda_l$ of $\mathbb{K}$, the matrix $\big[\mathrm{tr}(\lambda_i\lambda_j)\big]$ is therefore invertible.

Substituting the basis elements $\lambda_1, \ldots, \lambda_l$ in (7), we have

$$\lambda_i + \sigma(\lambda_i)\beta + \cdots + \sigma^{l-1}(\lambda_i)\beta^{l-1} = 0, \quad \text{for } i = 1, \ldots, l. \tag{10}$$

As in Method 1, we gather these l equations into a single matrix equation

$$\begin{bmatrix} \lambda_1 & \sigma(\lambda_1) & \cdots & \sigma^{l-1}(\lambda_1) \\ \lambda_2 & \sigma(\lambda_2) & \cdots & \sigma^{l-1}(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_l & \sigma(\lambda_l) & \cdots & \sigma^{l-1}(\lambda_l) \end{bmatrix} \begin{bmatrix} 1 \\ \beta \\ \vdots \\ \beta^{l-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{11}$$

Writing M for the given square matrix and $M^\top$ for its transpose, the product $MM^\top$ has ij entry

$$\lambda_i\lambda_j + \sigma(\lambda_i\lambda_j) + \cdots + \sigma^{l-1}(\lambda_i\lambda_j) = \mathrm{tr}(\lambda_i\lambda_j).$$

Thus $MM^\top$ is invertible, so M is invertible, in contradiction to (11).

Method 3. Using (5) and that $\beta^l \in \mathbb{K}$, we see that

$$\mathbb{D}_1 = \langle \mathbb{K}, \beta \rangle = \sum_{i=0}^{l-1} \mathbb{K}\beta^i. \tag{12}$$

We claim that the sum is direct. Note that our claim means that (6) cannot hold and thus establishes the lemma.

The claim follows from a powerful technique that goes back to Dedekind, often called "linear independence of characters." In detail, suppose we have a dependence relation

$$\lambda_1\beta^{i_1} + \cdots + \lambda_r\beta^{i_r} = 0 \tag{13}$$

for $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ and $0 \le i_1 < \cdots < i_r < l$. In search of a contradiction, we assume that r is minimal, so that each $\lambda_i \neq 0$.

Multiplying (13) on the right by $\gamma \in \mathbb{K}$ and rewriting via (5), we have

$$\lambda_1\sigma^{i_1}(\gamma)\beta^{i_1} + \cdots + \lambda_r\sigma^{i_r}(\gamma)\beta^{i_r} = 0, \quad \gamma \in \mathbb{K}. \tag{14}$$

Choose $\gamma_1 \in \mathbb{K}$ such that $\sigma^{i_1}(\gamma_1) \neq \sigma^{i_2}(\gamma_1)$. Replacing γ by $\gamma_1\gamma$ in (14) gives

$$\lambda_1\sigma^{i_1}(\gamma_1)\sigma^{i_1}(\gamma)\beta^{i_1} + \lambda_2\sigma^{i_2}(\gamma_1)\sigma^{i_2}(\gamma)\beta^{i_2} + \cdots + \lambda_r\sigma^{i_r}(\gamma_1)\sigma^{i_r}(\gamma)\beta^{i_r} = 0.$$

On the other hand, multiplying (14) on the left by $\sigma^{i_1}(\gamma_1)$ gives

$$\lambda_1\sigma^{i_1}(\gamma_1)\sigma^{i_1}(\gamma)\beta^{i_1} + \lambda_2\sigma^{i_1}(\gamma_1)\sigma^{i_2}(\gamma)\beta^{i_2} + \cdots + \lambda_r\sigma^{i_1}(\gamma_1)\sigma^{i_r}(\gamma)\beta^{i_r} = 0.$$

Subtracting the last equation from the previous one, we see that

$$\lambda_2\left(\sigma^{i_2}(\gamma_1) - \sigma^{i_1}(\gamma_1)\right)\sigma^{i_2}(\gamma)\beta^{i_2} + \cdots + \lambda_r\left(\sigma^{i_r}(\gamma_1) - \sigma^{i_1}(\gamma_1)\right)\sigma^{i_r}(\gamma)\beta^{i_r} = 0.$$

By our choice of γ_1, the coefficient of β^{i_2} is nonzero (for γ nonzero) which contradicts minimality of r in (13). Therefore (12) is a direct sum and we've proved the lemma once more.

Method 4. Consider $\mathbb{K}$ as a $\mathbb{K}_1$-vector space. The field $\mathbb{K}$ embeds in $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ via $m : \mathbb{K} \to \mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ where $m(\lambda)(\mu) = \lambda\mu$ (for $\lambda, \mu \in \mathbb{K}$). We identify $\mathbb{K}$ with its image under m. That is, for $\lambda \in \mathbb{K}$, we simply write λ for the map $m(\lambda) \in \mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$. We also have $\sigma \in \mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$. In parallel to (5), these various elements of $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ satisfy

$$\sigma^i\lambda = \sigma^i(\lambda)\sigma^i, \quad i \in \mathbb{Z},\ \lambda \in \mathbb{K}, \tag{15}$$

Now, by Dedekind's lemma (linear independence of characters), the sum

$$\sum_{i=0}^{l-1} \mathbb{K}\sigma^i \subseteq \mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$$

is direct, and hence has dimension l^2 over $\mathbb{K}_1$. The containment is therefore an equality. That is, each element of $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ admits a unique expression as $\sum_{i=0}^{l-1}\lambda_i\sigma^i$ with each $\lambda_i \in \mathbb{K}$. Using the same notation, we see that there is a well-defined map of $\mathbb{K}$-vector spaces

$$\sum_{i=0}^{l-1}\lambda_i\sigma^i \longmapsto \sum_{i=0}^{l-1}\lambda_i\beta^i : \mathrm{End}_{\mathbb{K}_1}(\mathbb{K}) \longrightarrow \mathbb{D}_1. \tag{16}$$

We claim that the map is a ring homomorphism. Note that the lemma follows. Indeed, given the claim, the kernel of (16) is a two-sided ideal in the simple ring $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ and hence is trivial. This means that (16) is injective and so the division ring $\mathbb{D}_1$ contains a copy of the matrix ring $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$—a contradiction. Alternatively, the map is surjective (its image, for example, contains $\mathbb{K}$ and β). Using simplicity of $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ again, the rings $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ and $\mathbb{D}_1$ must be isomorphic which is absurd.

To establish the claim, we appeal to the following principle whose proof is immediate.

(α) *Let R and S be rings and let $\phi : R \to S$ be a homomorphism of abelian groups. Suppose R is generated as an abelian group by a subset Γ and that*

$$\phi(\gamma_1\gamma_2) = \phi(\gamma_1)\phi(\gamma_2)$$

for all $\gamma_1, \gamma_2 \in \Gamma$. Then ϕ is a homomorphism of rings.

From our discussion, $\mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ is generated as an abelian group by the elements $\lambda\sigma^i$ for $\lambda \in \mathbb{K}$ and $i \in \mathbb{Z}$. Writing ϕ for the map in (16) and using (α), we see that we only have to check one family of relations:

$$\phi(\lambda\sigma^i\mu\sigma^j) = \phi(\lambda\sigma^i)\phi(\mu\sigma^j), \qquad \lambda, \mu \in \mathbb{K},\ i, j \in \mathbb{Z}.$$

By (15), $\lambda\sigma^i\mu\sigma^j = \lambda\sigma^i(\mu)\sigma^{i+j}$, and so the left side is $\lambda\sigma^i(\mu)\beta^{i+j}$. By (5), the right side is

$$\lambda\beta^i\mu\beta^j = \lambda\sigma^i(\mu)\beta^{i+j}.$$

Thus (16) is a ring homomorphism and the (fourth) proof is complete. □

Remark 5. Dedekind's lemma (linear independence of characters) is at the heart of Zassenhaus's lemma. Its central role in Methods 3 and 4 is evident. For Methods 1 and 2, the key is that the square matrices in equations (9) and (11) are invertible—and invertibility of these matrices is a quick consequence of Dedekind's lemma (a preferable approach perhaps to the more computational one used above). To check the implication, write C_i for the ith column of either matrix and suppose there exist $\gamma_1, \ldots, \gamma_l \in \mathbb{K}$ such that

$$\gamma_1 C_1 + \cdots + \gamma_l C_l = 0,$$

the zero column vector. In each case, the relation says that $\sum_{i=0}^{l-1} \gamma_i \sigma^i \in \mathrm{End}_{\mathbb{K}_1}(\mathbb{K})$ vanishes on a $\mathbb{K}_1$-basis of $\mathbb{K}$. Thus $\sum_{i=0}^{l-1} \gamma_i \sigma^i = 0$, so that $\gamma_i = 0$ for all i, and the matrices are indeed invertible via Dedekind's lemma. Noether's motto, quoted in [16], was "it is already all in Dedekind."

References

[1] M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, 6th ed. Springer, 2018.
[2] E. Artin, *The influence of J. H. M. Wedderburn on the development of modern algebra*. Bull. Amer. Math. Soc. 56 (1950), 65-72.
[3] R. P. Burn and D. M. Madaram, *Frobenius groups and Wedderburn's theorem*. Amer. Math. Monthly. 77(9) (1970), 983-984.
[4] R. W. Carter, *Nilpotent self-normalizing subgroups of soluble groups*. Math. Zeit. 75 (1961), 136-139.
[5] R. K. Dennis and B. Farb, *Noncommutative Algebra*. Graduate Texts in Math., 144, Springer-Verlag, 1993.
[6] S. Ebey and K. Sitaram, *Frobenius groups and Wedderburn's theorem*. Amer. Math. Monthly. 76(5) (1969), 526-528.
[7] W. Feit, *Characters of Finite Groups*. W. A. Benjamin, 1967.
[8] I. M. Isaacs, *Finite Group Theory*. Graduate Studies in Math., 92, Amer. Math. Soc., 2008.
[9] I. N. Herstein, *Topics in Algebra*, 2nd ed. John Wiley & Sons, 1975.
[10] N. Jacobson, *Basic Algebra II*, 2nd ed. W. H. Freeman and Co., 1989.
[11] T. J. Kaczynski, *Another proof of Wedderburn's theorem*. Amer. Math. Monthly. 71(6) (1964), 652-653.
[12] S. Mac Lane, *Mathematics at Göttingen under the Nazis*. Notices Amer. Math. Soc. 42(10) (1995), 1134-1138.
[13] C. Reid, *Hilbert*. Reprint of 1970 original. Copernicus, 1996.
[14] J.-P. Serre, *Finite Groups: An Introduction*. International Press, 2016.
[15] M. Chiara Tamburini, *The conjugacy conjecture for Carter subgroups*. Milan J. Math. 75 (2007), 357-377.
[16] B. L. van der Waerden, *On the sources of my book* Moderne Algebra. Hist. Math. 2 (1975), 31-40.
[17] J. H. M. Wedderburn, *A theorem on finite algebras*. Trans. Amer. Math. Soc. 6 (1905), no. 3, 349-352.
[18] J. H. M. Wedderburn, *On hypercomplex numbers*. Proc. London Math. Soc. (2) 6 (1908), 77-118.
[19] E. P. Vdovin, *Carter subgroups of finite groups*. Siberian Adv. Math. 19 (2009), no. 1, 27-74.
[20] H. J. Zassenhaus, *A group-theoretic proof of a theorem of Maclagan-Wedderburn*. Proc. Glascow Math. Assoc. 1 (1952), 53-63.

Alan Roche studied at University College Dublin and the University of Chicago, and has worked at the University of Oklahoma since 2001. His principal mathematical interests are in number theory and representation theory.

Department of Mathematics, University of Oklahoma, Norman, OK 73019, USA.
E-mail address: aroche@ou.edu