CLASSIFYING THE FINITE SIMPLE GROUPS

BY DANIEL GORENSTEIN[1]

CONTENTS

INTRODUCTION

CHAPTER IV: Outline of a Unitary Proof

INTRODUCTION

Although the title of this article is almost the same as that of an earlier paper in this journal, *The classification of the finite simple groups* [75], with the added similarity that each is divided into four chapters, nevertheless there is almost no overlap between the two, beyond the fact that both concern the recently completed determination of the finite simple groups.

My earlier paper focused on the past: its primary aim was to describe the finite simple groups (Chapters II and III) and also to present a picture of the principal techniques underlying the classification theorem (Chapter IV), with a brief global overview of the proof itself (Chapter I), added to put the entire discussion in perspective.

In contrast, the present paper looks to the future: its concern is, on the one hand, with the significance of the classification theorem both to future research in finite group theory and to its range of applicability in other areas of mathematics (topics discussed in Chapter I) and, on the other hand, to the fundamental task of constructing a shorter and more readily accessible "second generation" classification proof (this is the content of the remaining three chapters).

In preparation for a discussion of the meaning of a "satisfactory classification proof," in Chapter II we first illustrate the partially haphazard and partially inexorable evolution of the existing proof, developed as it was over a thirty-year period without benefit of a predetermined plan. Under such circumstances, it was inevitable that the final global design included considerable inefficiencies and, with hindsight, some avoidable duplications of effort. In Chapter III we present a critique of the present proof and discuss alternative strategies for constructing a satisfactory proof.

My own close examination of the present proof over the past several years, largely with Richard Lyons, and more recently with Ronald Solomon as well, has led to a particular global strategy which, utilizing only presently known

group-theoretic techniques, holds out the prospect of achieving as much as a five-fold reduction in length, with a commensurate over-all conceptual simplification. Chapter IV is devoted to a detailed outline of this proposed plan, whose implementation is well under way as of this writing.

Finally, for completeness, we have included in Chapter I a brief discussion of the known simple groups—in particular, the sporadic groups—which are, of course, also (more fully) described in [75]. However, concerning the sporadic groups, we have focused here on the distinction between the group-theoretic context in which each first arose and the specific method by which each was subsequently constructed, a point not stressed in the earlier article.

Although the present paper is self-contained, we should add that the nonexpert will certainly find the material in [75] helpful as background for understanding this discussion of the finite simple groups.

CHAPTER I. STATEMENT AND SIGNIFICANCE OF THE THEOREM

The classification theorem of the finite simple groups asserts that every finite simple group is on a completely specified list. Clearly, before we can discuss its proof, we must at least briefly describe the groups composing that list.

Since this classification theorem is generally regarded as a milestone of twentieth-century mathematics, I would also like to give some feeling for its profound impact on finite group theory and at the same time indicate the strong prospects, already partially realized, of its significant application to widely diverse areas of mathematics. Finally, in the last section of this part, I shall take a more philosophical perspective, discussing its significance within the broader context of classification theorems, in general, and shall also attempt to explain the inherent complexity of the theorem.

Our notation will be standard. However, terms unfamiliar to the nonexpert will be defined here. In particular, for any group X and subset or subgroup Y of X, $C_X(Y) = $ *centralizer* of Y in $X = \{ x \in X \mid xy = yx$ for all $y \in Y \}$, $Z(X) = $ *center* of $X = C_X(X)$, and $N_X(Y) = $ *normalizer* of Y in $X = \{ x \in X \mid x^{-1}Yx = Y \}$. Additional terminology will be introduced as we go along.

1. Statement of the theorem. We first state the classification theorem in a general form and shall then amplify its statement.

CLASSIFICATION THEOREM. *Every finite simple group is isomorphic to one of the following*:

(1) *A cyclic group of prime order*;

(2) *An alternating group*;

(3) *A member of one of sixteen infinite families of groups of Lie type*; *or*

(4) *One of twenty-six sporadic groups not isomorphic to any of the above groups*.

The group Z_p of prime order p is, of course, commutative. For brevity, the term "simple group" usually refers to a *nonabelian* simple group.

The alternating group A_n is the subgroup of even permutations of the symmetric group Σ_n on n letters. A_n is simple for all $n \geqslant 5$.

It is the groups of Lie type and the sporadic groups that require fuller descriptions.

A. *The groups of Lie type.* The sixteen families of simple groups of Lie type are finite analogues of (complex) Lie groups. They divide into three general types.

1. *The Chevalley groups.* There are four infinite families of Lie groups: the *linear* groups A_n, the *symplectic* groups C_n, and the *orthogonal* groups B_n, D_n, plus five *exceptional* Lie groups G_2, F_4, E_6, E_7, and E_8, the subscripts referring to their *Lie rank*—i.e., to the number of fundamental reflections generating their Weyl groups. [It is standard to use the notation A_n for both the alternating and linear groups. The context will make clear which group is intended.]

Each of these groups has representations by complex matrices. The finite analogues are obtained by using matrices over a finite field $GF(q)$ with q elements.

For example, the *general linear group* $GL_m(q)$ is the group of all nonsingular $m \times m$ matrices with coefficients in $GF(q)$. Its normal subgroup $SL_m(q)$ of matrices of determinant 1 is the *special linear group*, and the factor group $PSL_m(q) = SL_m(q)/$(group of scalar matrices) is the *projective special* linear group.

$PSL_m(q)$ is the finite analogue of the Lie group A_n for $n = m - 1$.

$PSL_m(q)$ is simple for all $m \geqslant 2$ except for $m = 2$ and $q \leqslant 3$ (in which cases it is solvable [just as the alternating groups of degree $\leqslant 4$ are solvable]).

It was Chevalley [42] who proved the existence of finite analogues of the simple Lie groups over any field—in particular over $GF(q)$, q any prime power—at the same time establishing their simplicity and showing that they have internal structures very similar to those of the corresponding Lie groups. Thus we have *nine* families of finite simple

CHEVALLEY GROUPS

$$A_n(q),\ B_n(q),\ C_n(q),\ D_n(q),\ G_2(q),\ F_4(q),\ E_6(q),\ E_7(q),\ \text{and}\ E_8(q).$$

2. *The Steinberg groups.* The Lie groups possess so-called real forms and those Lie groups with a *symmetric Dynkin diagram* have extra such forms: namely,

$$A_n \qquad \circ\!\!-\!\!\circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ$$

$$D_n \qquad \circ\!\!-\!\!\circ\!\!-\!\!\circ \cdots \circ\!\!-\!\!\circ\!\!\Big\langle{\circ \atop \circ}$$

and

$$E_6 \qquad \circ\!\!-\!\!\circ\!\!-\!\!\overset{\circ}{\circ}\!\!-\!\!\circ\!\!-\!\!\circ$$

which have symmetries of period 2. [The diagram of D_4 also has a symmetry of period 3, but as the complex numbers do not have an automorphism of period 3 over the reals, there is no corresponding real form.]

The *unitary* groups are such real forms of the linear groups. A matrix $X \in GL_m(\mathbf{C})$ is *unitary* if

$$(\overline{X}^t)^{-1} = X,$$

where $^-$ denotes complex conjugation and t denotes transpose. The set of unitary matrices form a subgroup $GU_m(\mathbf{C})$, the *general unitary* group, of $GL_m(\mathbf{C})$. One can then consider the corresponding groups $SU_m(\mathbf{C})$, and $PSU_m(\mathbf{C}) = SU_m(\mathbf{C})/\text{scalars}$, and one has

$$PSU_m(\mathbf{C}) = \text{real form of } A_n \text{ for } n = m - 1.$$

Note also that the map α: $X \to (\overline{X}^t)^{-1}$ is an *automorphism* of $GL_m(\mathbf{C})$ of period 2, and the group $GU_m(\mathbf{C})$ is precisely the subgroup of matrices fixed by α.

The unitary groups have finite analogues when $q = r^2$, for then the underlying field $GF(q)$ possesses an automorphism of period 2 (again denoted by $^-$), in which each element is raised to the rth power. Using this automorphism in place of complex conjugation, one obtains the finite unitary groups in the same way as the complex unitary groups (one usually writes $GU_m(r)$, etc., instead of $GU_m(q)$).

It was Steinberg [126] who showed that Chevalley's arguments carry over to produce finite analogues of real forms of A_n, D_n, E_6 whenever q is a square, as groups of fixed points of appropriate automorphisms of period 2 of the corresponding Chevalley groups. Moreover, when $q = r^3$, the cubing map gives an automorphism of $GF(q)$ of period 3, which can be used in place of complex conjugation to produce a second "real form" of $D_4(q)$, known as *triality* D_4. In this way we obtain *four* families of simple groups which are designated as follows:

<div align="center">

STEINBERG GROUPS

$${}^2A_n(q), \quad {}^2D_n(q), \quad {}^3D_4(q), \quad \text{and} \quad {}^2E_6(q).$$

</div>

Here the exponent indicates the order of the given symmetry, and the corresponding Chevalley groups are defined over $GF(q^2)$, $GF(q^2)$, $GF(q^3)$, $GF(q^2)$, respectively.

The linear, symplectic, orthogonal groups, and unitary groups are also known as the *classical* groups. We shall use the notation $PSL_m(q)$, $PSp_4(q)$, $PSO_m^\varepsilon(q)$, and $PSU_m(q)$ for them; here $\varepsilon = \pm 1$ and according as m is even or odd and $\varepsilon = +1$ or -1, $PSO_m^\varepsilon(q)$ corresponds to B_n, D_n, or ${}^2D_n$ for suitable n.

3. *The Suzuki and Ree groups.* When $q = 2^a$, 3^a, and 2^a, respectively, the groups $B_2(q)$, $G_2(q)$, and $F_4(q)$ possess an automorphism group twice as large as that predicted by the general Lie theory. Moreover, when a is odd, there is, in fact, an extra automorphism of period 2. Taking fixed points with respect to these involutory automorphisms yields three further families of simple groups for which there exist no complex analogues. Suzuki [132] had constructed the first of these families (usually denoted by $Sz(q)$) without being aware of their connection with the Lie theory. However, later Ree [112, 113] realized that connection and went on to apply the Steinberg method to the G_2 and F_4 cases.

Thus we have the *three* additional families of

SUZUKI AND REE GROUPS

$$Sz(q) = {}^2B_2(q), \quad q = 2^a; \qquad {}^2G_2(q), \quad q = 3^a;$$

$$ {}^2F_4(q), \quad q = 2^a; \qquad a \text{ odd}.$$

Together the Chevalley, Steinberg, and Suzuki-Ree groups constitute the (finite) groups of *Lie type*. They are all simple except in very low Lie rank and over small fields. Moreover, all isomorphisms among them, as well as with alternating groups, have been determined.

It should be mentioned that the three kinds of groups of Lie type all arise from a uniform construction, elucidated by Steinberg [129]: namely, if G denotes the Chevalley analogue over an algebraically closed field $\bar{F}_p$ of characteristic p of a complex simple Lie group, then G is, in fact a *linear algebraic group*. Moreover, if σ is a surjective endomorphism of G (as an algebraic group) whose fixed points $C_G(\sigma)$ form a finite group (e.g., σ: $x \to x^q$, $x \in \bar{F}_p$, $q = p^n$), then every nonabelian composition factor of $C_G(\sigma)$ is a (finite) group of Lie type, and all groups of Lie type arise this way for some σ.

B. *The sporadic groups.* The sporadic groups were discovered in the course of investigating several quite independent problems. As a result, even a minimal description of their origins requires a number of necessary terms, which we prefer to introduce at the same time.

Let X be an arbitrary finite group.

An element of X of order 2 is called an *involution*.

If X is a subgroup of Σ_n for some n, X is called a *permutation group* (of *degree n*).

We regard Σ_n as acting on the set $\Omega = \{1, 2, \ldots, n\}$.

X is *k-fold* transitive on Ω if any ordered k-tuple of distinct points of Ω can be transformed into any other by an element of X. One writes *transitive*, *doubly transitive*, etc. for 1-fold transitive, 2-fold transitive, etc.

The *stabilizer* X_a in X of the point $a \in \Omega$ is the subgroup of all elements of X leaving a fixed. If X is transitive, then X_a is determined up to conjugacy (and hence up to isomorphism) by a.

If X is transitive on Ω, then the *permutation rank r* of X is the number of distinct orbits of X_a in its action on Ω (i.e., the number of transitive constituents of X_a on Ω). r is independent of the choice of $a \in \Omega$. We have $r \geqslant 2$ (assuming $n \geqslant 2$) with $r = 2$ if and only if X is doubly transitive.

X is *primitive* on Ω if X is transitive on Ω and Ω cannot be decomposed as the union of $k \geqslant 2$ disjoint subsets $\Omega_1, \Omega_2, \ldots, \Omega_k$ that are transformed into each other by X. X is primitive on Ω if and only if X is transitive on Ω and X_a is a *maximal* subgroup of X for $a \in \Omega$.

The set $\mathscr{D}$ of all transpositions (ij), $i \neq j$, $1 \leqslant i, j \leqslant n$ has the following properties:

(1) The elements of $\mathscr{D}$ form a conjugacy class of involutions of Σ_n;
(2) Σ_n is generated by the elements of $\mathscr{D}$; and
(3) If $x, y \in \mathscr{D}$, then the order $|xy|$ of xy is 1, 2, or 3.

Motivated by these properties, Fischer made the following definition [56].

The abstract group X is called a 3-*transposition* group if X is generated by a (union) of conjugacy classes $\mathscr{D}$ of involutions such that the product of any two elements of $\mathscr{D}$ has order 1, 2, and 3.

Extending the definition, X is similarly called a $\{3,4\}$-*transposition* group if the product of any two involutions of the generating classes $\mathscr{D}$ has order 1, 2, 3, or 4.

Finally, a *lattice* Λ is a free abelian group of finite rank n together with a real-valued symmetric nonsingular bilinear form $(\,,\,)$. If $(\,,\,)$ is positive definite, then Λ can be embedded in $\mathbf{R}^n$. Λ is *integral* if (x, y) is an integer for all $x,\ y \in \Lambda$, and Λ is *even* if, in addition, (x, x) is an even integer for all $x \in \Lambda$. Moreover, Λ is *unimodular* if for some (any) basis $w_1,\ w_2, \ldots, w_n$ of Λ, the determinant of the matrix $((w_i, w_j))$ is ± 1.

The *Leech* lattice [98, 99] is a particular 24-dimensional lattice, which we shall not describe explicitly. We note, however, that Conway [43] has shown that it is the unique 24-dimensional positive definite unimodular even lattice in which no vector has square norm 2 (i.e., $(x, x) \neq 2$ for all $x \in \Lambda$).

$M_{11}, M_{12}, M_{22}, M_{23}, M_{24}$	The Mathieu groups
J_1, J_2, J_3, J_4	Janko's groups
.1, .2, .3	Conway's groups
M(22), M(23), M(24)'	The Fischer groups [M(24)' is the derived group of index 2 in M(24).]
HS	The Higman-Sims group
Mc	McLaughlin's group
Suz	Suzuki's (sporadic) group
Ru	The Rudvalis group
He	Held's group
Ly	Lyons' group
ON	O'Nan's group
F_5	Harada's group
F_3	Thompson's group
F_2	Fischer's "baby monster" group
F_1	The Fischer-Griess "friendly giant" or "monster" group

TABLE 1. The sporadic groups.

Likewise I shall not give the exact definition of the *Griess-Norton* algebra—a certain 196,883-dimensional commutative, nonassociative complex algebra, equipped with an associative form, which underlies the construction of Griess' "friendly giant," otherwise known as "the monster" [85].

Now we can give a schematic picture of the group-theoretic origins of the twenty-six sporadic groups and the terms in which they were initially constructed (Table 2). It should be pointed out that many were constructed only after a lengthy analysis of the particular group-theoretic context in which they arose. In fact, in some instances, the individual who found the initial *evidence* for the group was not the same one(s) who carried out the actual construction (primarily in those cases requiring computer calculations). Therefore, where appropriate, we list separately the original context and the basis for construction, likewise indicating by the term "(computer)" its dependence on computer calculations.

For brevity, we make no systematic attempt to give full attributions. In general, our notation refers to the individual(s) who found the initial evidence for the group, except for those cases in which other notation has become standard.

[Although F_1 itself is not one of the Fischer transposition groups and the initial evidence for its existence arose from a centralizer-of-involution problem, we prefer to include it with the transposition groups because of its close dependency on F_2, whose properties were crucial for its discovery. Note also that it is the commutator subgroup $M(24)'$ of index 2 in $M(24)$ that is a sporadic *simple* group.]

Griess' remarkable construction of F_1 was carried out entirely by hand, depending only upon the prior existence of .1 (which occurs as the homomorphic image of the centralizer of one of the involutions of F_1). In turn, construction of .1 requires only prior existence of M_{24}. Hence, as a corollary of his existence theorem, Griess immediately obtained a *noncomputer* proof of the existence of the following *seven* additional sporadic groups, each of which is involved in some fashion in F_1!

$$He, \quad M(22), \quad M(23), \quad M(24)', \quad F_5, \quad F_3, \quad F_2.$$

[J_2, HS, and Mc are also in F_1, but occur as subgroups of .1.]

Since the completion of the classification, Ru and J_3 have also been given noncomputer constructions, so that at the present time only the three groups

$$Ly, \quad ON, \quad J_4$$

depend on computer calculations for their existence.

2. Applications of the theorem. Some have suggested that the classification of the finite simple groups is an essentially isolated result that will have only a limited mathematical impact. The problem was there, it was interesting, so we solved it—much the same as the four-color problem. I believe this to be a serious misreading of the situation, for the evidence to date indicates, to the contrary, that the classification theorem is destined to have profound mathematical impact. Indeed, it has already found significant application in such

diverse fields as number theory, automorphic functions, finite geometry, model theory, algorithms, and coding theory as well as to finite group theory itself, and as the result becomes more widely understood and its usefulness appreciated, the number and range of such applications can certainly be expected to increase.

Origin	Construction
1. **Multiply transitive groups**	
M_{12}, M_{24}	5-fold transitive groups
M_{11}, M_{22}, M_{23}	point stabilizers
2. **Centralizers of involutions**	
J_1	Subgroup of $GL_7(11)$
J_2	Rank 3 primitive permutation group
J_3, Ly, He, ON	Rank > 3 primitive permutation groups (computer)
J_4	Subgroup of $GL_{112}(2)$ (computer)
F_5	Subgroup of $GL_{133}(C)$ (computer)
F_3	Subgroup of $E_8(3)$ (computer)
3. **Rank 3 primitive permutation groups**	
HS, Mc, Suz	Rank 3 primitive permutation groups
Ru	Subgroup of $GL_{28}(C)$ (computer)
4. **Automorphism groups of integral lattices**	
.1	Automorphisms of Leech lattice
.2, .3	Stabilizers of suitable Leech lattice vectors
5. **Transposition groups**	
M(22), M(23), M(24) (3-transposition groups)	Rank 3 primitive permutation groups
F_2 ({3,4}-transposition group)	Rank > 3 primitive permutation group (computer)
F_1 ("double cover" of F_2 as centralizer of involution)	Automorphism group of Griess-Norton algebra

TABLE 2

Here I shall limit myself to some selected applications of a general nature that will indicate the scope of the theorem.

A. *Finite group theory.* Many general questions about finite groups can be reduced to questions concerning simple groups, so that one would certainly expect the classification theorem to have important consequences for finite group theory. Here are five such applications.

1. *The Schreier conjecture.* The automorphism group of every simple group has been calculated, so that as a direct consequence of the classification theorem one obtains a proof of the celebrated Schreier conjecture:

THEOREM. *Every simple group has a solvable outer automorphism group.*

2. *Multiply transitive permutation groups.* Likewise one can easily determine which simple groups possess highly transitive permutation representations, which yields the following further consequence of the theorem:

THEOREM. *Every quadruply transitive permutation group is either an alternating, symmetric, or Mathieu group. In particular, the alternating and symmetric groups are the only 6-fold transitive permutation groups.*

3. *The Sims conjecture.* Some years ago Sims proved the following result [119]:

THEOREM. *If G is a primitive permutation group on a set Ω and if a one-point stabilizer G_1 has an orbit of length 3 on Ω, then $|G_1| = 2^a \cdot 3$ for some $a \leqslant 4$.*

Sims' proof was very interesting, reducing the problem to properties of so-called "cubic graphs," previously established by Tutte [148, 149]. [Subsequently, Warren Wong [153], using Sims' result, completely classified the possibilities for G.]

From this example, Sims was led to make the following important conjecture.

SIMS' CONJECTURE. There exists a function $f(d)$ such that if G is a primitive permutation group on a set Ω in which a one-point stabilizer G_1 has an orbit of length d, then $|G_1|$ is bounded by $f(d)$.

In effect, the conjecture gives a general property of the embedding of maximal subgroups of finite groups.

Using the classification theorem, Cameron, Praeger, Saxl, and Seitz [39] have verified the Sims conjecture. An interesting feature of their proof is that it involves a result of Gödel on logical compactness.

4. *Algorithms.* Given a subgroup G of Σ_n generated by a set $\mathscr{S}$ of permutations, one is interested in properties of G that can be computed from $\mathscr{S}$ in *polynomial time*—i.e., in which the number of steps in the calculation is a polynomial in n.

A fundamental algorithm of Sims [120] (independent of the classification) asserts that $|G|$ can be determined in polynomial time. Using the classification theorem, Kantor [95] and Luks [100] have established the following further result.

THEOREM. *A composition series for G and an element of prime order p dividing* $|G|$ *can be determined in polynomial time.*

[Standard proofs of Cauchy's theorem require exponential time for determining an element of order p when p is large. We also note that there are many questions about subgroups of a group G for which the question of the existence of polynomial-time algorithms remains open.]

5. *The groups of Lie type.* Finally, we wish to mention Seitz's work on the groups of Lie type, which involves a beautiful interplay between the classification theorem, on the one hand, and algebraic group theory, on the other. His first investigations focused on conditions which force a given subgroup to be contained in a proper parabolic [115]. However, as the statements of most of these general results include some exceptional cases, we shall not state them here. But in the next section we shall describe his subsequent work on maximal subgroups of the groups of Lie type.

B. *Number theory.* The classification theorem has already had a number of significant applications to number theory and automorphic forms. We describe three.

1. *Brauer groups.* If L and K are fields with $L \supset K$, the *relative Brauer group* $B(L/K)$ of L over K consists of all "Brauer classes of finite-dimensional central simple K-algebras split by L." Moreover, K is a *global field* if K is either an algebraic number field or an algebraic function field in one variable over a finite field.

Fein, Kantor, and Schacher have established the following general result [51], which utilizes the classification theorem in a critical way.

THEOREM. *If K is a global field and L is a nontrivial finite extension of K, then $B(L/K)$ is infinite.*

That the proof should require the classification theorem is remarkable since $B(L/K)$ is an *abelian* group. However, the theorem reduces to verification of the following purely group-theoretic assertion, the only known proof of which depends upon the classification theorem.

THEOREM. *If G is a transitive permutation group on a set Ω of cardinality > 1, then there exists a prime p such that some element of G of order a power of p fixes no point of Ω.*

2. *Galois groups.* Emmy Noether [109] was the first mathematician to establish a significant sufficient condition for a finite group to be realizable as the Galois group of some finite algebraic extension of the rationals $\mathbf{Q}$, and the fundamental question of whether every finite group is so realizable is often referred to as "Noether's problem." In a long and difficult paper Shafarevich [133] has verified the desired conclusion for arbitrary *solvable* groups.

However, very little positive was known about the realizability of (non-abelian) simple groups. Apart from the alternating groups, as recently as 1983 only the single family $PSL_2(q)$ (and then only for certain congruences on q) and the one group $Sp_6(2)$ had been shown to be Galois groups over $\mathbf{Q}$. [The case of M_{11} was not completed until mid-1984, by Matzat.]

It is certainly conceivable that·by sophisticated general methods of algebraic number theory and function theory, without recourse to any specific properties of groups, one could demonstrate that every finite group generated by involutions—in particular, every simple group—is a Galois group over $\mathbf{Q}$. However, such a wonderful result appears to be still far off.

The classification theorem thus provides added stimulus for considering the problem with a fresh eye. Indeed, in the last few years Thompson has brought his deep understanding of finite groups to bear on the question and his work has already produced striking results. He realized the importance of the concept of "rigidity" (previously introduced by Belyĭ [24]) and saw that it could be used in many explicit cases to show that a given group G is a Galois group over $\mathbf{Q}$ ("rational" rigidity) or over some cyclotomic extension of $\mathbf{Q}$ (general rigidity).

The definition of rigidity is slightly simpler when G has a trivial center, so we limit ourselves to this case.

First, a conjugacy class X of elements of G is said to be *rational* if for $x \in X$, $\chi(x) \in \mathbf{Q}$ for every irreducible (ordinary) character χ of G.

Next, let $X_1, X_2, \ldots, X_k$ be conjugacy classes of elements of G and set

$$A = A_G(X_1, X_2, \ldots, X_k) = \left\{(x_1, x_2, \ldots, x_k)\,\middle|\, x_i \in X_i,\ x_1 x_2 \cdots x_k = 1\right\}.$$

With this terminology, we say that $A = A_G(X_1, X_2, \ldots, X_k)$ is *rigid* provided:

(1) $|A| = |G|$; and
(2) If $(x_1, x_2, \ldots, x_k) \in A$, then $G = \langle x_1, x_2, \ldots, x_k \rangle$.
Moreover, we call A *rationally rigid* if, in addition,
(3) X_i is rational for all i, $1 \leqslant i \leqslant k$.
The importance of rigidity is shown by the following result.

THEOREM. *If G is a group with trivial center which is rigid with respect to $A = A_G(X_1, X_2, \ldots, X_k)$ for suitable conjugacy classes $X_1, X_2, \ldots, X_k$, then G is a Galois group over some well-specified cyclotomic extension of $\mathbf{Q}$, and if G is rationally rigid with respect to A, then G is a Galois group over $\mathbf{Q}$.*

In all known cases in which rigidity has been verified, $k = 3$.
Using this result, Thompson has proved [137].

THEOREM. *The Fischer-Griess monster F_1 is a Galois group over $\mathbf{Q}$.*

Thompson [138, 139] and Feit and Fong [53] have applied the theorem to obtain the following Galois realizations.

THEOREM. *The groups $PSL_3(p)$, p a prime, $p \equiv 1 \pmod 4$, and $G_2(p)$, p a prime are Galois groups over $\mathbf{Q}$.*

Feit, Fong, and Srinivasan [52, 54], Thompson [140], and Walter [150] have shown that certain other classical groups, in suitable dimensions and over some fields, are Galois groups over $\mathbf{Q}$.

Finally, the combined work of Hunt, Hoyden, Matzat, Thompson, and Zeh [**87, 88, 89, 107**] has now shown that no less than eighteen of the sporadic groups are Galois groups over $\mathbf{Q}$ [namely, M_{11}, M_{12}, M_{22}, J_1, J_2, HS, Suz, ON, .3, .2, .1, $M(22)$, $M(23)$, $M(24)'$, F_5, F_3, F_2, F_1].

Strictly speaking, these Galois group results do not constitute applications of the classification theorem per se, but rather of the knowledge of the finite simple groups gained in the course of establishing the theorem. However, it should be mentioned that Matzat [**106**] has recently obtained some partial results on the "extension problem", proving under certain conditions that if X and Y are Galois groups over $\mathbf{Q}$, then so is an extension of X by Y. Thus, if in the future all simple groups were to be shown to be Galois groups over $\mathbf{Q}$ and if the extension problem were fully resolved as well, then the classification theorem *would* come to play a crucial role in deriving a solution of Noether's problem concerning the realizability of every finite group as a Galois group of $\mathbf{Q}$.

3. *Modular forms, graded Lie algebras, and the group* F_1. Despite the fact that the theory of elliptic functions has been thoroughly investigated over some 150 years, no one seems to have even guessed that the coefficients of the expansion of the Jacobi modular function at ∞ had any special significance:

$$J(q) = q^{-1} + 744 + 196{,}884q + 21{,}493{,}760q^2 + 864{,}299{,}970q^3 + \cdots ;$$

but shortly after the monster F_1 had been discovered, John McKay noticed that

$$196{,}884 = 1 + 196{,}883,$$

the sum of the degrees of the trivial representation and of the smallest faithful irreducible complex representation of F_1. Thompson then checked that the next few coefficients were likewise sums of degrees of irreducible representations of F_1 [**136**]: thus,

$$21{,}493{,}760 = 1 + 196{,}883 + 21{,}296{,}876$$
$$864{,}299{,}970 = 2 \cdot 1 + 2 \cdot 196{,}883 + 21{,}296{,}876 + 842{,}609{,}326.$$

Obviously something more than coincidence must be at work here!

The ultimate meaning of these connections remains to be determined, but deep relationships between the group F_1 and automorphic forms, on the one hand, and infinite-dimensional graded Lie algebras, on the other, have already been established. We should like to indicate the general nature of these results, without getting into technicalities.

There is a natural action of the group $GL_2^+(\mathbf{R})$, the group of 2×2 real matrices of positive determinant, on the upper half plane $\mathscr{H} = \{ z \mid \text{im } z > 0 \}$, given by

$$z \rightarrow \frac{az + b}{cz + d} \quad \text{for} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2^+(\mathbf{R}).$$

One also sets $\Gamma = SL_2(\mathbf{Z})$, and $\Gamma_0(N) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ with $c \equiv 0 \pmod{N}$ for any positive integer N.

If G is a subgroup of $GL_2^+(\mathbf{R})$ whose intersection with Γ is of finite index in both G and Γ, then the orbit space of G on $\mathscr{H}$ can be completed to a compact Riemann surface X_G. Thus associated to G is a field of meromorphic functions on X_G. Moreover, this field amounts to the field of G-*invariant* meromorphic functions on $\mathscr{H}$—i.e., such functions $f(z)$ on $\mathscr{H}$ satisfying

$$f(gz) = f(z) \quad \text{for all } g \in G.$$

We denote this field by $\mathscr{F}_0(X_G)$.

In particular, this applies with $G = \Gamma_0(N)$ and, for brevity, we set $\mathscr{F}_0(N) = \mathscr{F}_0(X_{\Gamma_0(N)})$.

Now consider the group F_1. Given "generalized" characters μ_n, $n \geqslant 1$, of F_1 (i.e., differences of ordinary characters), we form the formal power series

$$M_{F_1} = M = 1_{F_1} q^{-1} + \sum_{n \geqslant 1} \mu_n q^n,$$

where 1_{F_1} denotes the trivial character of F_1. We can evaluate this formal series for each $x \in F_1$, obtaining in each case a power series

$$M(x) = q^{-1} + \sum_{n \geqslant 1} \mu_n(x) q^n.$$

Beginning with John McKay's serendipitous observation and continuing with Thompson's observations, as well as the remarkable numerology of Conway and Norton [**44**], the following result linking F_1 to such "meromorphic forms" M_{F_1} has now been established (cf. Fong [**57**], S. Smith [**125**] and G. Mason [**105**]).

THEOREM. *There exist generalized characters* μ_n, $n \geqslant 1$, *of* F_1 *such that the corresponding formal power series M has the following properties:*

(i) *For each* $x \in F_1$, *there exists a positive integer* N_x *such that* $M(x) \in \mathscr{F}_0(N_x)$—*i.e.,* $M(x)$ *is a* $\Gamma_0(N_x)$-*invariant meromorphic function on* $\mathscr{H}$;

(ii) $N_x = |x| \cdot h$, *where* h *is a divisor of g.c.d.* $(24, |x|)$, *depending on* x; *and*

(iii) *If* $x = 1$, *then* $M(1) = J(q) - 744$, *where* $J(q)$ *denotes the Jacobi modular function.*

Furthermore, for each $x \in G$, *there is a subgroup* $\Gamma_0(N_x) \leqslant H_x \leqslant GL_2^+(\mathbf{R})$ *such that*

(1) X_{H_x} *has genus zero* (*i.e., is topologically a sphere*);

(2) $M(x)$ *is meromorphic on* X_{H_x}; *and*

(3) *Every meromorphic function on* X_{H_x} *is a rational function of* $M(x)$ (*so that* $M(x)$ *is a normalized generator of the field of meromorphic functions on* X_{H_x}).

There is very strong evidence that the μ_n are, in fact, ordinary characters.

In a major recent work, Frenkel, Lepowsky, and Meurman [**59**] have succeeded in relating F_1 and its associated Griess-Norton algebra B to so-called "vertex operators" on infinite-dimensional graded vector spaces. Here we can only give a conceptual description of their results.

First, with each even lattice Λ, they associate in a natural way an infinite-dimensional graded symmetric algebra, which they then tensor with the group algebra of Λ over a field of characteristic 0 containing suitable roots of unity. This yields a graded vector space V_Λ. For each $\alpha \in \Lambda$, the *vertex operator*

$X(\alpha, \zeta)$ is a precisely defined map from V_Λ to the set $V_\Lambda\{\zeta\}$ of all formal Laurent series $\sum_{n=-\infty}^{\infty} v_n \zeta^n$, $v_n \in V_\Lambda$, in the indeterminate ζ, defined in terms of suitable exponential series. The coefficient $x_\alpha(n)$ of ζ^n is an operator on V_Λ for each n. Frenkel, Lepowsky and Meurman refer to this as the *homogeneous representation* or *picture*.

A variation of this construction leads to a second space V_Λ', using a slight modification of the infinite-dimensional graded symmetric algebra from which V_Λ was defined as well as a suitable finite-dimensional module T for a central extension $\hat{\Lambda}$ of Λ. The $x_\alpha(n)$ also operate on V_Λ', and they refer to this as the *twisted representation* or *picture*.

Finally, they define $X^{\pm}(\alpha, \zeta) = X(\alpha, \zeta) \pm X(-\alpha, \zeta)$ and denote the coefficient of ζ^n in $X^{\pm}(\alpha, \zeta)$ by $x_\alpha^{\pm}(n)$.

This describes the formal context in which Frenkel, Lepowsky, and Meurman operate. Their first result, which is the key to their work, is quite general: with suitable restrictions on Λ and its module T, they are able to determine the operator

$$\left[x_\alpha^+(m), x_\beta^+(n) \right] \quad \text{for } \alpha, \beta \in \Lambda \text{ and } m, n \in \mathbf{Z}$$

in either the homogeneous or twisted picture when the inner product (α, β) of α, β is 0, -1, -2, or -4, with the stipulation that $\beta = -\alpha$ in the last case. [Here $[x_\alpha^+(m), x_\beta^+(n)]$ denotes the Lie product of $x_\alpha^+(m)$ and $x_\beta^+(n)$.]

They then apply this entire machinery to the case $\Lambda = $ Leech lattice, with T the uniquely determined 2^{12}-dimensional irreducible module for the 2-fold cover $\hat{\Lambda}$ of Λ. [In its action on T, $\hat{\Lambda}/\mathrm{ker}(\hat{\Lambda})$ is an "extra-special" group of order 2^{25}.] In this case there is an involution θ acting on V_Λ and V_Λ', whose ± 1-eigenspaces split V_Λ, V_Λ' into $V_\Lambda^+ + V_\Lambda^-$, $V_\Lambda'^+ + V_\Lambda'^-$, respectively, and the crucial space for the action of F_1 and its Griess-Norton algebra B is

$$V = V_\Lambda^+ + V_\Lambda'^-.$$

V is a graded vector space with gradings running from $+1$ to $-\infty$:

$$V = \sum_{n=-1}^{\infty} V_{-n}.$$

Furthermore, from the definition of θ, it is immediate that the $x_\alpha^+(n)$ act on V.

Here are the two principal results which Frenkel, Lepowsky, and Meurman prove. First, let $f_V(q)$ be the formal power series

$$f_V(q) = \sum_{n=-1}^{\infty} (\dim V_{-n}) q^n.$$

THEOREM. *We have* $f_V(q) = J(q) - 744$.

Thus $\dim V_0 = 0$ and for $n \neq 0$, $\dim V_{-n}$ is precisely the coefficient of q^n in the expansion of $J(q)$. Hence the theorem gives a functorially described graded module associated with the modular function $J(q)$.

Next, we write (F_1, B) for F_1 acting on B.

THEOREM. *V is a graded module for* (F_1, B) *in a precisely defined sense.*

There then is the marvelous infinite-dimensional Lie algebra connection of F_1 with the modular function. It is clear that it will take some time for the full significance of these spectacular results to be understood.

C. *Model Theory.* The classification theorem has already found two striking applications in model theory. The first of these was proved by Cherlin, Harrington, and Lachlan [40] and independently by B. L. Zilber [154].

THEOREM. *A totally categorical theory with no finite models cannot be finitely axiomatized.*

A theory is said to be *totally categorical* if any two models of the theory of the same cardinality are necessarily isomorphic.

The proof of the theorem involves a study of so-called "strongly minimal sets" (which I shall not define) and their automorphism groups, which in turn reduces to the classification of certain types of primitive Jordan groups. Here a group G acting faithfully as permutations of a set Ω (of arbitrary cardinality) is said to be a *Jordan* group provided there exists a finite subset X of Ω with $|X| > 1$ and $|\Omega - X| > 1$ such that the subgroup G_X of elements of G leaving X set-wise invariant acts transitively on $\Omega - X$. [Double transitivity corresponds, of course, to the case in which the given condition holds with $|X| = 1$.]

The primitive Jordan groups arising as automorphism groups of strongly minimal sets turn out to be "geometrically locally finite" in a well-defined sense (they arise as limits of certain sequences of finite primitive Jordan groups). The classification of the doubly transitive finite simple groups enables one to determine all finite primitive Jordan groups and, in turn, all infinite such groups. [There are precisely three types of infinite primitive Jordan groups that are geometrically locally finite, corresponding to the cases in which Ω has no structure, Ω has the structure of an infinite-dimensional affine vector space over some $GF(q)$, or an infinite-dimensional projective space over some $GF(q)$ and correspondingly $G = \Sigma_\Omega$, $AGL(\Omega) \leqslant G \leqslant A\Gamma L(\Omega)$, or $PGL(\Omega) \leqslant G \leqslant P\Gamma L(\Omega)$.]

It should be noted that Zilber has given a (very difficult) proof of the classification of geometrically locally finite primitive Jordan groups, using a combination of combinatorial analysis, logic, and algebra, so that the classification of the simple double transitive groups can be avoided here. However, the theorem effectively illustrates how a question in logic reduces to one concerning finite groups.

The second result, due to Cherlin and Lachlan [41], at present requires the full simple group classification theorem and has no other proof. I shall not state the theorem precisely; it provides a classification of "stable theories in finite relational languages" in which it is possible to "eliminate quantifiers."

However, I shall state the permutation-theoretic result on which the theorem depends, the proof of which uses the classification theorem.

PROPOSITION. *Let r and s be positive integers. Then there exists a positive integer $f(r, s)$ such that for any transitive permutation group G on a set Ω of cardinality $\geqslant f(r, s)$ for which the number of orbits of G on the Cartesian product $\Omega^5 = \Omega \times \Omega \times \Omega \times \Omega \times \Omega$ is less than s, there exists a subset X of Ω of cardinality r such that $G_X/G^X \cong \Sigma_r$, where G^X denotes the normal subgroup of G_X of elements leaving each point of X invariant.*

[G_X/G^X is the permutation group induced on X by its set stabilizer G_X.]

These results suggest that the classification of the simple groups may well play an important role in future developments in mathematical logic.

3. Whither finite group theory? For a time there was a feeling that completion of the classification proof might well lead to the demise of finite group theory—not to those areas of the subject such as representation and character theory, solvable groups, nilpotent groups, etc., which had always had a life independent of the classification theorem, but rather to those parts that had developed out of the "local group-theoretic" techniques underlying the classification proof. And it is true that some simple group theorists have switched their interests to such fields as number theory, infinite group theory, or geometry. But as the number of significant applications of the theorem has grown, it has become increasingly apparent that the vitality of finite group theory has remained intact, for the classification theorem and the techniques its proof generated are opening up entire new research directions, not only of intrinsic interest, but ones that can be expected to yield significant applications in the future.

I shall limit myself here to two particularly active areas of current research: the first concerns the interrelationship between representations of groups and maximal subgroups, the second describes recent developments in the theory of groups of "weak Lie type," a broad generalization of the notion of a group of Lie type, whose study involves a delicate interplay of local group-theoretic and geometric techniques.

A. *Representations and maximal subgroups.* Aschbacher [17] has suggested that, in view of the classification theorem, it is important to develop techniques which reduce questions about arbitrary finite groups or their representations to questions about simple groups or more precisely about "almost" simple groups.

The latter is an important group-theoretic notion, so let me define it. First, a group K is said to be *quasisimple* if K is a perfect central extension of a simple group—i.e., $K = [K, K]$ and $K/Z(K)$ is simple.

In particular, simple groups are quasisimple (the case $Z(K) = 1$). However, the groups $K = SL_n(q)$, $n \geqslant 3$ or $n = 2$ and $q > 3$, for example, are quasisimple, but not necessarily simple. Indeed, if, say, n is even and q is odd, the matrix $-I$ is a nontrivial element of $Z(K)$.

A group X is said to be *almost simple* if X contains a normal quasisimple subgroup K such that $X/Z(K)$ is a subgroup of the automorphism group

Aut($K/Z(K)$) of the simple group $K/Z(K)$. [Note that $Z(K)$, being char-
acteristic in K, is normal in X, so $X/Z(K)$ is well defined.] Thus an almost
simple group is obtained from a simple group by adding "decoration" at the
"bottom" ($Z(K)$) and at the "top" (outer automorphisms).

Now consider Aschbacher's observation for permutation representations of a
group G. The *indecomposable* permutation representations of G are its transi-
tive such representations and these are in a natural one-one correspondence
with the conjugacy classes of its subgroups, while the *irreducible* permutation
representations of G are its primitive representations, which are in one-one
correspondence with the conjugacy classes of its maximal subgroups. Thus to
reduce questions about such indecomposable representations to questions
about such irreducible representations, one needs information about maximal
subgroups containing a given subgroup—i.e., maximal "overgroups" of sub-
groups. One would also like to be able to reduce questions about maximal
subgroups of G to questions about almost simple groups.

These overgroup and reduction-to-simplicity questions have direct analogues
for *linear* representations of groups, both over fields of characteristic 0 or
$p > 0$.

Depending on the specific situation being studied, the analysis may involve
the classification theorem, general group-theoretic methods, the geometry of
the classical groups, the theory of algebraic groups, or properties of the
sporadic groups.

I shall describe a few of the results obtained in these directions, the first
being a theorem of O'Nan and Scott [21].

THEOREM. *Let G be a primitive permutation group acting on a set Ω and let N
be the product of all minimal normal subgroups of G. Then one of the following
holds*:

(i) *N is an elementary abelian p-group for some prime p acting regularly on Ω*
(*i.e.*, *$N \cong Z_p \times Z_p \times \cdots \times Z_p$, N is transitive on Ω, and no nonidentity
element of N fixes any point of Ω*); *or*

(ii) *$N = N_1 \times N_2 \times \cdots \times N_m$ for suitable isomorphic nonabelian simple
groups N_i, $1 \leqslant i \leqslant m$.*

In the case of (ii), the theorem gives much more detailed information,
asserting that G has one of three possible actions and general structures. We
mention only the first of these, for it shows how the analysis of the general
primitive permutation group reduces to the study of such almost simple
groups.

This alternative asserts that there exists a subgroup G_1 of the automorphism
group of N_1 with the following properties

(1) if $|\Omega| = n$, then $n = n_1^m$ for some integer n_1 and G_1 acts as a primitive
permutation group on a set Ω_1 of cardinality n_1; and

(2) If $G_1, G_2, \ldots, G_m$ denote isomorphic copies of G_1, B denotes the
symmetric group on the set $\{G_1, G_2, \ldots, G_m\}$, and A denotes the direct
product $G_1 \times G_2 \times \cdots \times G_m$, then G is isomorphic to a subgroup of
the "semidirect product" of A by B [in the language of group theory,
$G \leqslant (G_1 \text{ wreath } \Sigma_m)$].

Thus (1) reduces the action of the primitive group G to that of its primitive almost simple subgroup G_1, while (2) describes the general relationship between the structure of G and that of G_1. The other two possibilities of the theorem likewise provide reductions for the analysis of primitive permutation groups.

As a corollary of the theorem, together with the classification of the simple groups, one obtains an essentially complete solution of a long-standing permutation group problem: the determination of all primitive permutation groups of prime power degree not containing a regular normal subgroup (in particular, all such simple groups).

Furthermore, in this same terminology, Aschbacher and Scott [21] have given a complete general description of the maximal subgroup M of G determining the given permutation representation.

We turn next to linear representations. Aschbacher has obtained the following result for the classical groups [12].

THEOREM. *Let G be a finite classical group of characteristic p and let V be the natural module for G. Then there exists a prescribed collection $\mathcal{M}(G)$ of subgroups of G such that if X is any subgroup of G, one of the following holds:*
   (i) *X is contained in some element of $\mathcal{M}(G)$; or*
   (ii) *X is almost simple and acts absolutely irreducibly on V.*

The set $\mathcal{M}(G)$ consists of stabilizers of subspaces or sets of subspaces, normalizers of classical groups over subfields or extension fields of the field of definition of V, normalizers of subgroups preserving a tensor product decomposition of V, or normalizers of certain so-called "extra-special" subgroups.

Seitz has been systematically analyzing the absolutely irreducible subgroups X of the general linear group $G = GL(V)$, with V of characteristic $p > 0$ and X a group of Lie type of characteristic p, reducing questions to those for the corresponding algebraic groups $\tilde{X}$ and $\tilde{G} = SL(\tilde{V})$, where $\tilde{V}$ is defined over the algebraic closure $\overline{F}_p$ of $GF(p)$.

Here is a first result of Seitz [94] in the same spirit as the above.

THEOREM. *If Y is a subgroup of G containing X, then with a finite number of (low Lie rank) exceptions for X, the normal closure of X in Y is an elementwise commuting product of groups $Y_1, Y_2, \ldots, Y_n$, where each Y_i is either a group of Lie type of characteristic p or an alternating group.*

In essence, this reduces the general problem to the almost simple case $Y = Y_1$. In this case Seitz's goal is to reduce to the following algebraic group problem.

$(*)$
   (1) $\tilde{X} \leqslant \tilde{Y} \leqslant \tilde{G} = SL(\tilde{V})$;
   (2) $\tilde{X}$ and $\tilde{Y}$ are simply connected algebraic groups;
   (3) The representation of $\tilde{X}$ on $\tilde{V}$ is irreducible; and
   (4) The representation of $\tilde{Y}$ on $\tilde{V}$ is rational.

Concerning the algebraic group problem (∗), Dynkin [**47, 48**] has obtained a complete solution of the corresponding characteristic 0 problem in two long and difficult papers. In particular, as a corollary, he has determined all maximal connected subgroups of classical groups in characteristic 0. However, as Seitz notes, the problem in characteristic p is considerably more involved because some of Dynkin's reductions fail and, in addition, there are embeddings with which Dynkin did not have to contend.

I conclude with one of Seitz's algebraic group results, which is important for his general theory [**116**].

THEOREM. *If $\tilde{X}$ and $\tilde{Y}$ have the same rank as algebraic groups, then the following holds*:

  (i) $\tilde{X} = B_n(F_2)$, $C_n(F_2)$, $F_4(F_2)$, *or* $G_2(F_3)$;

  (ii) $\tilde{X}$ *is generated by a suitable subsystem of root subgroups in a fixed root system of $\tilde{Y}$; and*

  (iii) $\tilde{V}$ *is completely specified as a $\tilde{Y}$-module, determined by a suitable dominant weight.*

These results should give a good indication of the direction of current research in this important topic.

B. *Groups of weak Lie type and parabolic systems*. The groups of Lie type possess a very precise, uniformly described *internal* structure, determined by generation by their *parabolic* subgroups. Moreover, with the set of parabolics of such a group G there is associated a natural geometry, on which G acts as a group of automorphisms. These geometries are examples of what Tits calls "spherical buildings" of rank r, where r is the Lie rank of the given group G—i.e., the rank of the Weyl group W of G. A fundamental result of Tits [**145**] gives a complete geometric classification of spherical buildings of rank $\geqslant 3$, showing, in effect, that every such building arises as the associated building of some group of Lie type.

In the course of the simple group classification proof, finite group theorists were forced to investigate certain groups generated by a set of "parabolic subgroups" satisfying conditions somewhat looser than those of a group of Lie type. These groups, too, possess natural associated geometries, which were then studied abstractly and the results fed back into the classification proof.

Moreover, beginning with the work of Buekenhout [**38**] and of Ronan and S. Smith [**114**], there developed, quite apart from the classification proof, an interest in finding such geometric descriptions of the sporadic groups in terms of their subgroup structure. I might also mention here Timmesfeld's subsequent lovely work on "single-bond" geometries [**143**] and Aschbacher's comprehensive investigation of the subgroups of sporadic groups containing a given Sylow p-subgroup and their associated geometries [**19**]. In fact, since completion of the classification, there has been a very intense effort to place this "parabolic-geometric" chapter of simple group theory in a general setting.

I shall refer here to the class of groups involved as groups of *weak* Lie type. Such groups are generated by a *parabolic system* of subgroups. The associated geometries are known as *chamber systems*.

At one end of the spectrum, one stays as close to the given group as possible, analyzing its parabolic structure using only local group-theoretic methods. At the other end, one studies abstract chamber systems geometrically, divorced from any group-theoretic context. In between, there are problems of group theory which require a combination of local group-theoretic and geometric methods for their analysis.

This is a very rich subject, which I fully expect to have lasting impact on both simple group theory and geometry. One of its natural goals is to define a class of geometries capable of classification, but broad enough to encompass both the groups of Lie type and the sporadic groups as transitive automorphism groups. Such a result, although of considerable conceptual value for understanding the simple groups, would not directly affect the classification theorem. However, there are other parts of the subject where progress, already substantially achieved, will produce significant improvements in portions of the existing classification proof.

I shall limit most of my comments here to the parabolic-geometric situation since it is the part of the subject closest to the group-theoretic context.

First, consider a group G of Lie type of characteristic p. Then G has a decomposition entirely analogous to the Bruhat decomposition of a simple Lie group. Let U be a Sylow p-subgroup of G (the letter U because the elements of U are "unipotent" in the Lie sense). The normalizer B of U in G is called a *Borel* subgroup. By Sylow's theorem, any two Borel subgroups are conjugate in G. Furthermore, B has the form $B = HU$, where $H \cap U = 1$ and H is abelian. H is called a *Cartan* subgroup of G. By the Schur-Zassenhaus theorem [**66**, Theorem 6.2.1], all Cartan subgroups of B (i.e., all complements to U in B) are conjugate in B, so all Cartan subgroups are likewise conjugate in G.

The basic structure of G is determined by B and a proper subgroup N with the following properties:

(*)
> (1) $G = BNB$; in particular, G is generated by B and N;
>
> (2) H is normal in N; and
>
> (3) $W = N/H$ is a group "generated by fundamental reflections" w_1, $w_2, \ldots, w_r$.

The group W is called the *Weyl* group of G (in view of the above conjugacies, uniquely determined by G). In particular, the elements w_i are involutions, which together generate W. I shall not give the precise definition of a group generated by reflections, but an example is the symmetric group Σ_{r+1} with w_i corresponding to the transposition $(i, i + 1)$, $1 \leqslant i \leqslant r$. Σ_{r+1} is, in fact, the Weyl group of the linear groups $A_r(q)$. In general, the integer r is called the *Lie rank* of G.

There is also a precise rule for multiplying any two elements of G of the form xyz for $x, z \in B$ and $y \in N$ (which we also omit), so that, in effect, the multiplication table of G is completely determined by B and N.

Next, set $I = \{1, 2, \ldots, r\}$. For any subset J of I, let W_J be the subgroup of W generated by the w_j for $j \in J$ and let N_J be the complete inverse of W_J in N. The subgroup P_J generated by B and N_J is called a *parabolic* subgroup of G. We have, in fact,

$$P_J = BN_J B.$$

Moreover, if U_J denotes the unique largest normal p-subgroup of P_J—the so-called " unipotent radical" of P_J—then we have

(1) U_J contains its own centralizer in P_J; and

(2) The subgroup $O^{p'}(P_J/U_J)$ of P_J/U_J generated by its p-elements is a group of Lie type.

If $J = \{j\}$, P_J is called a *minimal* parabolic and one writes P_j, U_j for P_J, U_J, respectively. Thus P_j is determined by the single reflection w_j and $O^{p'}(P_j/U_j)$ is a group of Lie type of Lie rank 1. This last condition implies that B is a maximal subgroup of each P_j, $j \in I$.

For any subset J of I, one clearly has

$$P_J = \langle P_j \mid j \in J \rangle,$$

so that all parabolics are determined from the minimal parabolics. We call $\{P_j, 1 \leqslant j \leqslant r\}$ a *parabolic system* for G.

The Bruhat decomposition of G, as described in (∗), is reflected in very tight interdependencies among the parabolics P_J. This should provide a rough picture of the parabolic structures of the groups of Lie type.

If one focuses solely on the generational aspects of the parabolics, dropping the special structural conditions specified by (∗), one obtains a broad generalization of the notion of a group of Lie type. We shall say that a group G possesses a *parabolic system* $\{P_i, 1 \leqslant i \leqslant r\}$ *centered* at B for some set of proper subgroups P_i of G, $1 \leqslant i \leqslant r$, provided

(1) $G = \langle P_i \mid 1 \leqslant i \leqslant r \rangle$; and

(2) B is a maximal subgroup of P_i for all i, $1 \leqslant i \leqslant r$.

One can replace (2) by the weaker condition $B \leqslant P_i$ for all i, but in its present form the P_i more closely resemble minimal parabolics. The integer r is called the *rank* of G. Since the P_i are proper, condition (1) forces $r \geqslant 2$.

As in the Lie type case, we set $I = \{1, 2, \ldots, r\}$, and for any subset J of I, we define

$$P_J = \langle P_j \mid j \in J \rangle.$$

The P_J are again called *parabolic* subgroups of G.

Without additional restrictions on the structure of the P_J and of B, it appears to be very difficult to derive substantial information about the structure of G. In practice, the cases of primary interest are those in which the

P_i have structures resembling those of the minimal parabolics in a group of Lie type.

This leads to the following definition. A group G will be said to be of *weak Lie type* of *characteristic p* provided the following conditions hold.

(1) G has a parabolic system $\{ P_i \mid 1 \leqslant i \leqslant r \}$ centered at B;

(2) If U is a Sylow p-subgroup of B, then B is the normalizer of U in P_i for all i, $1 \leqslant i \leqslant r$; and

(3) If U_i denotes the unique largest normal p-subgroup of P_i, then U_i contains its own centralizer in P_i, $1 \leqslant i \leqslant r$; and

(4) $O^{p'}(P_i/U_i)$ is a group of Lie type of Lie rank 1 (or when $p = 2$ a dihedral group), $1 \leqslant i \leqslant r$.

With any group G possessing a parabolic system, there are associated certain natural incidence geometries determined by the coset spaces of the parabolics of G. In the above notation, one can consider the space $\mathscr{G}$ of all left cosets gP_J as J ranges over all subsets of I and g ranges over G, with two objects $X, Y \in \mathscr{G}$ defined to be *incident* if and only if $X \cap Y \neq \varnothing$. Moreover, the elements of G induce an *incidence-preserving* action on the objects of $\mathscr{G}$ by left multiplication, so that G acts as a group of automorphisms of $\mathscr{G}$.

It is this incidence geometry that leads to the spherical building description of the groups of Lie type.

Tits [121] has associated a second natural geometry with any such group G, defined solely in terms of the parabolic system P_i and its center B, on which G acts as a *transitive* group of automorphisms. The objects of $\mathscr{C}$ are the left cosets of B (on which G does indeed act transitively) and for each $i \in I = \{1, 2, \ldots, r\}$, two cosets gB, $g'B$ with g, $g' \in G$ are said to be *i-adjacent* provided $gP_i = g'P_i$. Clearly, left multiplication by elements of G preserves *i*-adjacency, so again G acts as a group of automorphisms on $\mathscr{C}$.

$\mathscr{C}$ is called the associated *chamber system* of G. The objects of $\mathscr{C}$, called *chambers*, are the elements of the left coset space G/B, while incidence is determined by the left coset spaces G/P_i, $i \in I$. Observe that the left cosets of each P_i define a *partition* of the set $\mathscr{C}$ of left cosets of B, with gB, $g'B$ being *i*-adjacent if and only if they lie in the same element of this partition.

Tits has used this last observation as the basis of a general notion of a *chamber system* as a set $\mathscr{C}$, called *chambers*, together with partitions π_i of $\mathscr{C}$, i ranging over some index set I, which determine incidences in $\mathscr{C}$ by the condition c, $c' \in \mathscr{C}$ are *i*-adjacent if and only if c and c' lie in the same element of the partition π_i.

Chamber systems can be considered independently of any group-theoretic context and analyzed geometrically. This is presently a very active area—the kind of assumptions imposed on $\mathscr{C}$ are related to the structures of certain associated "residue" geometries (such residue geometries also play an important role in studying chamber systems arising from groups).

Thus one can analyze a given group G having a parabolic system either *internally*, in terms of the structure of its parabolic subgroups, or *externally*, in terms of its action on its associated chamber system. Depending on the context, both perspectives have been effective in obtaining structural information about G.

In the rank $\geqslant 3$ case, the configurations leading to groups of Lie type can usually be identified geometrically. In other cases, the configurations investigated usually lead to an "affine" diagram, so that at least an infinite "covering group" $\overline{G}$ of G can be determined. Moreover, a fundamental result of Tits (unpublished) classifies all such affine buildings of rank $\geqslant 4$, so that the possibilities for G in this case are determined *externally* by his theorem.

Thus the ranks 2 and 3 cases are exceptional, and their analyses seem to require a combination of local group theory and geometry. Furthermore, by itself that analysis may not determine the possible isomorphism types of G, but only the structure of the parabolics P_1, P_2 in the rank 2 case and the parabolics $\langle P_i, P_j \rangle$, $1 \leqslant i, j \leqslant 3$ in the rank 3 case. In some instances, it yields their exact isomorphism types, while in others only their orders, composition factors, etc., so that further analysis is necessary to pin down G itself.

The first rank 2 parabolic system to be considered arose in Sims' work on primitive permutation groups G in which a one-point stabilizer P_1 has an orbit or rank 3. It follows directly from his conditions that for some $g \in G$, if we set $P_2 = P_1^g$, then G is a group of weak Lie type with parabolic system $\{P_1, P_2\}$ of characteristic 2 with center $B = P_1 \cap P_2$ a Sylow 2-subgroup of G and such that

$$P_i / U_i \cong PSL_2(2) \left(\cong \Sigma_3 \right), \qquad i = 1, 2.$$

Goldschmidt [63] later investigated the same general situation, but without the restriction that P_1 and P_2 be G-conjugate (also requiring only that B be Sylow in P_i, but not necessarily in G). Goldschmidt's paper is fundamental for all subsequent amalgam-type analyses of groups of weak Lie type of ranks 2 or 3. He considered the "free product $(P_1 * P_2, B)$ amalgamated at $B = P_1 \cap P_2$ (i.e., the free group generated by P_1 and P_2, subject to the identification that B be a subgroup of both P_1 and P_2). Associated with this infinite *amalgam* $(P_1 * P_2, B)$ is a *tree* on which G acts as a group of transformations. By analyzing this action, Goldschmidt was eventually able to determine the possible isomorphism types of P_1 and P_2.

His answers include, of course, the parabolic systems of the Chevalley groups of Lie rank 2 over the prime field $GF(2)$: namely, $PSL_3(2)$, $PSp_4(2)$, and $G_2(2)$, but a number of other possibilities occur as well (e.g., M_{12}).

In the past few years, the groups of weak Lie type of rank 2 have been closely studied, and their possible parabolic systems have now been essentially completely determined. The definitive result is due to Delgado and Stellmacher [46], based on a simplified approach to the Goldschmidt method, in which his geometric so-called "track" analysis is replaced by an easier one more closely related to local group-theoretic analysis. We conclude this discussion with the statement of their main theorem.

For brevity, if G has a parabolic system $\{P_1, P_2\}$, we shall say that G is of *parabolic type* G^* for some known simple group G^* with parabolic system $\{P_1^*, P_2^*\}$ provided the structure of P_i "closely resembles" that of P_i^*, $i = 1, 2$ (in particular, same orders and same composition factors).

THEOREM. *If G is a simple group of weak Lie type of characteristic p and rank 2, then one of the following holds.*

(i) *G is of parabolic type* $PSL_3(p^n)$, $PSp_4(p^n)$, $PSU_4(p^n)$, $PSU_5(p^n)$, $G_2(p^n)$, *or* $^3D_4(p^n)$ *for some n;*

(ii) $p = 2$ *and G is of parabolic type* $^2F_4(2^n)$, $n > 1$, $^2F_4(2)'$ (*of index* 2 *in* $^2F_4(2)$), $G_2(2)'(\cong PSU_3(3))$, M_{12}, *or* J_2; *or*

(iii) $p = 3$ *and G is of parabolic type* F_3.

4. The general significance of classification theorems.

In the course of a recent long and very interesting interview in the *Mathematical Intelligencer*, Michael Atiyah [22] comments on several broad philosophical questions about the classification of the finite simple groups and classification theorems, in general, which raise some issues that will be central to us here: (A) How significant are classification theorems for mathematics as a whole? (B) Is an internal or external approach to the proof of a classification theorem likely to be more effective? (C) Is the central aim of mathematics to establish broad general theorems involving no exceptional cases? (D) Doesn't the extreme length of the simple group classification theorem signify that group theorists have taken the wrong approach?

I feel that a detailed response to these fundamental questions will serve to place the classification of the simple groups in its proper mathematical perspective. I shall discuss the four questions separately.

A. *Classification as structure.* Atiyah expresses the view that classification theorems are of less importance than the action of the given system on other mathematical objects. Thus the theory of group representations becomes more fundamental than the classification of the simple groups.

Part of the problem here is the use of the word *classification*, which conveys the idea of simply determining the *list* of objects satisfying some specified axiom system. However, to me, classification refers to an attempt to understand the *intrinsic structure* of a given mathematical system.

I can understand a preference for the kind of problems arising from one perspective rather than another, but surely determining the basic structure of a mathematical system is as inherently significant as studying the way in which the system acts. For Brauer the primary objects of interest were first algebraic number fields and later finite groups—and he always regarded representation and character theory as *tools* to increase his understanding.

But beyond this, I think the whole dichotomy between *internal properties* and *external actions* is an artificial one, for in reality they often represent two *essentially equivalent* ways of viewing the same mathematical object. Thus

a. Every irreducible representation of a finite-dimensional associative semi-simple complex algebra A is equivalent to one determined internally by a minimal right ideal of A.

b. Every primitive permutation representation of a finite group G is equivalent to one determined internally by a maximal subgroup of G.

c. The solvability of a given irreducible equation over **Q** by radicals (internal property) is equivalent to the solvability of the Galois group of the corresponding number field determined by its roots (external action).

What is really at issue is the effectiveness of a given perspective in providing tools for understanding a particular system or for solving a particular problem.

B. *External versus internal approach to classification.* Many mathematicians have suggested that the enormous length and complexity of the classification of the simple groups is a consequence of our insistence on working entirely "inside the system." It should be more effective to study simple groups through their representations or some geometric realization, for a much wider array of tools would then be available, which will ultimately yield a deeper understanding and a shorter classification proof.

Such an external viewpoint may well be the most appropriate for analyzing many mathematical systems, but for the finite simple groups the evidence strongly suggests that the techniques developed through the internal approach are of much greater power than any that an external perspective is capable of providing. How else can one explain the fact that the internal approach has not only led to the discovery of eighteen of the twenty-six sporadic groups (the five Mathieu and three Conway groups arose from external considerations), but has proved that the present list of simple groups is complete, whereas in the 150 years since Galois, all external manifestations of finite groups have yielded only extremely limited information about the structure of the general simple group?

The odd order theorem has been reduced to questions in both topology and ring theory, but twenty years after the Feit-Thompson original proof [55], their internal strategy remains the only known viable approach.

I believe that in the end it will be our internal approach to the study of finite groups that will provide deep insights into their external realizations rather than the other way around. The various examples I have described in §2 and §3 give strong evidence in support of this judgment.

The difficulty of the classification theorem also raises the question of whether the theorem was worth the effort, since it is often the case that in a given area of mathematics, the most important things may well be of an elementary, general nature, for they are the ones with the widest impact.

But there is also the matter of *depth* as well as *breadth*. Many years ago George Mackey remarked to me that the best mathematics consists of easily understood statements whose proofs involve deep and interesting mathematical ideas.

The Feit-Thompson theorem epitomizes Mackey's observation: its statement "Every group of odd order is solvable" takes only a single line, while its long and deep 255-page proof is filled with wonderful new ideas, foreshadowing the developments that led to the ultimate classification of the simple groups.

C. *General versus singular mathematics.* In his interview, Atiyah tends to down-play the significance of "singular" solutions to problems. Since, in practice, one usually encounters only the *classical* groups, he perceives the *exceptional* Lie groups as a somewhat troublesome complication rather than as objects of intrinsic interest.

Apart from the fact that we have no control over mathematical reality—the five exceptional Lie groups and the twenty-six sporadic simple groups are,

after all, fixed stars of our universe that no amount of speculation will remove —I for one feel it is the existence of singular solutions that gives mathematics its rich texture and their suppression, were that possible, would make mathematics too homogenized for my taste.

Beyond their intrinsic interest, singular solutions often provide insight into the nature of general solutions. Tits remarked to me that for him E_8 is the fundamental testing ground for any general property of simple Lie groups he may hope to establish. Only when he can construct a proof that works for E_8 does he feel he has reached a true understanding of the property in question. His point is that proofs which work for E_8 as well as for the classical groups are bound to produce deeper insights. For this reason, Tits views the urge to find such proofs as a "moral" principle for Lie theorists.

In a recent article, "Unfashionable Pursuits," Freeman Dyson [49] speculates about the future of physics in a way that vividly captures the appeal of singular solutions in mathematics: "I have to confess to you that I have a sneaking hope, a hope unsupported by any facts or evidence, that sometime in the twenty-first century physicists will stumble upon the monster group, built in some unsuspected way into the structure of the universe. This is of course only a wild speculation, almost certainly wrong. The only argument I can produce in its favor is a theological one. We have strong evidence that the creator of the universe loves symmetry, and if he loves symmetry, what lovelier symmetry could he find than the symmetry of the Monster?"

D. *The inherent complexity in classifying the simple groups*. I have often heard the view expressed that the present proof of the classification theorem must certainly be the wrong one—no theorem in mathematics can possibly require such an absurdly long and technically complex proof! With this in mind, Thompson once remarked, only half in jest: "We finite group theorists are either unbelievably stupid or very clever."

I would like to explain the intuitive basis for my deeply held belief that the classification of the finite simple groups is an *inherently difficult* theorem. The perspective I shall present here reflects the core of my understanding of the finite simple groups and will be critical for all that follows.

Since the bulk of the simple groups are finite analogues of the Lie groups, it is natural to compare their classification with that of the relatively short proof that suffices to classify the simple Lie groups—in view of Lie's duality theorem, to classify the finite-dimensional complex simple Lie algebras.

There are several essentially equivalent ways of classifying the finite-dimensional simple Lie algebras L over any field F of characteristic 0, each involving the following five basic components.

1. The existence of Cartan subalgebras H in L.

2. The Killing form and its realization on the dual space H^* of H.

3. Root space decompositions of L relative to the action of H^*

4. The nondegeneracy criterion on the Killing form for the semisimplicity of L (i.e., for the triviality of the radical Rad(L)).

5. The decomposition of a semisimple algebra as the direct sum of simple algebras.

Taken together, these five results reduce the classification to the solution of a specific geometric problem concerning a spanning set of vectors of the Euclidean space H^* — the so-called "root lattice" of L — which in turn reduces to the classification of "connected" Dynkin diagrams. Indeed, it is ultimately shown that Lie multiplication of the elements of a suitable vector space basis of L is *uniquely* determined by its associated diagram.

Thus the total classification proof is almost entirely *internal Lie-theoretic*, with a short *external geometric* argument at the end.

[Uniqueness proves only that there is *at most* one simple algebra with a given connected diagram, and is independent of the question of whether there exists an algebra for each such diagram, which requires separate verification.]

Before comparing Lie algebras with finite groups, it will be useful to examine the case of finite-dimensional simple Lie algebras over algebraically closed fields of characteristic $p > 0$, which I visualize as a "half-way station." The classification of all such simple Lie algebras is by no means complete, but recently Block and Wilson [31] have determined all so-called *restricted* simple algebras when $p > 7$: every such algebra is either "classical" (a characteristic p analogue of a finite-dimensional complex simple algebra) or of so-called "Cartan type," a class of algebras introduced by Kostrykin and Shafarevich [94], who first conjectured the Block-Wilson theorem. [Right Lie multiplication of L by the element $x \in L$ induces a derivation of L, called the *adjoint transformation* and denoted by ad x. A Lie algebra L of characteristic $p > 0$ is said to be *restricted* if $(\text{ad } x)^p$ is an inner derivation for all $x \in L$. I note also that the Killing form of L (in any characteristic) is the bilinear form trace $(\text{ad } x \cdot \text{ad } y)$ for $x, y \in L$.]

In the characteristic $p > 0$ case, there still exist Cartan subalgebras, but they are no longer necessarily "tori" nor need they all be conjugate. Likewise the corresponding dual spaces exist, as do root decompositions of L, but now there is no nondegeneracy criterion for semisimplicity. Furthermore, semisimple algebras have a far more complex structure than in the classical characteristic 0 case. Thus Block and Wilson have available only a more complex decomposition theorem. As a result, their analysis is much more involved and technically difficult than in the classical case. In fact, its flavor is not unlike that of local group theory. Indeed, they express L as a sum (not necessarily direct) of rank 2 subalgebras $L_{\alpha\beta}$, where α, β are suitable independent roots (relative to a suitably chosen Cartan subalgebra). Thus they first need a complete description of all such restricted semisimple rank 2 algebras. When the Cartan subalgebra is a torus, they show that there are twenty-three such rank 2 algebras, while in the general case there turn out to be twenty-three suitably parametrized such semisimple algebras. Then by a delicate analysis of the global interdependency of the subalgebras $L_{\alpha\beta}$, $L_{\beta\gamma}$, $L_{\alpha\gamma}$ as α, β, γ range over suitable independent roots, they are able to eliminate most potential possibilities for $L_{\alpha\beta}$ from the given restricted simple Lie algebra L (of rank ≥ 3).

"Lining up" the $L_{\alpha\beta}$'s enables Block and Wilson to determine the general structure of a certain *maximal* subalgebra L_0 of L and the general nature of the adjoint action of L_0 on the factor space L/L_0. This information must now be used to identify the possibilities for L. However, this is a far more elaborate

task than in the characteristic 0 case. In fact, it is reasonable to view the existence of L_0 as a "half-way" point in the analysis, for in the case that L turns out to be classical, L_0 is precisely the *centralizer* of a suitable element in a Cartan subalgebra. Thus the classification theorem splits into two parts: first, determining the general shape of the centralizer L_0 of a suitable Cartan element (as well as of its action on L/L_0); then showing that this information suffices to force the possibilities for L.

The second part of the analysis depends critically on a major classification theorem of Kac concerning graded Lie algebras [93, 94]. Indeed, L_0 determines a natural 'filtration" of subspaces of L_i of L:

$$\cdots \supseteq L_{-2} \supseteq L_{-1} \supseteq L_0 \supseteq L_1 \supseteq L_2 \supseteq \cdots .$$

This filtration is completely determined by L_{-1}/L_0, which can be taken to be an arbitrary irreducible L_0-submodule of L/L_0. Considering the direct sum of the factor spaces $A_i = L_i/L_{i+1}$, one obtains a graded Lie algebra A, which Block and Wilson show satisfies the hypothesis of Kac's theorem. In the present case his theorem implies that A is either classical or of Cartan type, thereby reducing the classification theorem to showing that L and A are isomorphic (more precisely, since A need not be simple, that L is isomorphic to the second derived algebra $A^{(2)}$ of A). This follows directly from Kac's theorem when A is classical. However, the proof is quite elaborate when A is "sporadic"—i.e., of Cartan type. The strategy is to plug into a general machine designed to determine all simple Lie algebras having a given associated graded algebra.

First, A is of one of four possible types: $X = W, S, H,$ or K (whose definitions we omit). Moreover, there is a natural embedding of A into a suitable subalgebra $X(n)$ of the infinite-dimensional (nonrestricted) so-called Jacobson-Witt algebra $W(n)$, where $n = \dim A$, which in turn yields an embedding of L into $W(n)$ compatible with the given filtration of L. On the other hand, a theorem of Blattner [30] produces a second embedding of L into $W(n)$ that depends only on the existence of L_0 and that likewise preserves the given filtration of L.

Blattner's results also yield that the two filtration-preserving embeddings of L into $W(n)$ are necessarily conjugate under some automorphism θ of $W(n)$, which implies that L is isomorphic to a subalgebra of $\theta(X(n)) \cap W(n,1)$, where $W(n,1)$ is a specified subalgebra of $W(n)$. In the restricted case, Kac has shown under these conditions that θ must, in fact, be the identity automorphism, in which case $L \subseteq X(n,1) = X(n) \cap W(n,1)$, and now it is easily verified that L is uniquely determined by X and n: namely, $L \cong X(n,1)^{(2)}$.

Thus again there is an external realization of the generic simple Lie algebra L, but it is reached only after a far lengthier internal analysis than in the classical case. Moreover, the external phase of the analysis, beginning with the construction of the associated graded Lie algebra A, is likewise much more intricate than in the classical case. If one includes the proofs of the various supporting theorems as well as the rank $\leqslant 2$ classification, Block and Wilson's total argument certainly approaches 500 pages.

Furthermore, however impressive their result, a complete classification of finite-dimensional simple Lie algebras is still a long way off, requiring extensions of their arguments to nonrestricted algebras of arbitrary characteristic. Moreover, the desired result will inevitably be more complicated since simple algebras exist of characteristic 2, 3, or 5 that are of neither classical nor Cartan type. In addition, if the classification of the simple groups is any guide, the case $p = 2$ promises to be especially difficult. Thus we can reasonably anticipate both the internal Lie-theoretic and external graded Lie algebra phases of the analysis to be of very substantial length.

What then of the finite simple groups, which possess neither root decompositions nor associated graded groups and, in addition, come with no prespecified characteristic? Yet with *no premeditated planning* their classification followed a basic pattern very analogous to the Block-Wilson proof.

1. "Low-rank" groups required separate analysis, a necessary prerequisite before attacking the generic simple group G.

2. The first phase of the generic case analysis again involved an internal determination of the structure of the centralizer of a suitable "semisimple" element of G.

3. The bulk of the second phase was likewise internal: the given centralizer information was used to force the exact shape of other portions of the subgroup structure of G.

4. The total subgroup information was used finally to show that G possessed a uniquely determined multiplication table.

However, the analysis was considerably more complex than in the Lie algebra case, for it was necessary to develop ab initio a number of new techniques to serve as substitutes for the lack of root decompositions in G and the inability to associate a graded group with the "coset space" of the centralizer of a semisimple element of G. In addition, we had to develop an internal criterion for deciding whether G had "even" or "odd" characteristic. Only at the very last step were we able to achieve a geometric realization for G, equivalently expressed in terms of a presentation by generators and relations.

Moreover, in part because of the existence of the sporadic solutions, the treatment of the low-rank cases was exceedingly elaborate, requiring, in particular, a division into "small" and "intermediate" subcases. In the end, the sporadic groups all emerged in one of the following two guises:

SPORADIC SOLUTIONS OF LOW RANK CASES

Internal odd characteristic: J_1, Mc, ON, $.3$, $M(23)$, Ly.

Internal characteristic 2: M_{11}, M_{12}, M_{22}, M_{23}, M_{24}, J_2, J_3, J_4,

HS, He, Ru, Suz, $.2$, $.1$, $M(22)$, $M(24)'$, F_5, F_3, F_2, F_1.

There you have the basis for my belief that the classification of the finite simple groups is an inherently difficult problem. Moreover, if some 100 years after the classification of the finite-dimensional complex Lie algebras, the proof is still almost entirely internal, in my opinion it is an illusion to imagine that it will ever be possible to associate a geometry with either finite-dimensional simple Lie algebras or finite simple groups without preparing the way by

an extensive process of *internally distinguishing simplicity*. In the simple group case, this process has three basic phases which I shall suggestively refer to as (a) *internal elimination of* Rad(G) (the unique largest solvable normal subgroup of G), (b) *internal reduction to the simplicity of* Soc(G) (the product of all minimal normal subgroups of G), and (c) *internal construction of* Soc(G). The only meaningful question is what is the minimum length required to carry out the process, not its necessity.

The next three chapters should therefore be viewed as an elaboration of these various points relative to both the existing and other possible approaches to the classification of the finite simple groups.

Chapter II. The Evolution of the Proof

The estimated 500 articles encompassing the proof of the classification theorem represent no more than a quarter of the papers written on simple groups. This is not surprising since it was not until the very end of the period that a global strategy was formulated—before then there existed at best a sense of direction.

As a consequence, during much of the period, various paths were explored without any certainty of·their ultimate relationship to the classification. Some did produce major advances, others completely failed, while some were pursued for their own sakes, and still others were bypassed by the ensuing developments.

If the Mathieu groups were discovered as highly transitive permutation groups and the Conway groups from automorphisms of the Leech lattice, what could be more natural and important than to study multiply transitive permutation groups and automorphism groups of other integral lattices? Unfortunately, no further simple groups were discovered in either direction. Similarly, if groups with abelian or dihedral Sylow 2-subgroups could be classified, why not consider other types as well? Indeed, for a time it appeared we were heading towards a characterization of every simple group or family of groups in terms of the structure of its Sylow 2-subgroups—a result which, in its general form, was not needed for the final proof. And so it went....

Here I shall focus on the evolutionary nature of the mainstream of developments. This will provide an appropriate perspective for the discussion of Chapters III and IV.

1. Discovery versus classification. Obviously one cannot even state the full classification theorem without a complete list of the simple groups, so that the existence of the simple groups constitutes *logically* a separate chapter of simple group theory that precedes their classification.

However, *historically*, there was considerable interplay between discovery of new simple groups and classification results. For example, Suzuki's family $Sz(2^n)$ emerged in the course of an attempt to determine all simple groups in which the centralizer of every involution was a 2-group [131]. His conjectured solution—the groups $PSL_2(q)$ for suitable q, plus the single group $PSL_3(4)$—was obtained by checking the list of then-known simple groups. Arguing by contradiction, he determined the *internal* structure of a minimal counterexample G to his proposed theorem: A Sylow 2-subgroup U of G (of

order 2^{2n} for some odd n) had a uniquely determined structure as did its
normalizer B in G (of order $2^{2n}(2^n - 1)$). [A theorem of G. Higman [86] which
gives a precise description of all 2-groups admitting an automorphism that
transitively permutes the involutions of the group played a crucial role in
forcing the structure of B.] Moreover, the permutation representation of G
on the cosets of B was doubly transitive of degree $2^{2n} + 1$ (whence $|G| =
2^{2n}(2^n - 1)(2^{2n} + 1)$) with only the identity fixing three cosets. This was the
evidence phase of his analysis.

At this point, Suzuki could see that G had an internal structure similar to
that of a group of Lie type of Lie rank 1, but had no idea whether such a group
G existed for any allowed value of n. In particular, there was then no known
group of Lie type with such an internal structure. Eventually Suzuki in a
brilliant insight discovered a group of 4×4 matrices defined over $GF(2^n)$ that
satisfied all the conditions of his group G. Thus his actual construction of the
groups $Sz(2^n)$ was carried out *externally*. Moreover, it was this matrix repre-
sentation that enabled Ree to connect Suzuki's groups with the family $^2B_2(2^n)$
of groups of Lie type and thereby led him to the Lie-theoretic construction of
the families $^2G_2(3^n)$ and $^2F_4(2^n)$.

Much of the time, the discovery of new simple groups was keeping pace with
the new classification theorems being proved. Moreover, the linkage between
the two developments cut in both directions. For example, the techniques
Fischer introduced in his analysis of 3-transposition groups led eventually to
the classification of groups of so-called "$GF(2)$-type," a very important class of
simple groups [see §4 below].

Another aspect of the historical development was that throughout most of
the period simple group theorists were convinced that there were still many
simple groups out there, waiting to be discovered, so that they were never quite
certain whether a particular tight configuration of subgroups (such as in
Suzuki's situation above) would lead ultimately to a contradiction or to one or
more new simple groups. Some simple group theorists, myself included, who
were primarily interested in general classification theorems, at times attempted
to avoid the problem of undetected sporadic groups by proceeding *axiomati-
cally*, so that our arguments would remain valid provided only that any
yet-to-be-discovered simple group satisfied the given set of axioms. On the
other side, there were those who spent great effort searching for new sporadic
groups similar in structure or construction to those already existing.

This uncertainty over the completeness of the list of simple groups persisted
until about 1980. It was only then with the determination of the simple groups
of $GF(2)$-type that finite group theorists at last began to feel that all sporadic
groups had very likely been found.

No matter how intertwined existence and classification may have been, it is
preferable to regard them as two independent parts of simple group theory.
Thus *classification* shall refer solely to the problem of proving that an existing
list of simple groups is *complete*. Only the issue of actual construction is
excluded from classification. This is a natural division, since classification
appears to be an internal problem, while existence is an external one.

2. Local group-theoretic analysis. As has been stressed, the existing classification proof is internal, focusing almost completely on the proper subgroup structure of the simple group G under investigation. Some of the results making up the total proof have *noninductive* hypotheses, while others are *inductive*.

As an example of the latter, the odd order theorem is inductive, since the hypothesis carries over to subgroups and homomorphic images. Thus Feit and Thompson began with a *minimal* counterexample G—i.e., a counterexample of least order—concluding at once that G is simple (of odd order) and each of its proper subgroups is solvable. To establish the theorem, it was necessary for them to derive a contradiction *from this information alone*. Every inductive classification theorem leads to similar initial general structural assumptions about proper subgroups of a minimal counterexample.

In noninductive problems, the hypothesis typically specifies that some proper subgroup H of G, such as the centralizer of an involution, has a structure similar (perhaps even identical) to that of some known simple group G^* (in some instances, similar to that of *no* known simple group). From this initial information about H, one proceeds by a process of "analytic continuation" to obtain structural information about other proper subgroups, ultimately forcing a sufficient portion of G's internal structure to so "closely resemble" that of G^* that one can then go on to prove that $G \cong G^*$ (or to derive a contradiction when no such G^* exists). This is the meaning of the second step in internally distinguishing simplicity.

In inductive problems, the initial assumptions are, on the one hand, *global*, since they apply to all proper subgroups of a minimal counterexample G, but on the other, *loose* since they specify only the general shape of these subgroups. As a result, the first objective is to force the internal structure of G to satisfy precisely the kind of conditions imposed in the noninductive cases, and this is the meaning of the first step in internally distinguishing simplicity.

Furthermore, almost all the techniques developed over the years to study simple groups were designed to enable us to make assertions about the subgroup structure of G, thus facilitating the process of internally distinguishing simplicity. Primarily, these concerned only the *local* subgroups of G. Here by definition the normalizer in G of a nontrivial *solvable* subgroup of G is called a *local* subgroup. Furthermore, the study of a simple group's local structure, including the underlying body of techniques, came to be called *local group-theoretic analysis*.

The term provides a more explicit description of the internal approach to the classification: the existing proof consists of an analysis of the structure and embedding of the local subgroups of the simple group under investigation. This statement characterizes the essential nature of the classification theorem.

3. The Brauer principle. The first significant general inroad into the classification problem was made by Richard Brauer, who early on came to the following realization:

If the centralizer of an involution of an abstract simple group G closely

*approximates that of an involution in a known simple group G^*, then methods can
be developed for proving that $G \cong G^*$.*

The statement must be read somewhat loosely, since in a few cases distinct
simple groups have identical centralizers of involutions—e.g., the groups
$PSL_3(3)$ and M_{11} and the groups $PSL_5(2)$, M_{24}, and *He*.

Using the block theory methods he had developed in the 1940s and 1950s,
Brauer first validated this principle for the family $PSL_3(q)$, q odd, characteriz-
ing it in terms of the structure of the centralizers of its involutions [34].
Brauer's objective was geometric: to construct a projective 2-space π on which
G acted as a group of projective transformations, in which case he could
"identify" G as $PSL_3(q)$. However, his geometry was defined *internally* in
terms of the set of involutions of G and their centralizers and related
subgroups. Properties of the geometry were then established by a combination
of character theory and internal analysis. Moreover, he encountered an excep-
tional configuration in the $PSL_3(3)$ case, which gave rise to the M_{11} solution.

As the *Brauer principle*, as I shall refer to it, came to be verified in more and
more cases, his geometric approach was superseded, except for some of the
sporadic groups, by a *generator-relations* approach. In particular, for char-
acterizations of the groups of Lie type, the aim of the analysis was to force the
group G under investigation to satisfy the precise conditions of the so-called
Steinberg presentation [**127, 128**] (cf. also [**76**])—equivalently, to force G to
have the structure of a (B, N)-*pair* (Tits [**145**]; cf. also [**76**]). However, Brauer's
geometric approach remained the more effective for many of the sporadic
group characterizations, in which cases the aim was to construct a graph π on
which G acted as a group of automorphisms and then to argue that π was
uniquely determined, thereby forcing the isomorphism type of G to be unique.

In reality, the geometric and generator-relation approaches are
equivalent—one can pass directly back and forth between them. The critical
point is that one must force the structure of the abstract group G to be
uniquely determined; the particular description simply provides the means by
which one "recognizes" G as a specific known simple group.

Brauer's discovery was fundamental, for it provided an operative, inductive
procedure to follow in every general classification theorem:

(A) *Force the centralizer C_x of an involution x in the simple group G under
investigation to closely resemble that in some known simple group.*

(B) *Determine the isomorphism type of G by validating the Brauer principle for
the centralizer C_x.*

This division came to dominate the entire classification proof. In fact, in
time the Brauer principle was broadened to include centralizers of elements of
odd prime order.

(A) and (B) relate to the internally distinguishing simplicity process in the
following way: (A) covers the internal elimination of Rad(G) as well as the
internal reduction to a simple Soc(G), while (B) covers the internal construc-
tion of Soc(G), including the uniqueness of its multiplication table.

In groups of low Lie rank such as $PSL_3(q)$, q odd, all involutions are
conjugate, so Brauer had no option on which class to base his analysis.
However, the number of conjugacy classes in the groups of Lie type increases

with the Lie rank. It was gradually realized that validation of the Brauer principle worked most smoothly for centralizers C_x of involutions x whose order was as large as possible and which contained a normal quasisimple subgroup. The customary term in this case is *standard form* with *standard component K* (the general notion due to Aschbacher [6]). However, for emphasis, I shall use the more suggestive terms *Brauer form* and *Brauer component*, including in the definition the case in which x has arbitrary prime order p.

The typical shape of a centralizer C_x in Brauer form with $|x| = p$ is as follows:

(1) The Brauer component K of C_x has order divisible by p; and

(2) $H = C_{C_x}(K)$ has cyclic (or for $p = 2$ quaternion) Sylow 2-subgroups.

To illustrate the existence of Brauer form, take $G = GL_n(q)$, $n \geqslant 3$ (even though G is not simple, the essential structure remains visible, and it is easier to make the calculations than in the corresponding simple group $PSL_n(q)$). Any element x of prime order p in G *coprime* to q can be diagonalized over some finite extension of the underlying field $GF(q)$. For simplicity, we assume this already occurs over $GF(q)$. Thus, up to a permutation of coordinates,

$$
x = \begin{pmatrix}
\alpha_1 \overset{k_1}{\underset{\alpha_1}{\searcurrow}} & & & & 0 \\
 & \alpha_2 \overset{k_2}{\underset{\alpha_2}{\searcurrow}} & & & \\
 & & \ddots & & \\
0 & & & & \alpha_r \overset{k_r}{\underset{\alpha_r}{\searcurrow}}
\end{pmatrix}
$$

where $\alpha_1, \alpha_2, \ldots, \alpha_r$ are distinct nonzero elements of $GF(q)$, and

$$
C_x = \left\{ \begin{pmatrix}
A_1 & & & & \\
 & A_2 & & & \\
 & & \ddots & & \\
 & & & \ddots & \\
 & & & & A_r
\end{pmatrix} \right\}
$$

where $A_1, A_2, \ldots, A_r$ are arbitrary nonsingular $k_1 \times k_1$, $k_2 \times k_2, \ldots$ matrices.

Direct calculation will yield that C_x consists of block matrices, as described above. We see then that

$$
C_x \cong GL_{k_1}(q) \times GL_{k_2}(q) \times \cdots \times GL_{k_r}(q).
$$

To obtain an element of order p with centralizer in Brauer form, we specialize to the case $k_1 = n - 1$ and $k_2 = 1$, whence

$$
C_x \cong GL_{n-1}(q) \times GL_1(q).
$$

Assuming $n \geqslant 4$ or $n = 3$ and $q > 3$, the $SL_{n-1}(q)$ subgroup K of $GL_{n-1}(q)$ will be quasisimple and normal in C_x; moreover, the second factor $GL_1(q)$ is cyclic, isomorphic to the multiplicative group of the field $GF(q)$. Thus indeed

C_x has the general shape of the Brauer form. To obtain its precise form, one would have to cut down to $C_x \cap SL_{n-1}(q)$ and then factor out by the central subgroup of scalar matrices.

4. Minimal simple groups. In the next two sections, I shall illustrate the evolutionary development of the classification proof with a number of examples describing the more or less piecemeal manner in which major chapters of simple group theory grew out of particular aspects of specific classification theorems.

It will be best to begin with some general terminology that will be needed for this discussion as well as in later chapters.

For any group X and prime p, X is a *p-group* (*p'-group*) if X has order a power of p (order relatively prime to p).

$O_p(X) =$ unique largest normal p-subgroup of X.

$O_{p'}(X) =$ unique largest normal p'-subgroup of X, called the *p'-core* of X.

$O(X) = O_{2'}(X)$, called the *core* of X.

A subgroup Y of X is called a *p-local*, *p'-local*, or *O-local* subgroup of X if $Y = N_X(W)$ for some nontrivial p-subgroup, p'-subgroup, or subgroup W of odd order, respectively. Thus O-local subgroup $=$ 2'-local subgroup; moreover, p-local and O-local subgroups are local subgroups (as W is then solvable).

If X is an abelian p-group, $m(X) = rank$ of $X =$ minimum number of generators of X.

If X is a p-group, $m(X) = rank$ of $X = \max\{m(A)|\ A$ an abelian subgroup of $X\}$.

$n(X) = normal\ rank$ of $X = \max\{m(A)|\ A$ an abelian normal subgroup of $X\}$.

$r(X) = sectional\ rank$ of $X = \max\{m(A)|\ A$ an abelian section of $X\}$. [A homomorphic image of a subgroup of any group X is called a *section* of X.]

For arbitrary groups X, $m_p(X)$, $n_p(X)$, $r_p(X) = p$-*rank*, *normal p-rank*, *sectional p-rank* of X, respectively, $= m(P)$, $n(P)$, $r(P)$, respectively, where P is a Sylow p-subgroup of X.

A 2-group X is (a) *dihedral*, (b) *quaternion*, or (c) *quasidihedral* if X is generated by elements x, y subject to the relations $x^{2^n} = 1$ and correspondingly

(a) $y^2 = 1$, $x^{-1}yx = y^{-1}$, $n \geq 1$;

(b) $y^4 = 1$, $x^{-1}yx = y^{-1}$, $y^2 = x^{2^{n-1}}$, $n \geq 2$; or

(c) $y^2 = 1$, $x^{-1}yx = y^{-1+2^{n-1}}$, $n \geq 3$.

A 2-group X is *wreathed* if X is generated by elements x, y, z subject to the relations $x^{2^n} = y^{2^n} = z^2 = 1$, $[x, y] = 1$, $z^{-1}xz = y$, $n \geq 2$.

If X is a p-group, then X is *extra-special* if $Z(X) = [X, X]$ is of order p and $X/Z(X)$ is elementary abelian (i.e., the direct product of subgroups of order p). Moreover X is said to be *of symplectic type* if X is the product of two subgroups X_1, X_2 with X_1 extra-special, $Z(X_1) \leq Z(X_2)$, $[X_1, X_2] = 1$, and either X_2 is cyclic or $p = 2$ and X_2 is dihedral, quaternion or quasidihedral of order ≥ 16. For completeness, we include the possibility $X_1 = 1$, $X = X_2$ in the definition of symplectic type.

The significance of this last class of p-groups arises from the following result of P. Hall [**74**, Theorem 5.4.9].

PROPOSITION. *If a p-group X contains no noncyclic characteristic abelian subgroups, then X is of symplectic type.*

Finally, for arbitrary X, $X^{\#} = X - \{1\} =$ set of nonidentity elements of X. Also, for the ambient simple group G under investigation, we write, for simplicity, C_Y, N_Y for $C_G(Y)$, $N_G(Y)$, respectively, for any subgroup or subset Y of G.

Our first set of examples is connected with Thompson's fundamental classification of the *minimal* simple groups [110]—simple groups of *even* order each of whose proper subgroups are solvable. Once the odd order theorem had been completed, it was natural for Thompson to consider this family of simple groups. The group A_5, of order 60—the simple group of smallest order—is clearly an example of a minimal simple group.

Given the local methods that Thompson had used in the odd order theorem, it soon became apparent to him that he did not require this solvability assumption on *all* subgroups of G, but only on its *local* subgroups. Thompson dubbed a group satisfying this weaker condition an *N-group* and his goal was to determine the isomorphism types of all simple N-groups.

I shall describe five obstructions (out of a considerably greater number) that Thompson was forced to overcome to establish his celebrated classification theorem. As we shall see, each of these subsequently developed into a major chapter of simple group theory.

A. *The 2-rank* $\leqslant 2$ *case.* In the odd order proof, Thompson was forced into quite distinct lines of argument to pin down the "p-structure" of G according as $m_p(G) \leqslant 2$ or $m_p(G) \geqslant 3$ for a given odd prime p. This rank distinction persisted in the N-group case. In particular, Thompson soon found that he had to single out the $m_2(G) \leqslant 2$ case for separate treatment.

It is an elementary fact that a 2-group of rank 1 is either cyclic or quaternion. It follows therefore from the odd theorem for the 2-rank 0 case, and from a result of Brauer and Suzuki [37], for the quaternion case, (plus an elementary transfer result of Burnside for the cyclic case) that there are no (nonabelian) simple groups of 2-rank $\leqslant 1$, so it was the case $m_2(G) = 2$ that concerned Thompson.

Using so-called *fusion-theoretic* methods, Thompson's objective was to force the possible isomorphism types of a Sylow 2-subgroup of G.

Here if p is a prime and P a Sylow p-subgroup of G, *p-fusion* refers to the conjugacy in G of distinct elements or subsets of P. There already existed a number of classical transfer results, giving conditions on p-fusion for a group to possess a normal subgroup of index p. Thompson added his own elementary but important transfer result, applicable solely to the prime 2:

LEMMA. *If S is a Sylow 2-subgroup of a group G, T a subgroup of index 2 in S, and x an involution of S − T, then either G has a normal subgroup of index 2 (not containing x) or else x is G-conjugate to an involution of T.*

At about this time, Glauberman, extending the Brauer-Suzuki character-theoretic quaternion argument, established his fundamental Z^**-theorem* [62], a result that was destined to become an essential tool throughout the

entire classification proof, and which subsequently went through many gener-
alizations. Again it applies only to the prime 2; it derives its name from the
term $Z^*(G)$, which denotes the preimage in G of the center of the group
$G/O(G)$.

THEOREM. *If S is a Sylow 2-subgroup of a group G and x an involution of S,
then either $x \in Z^*(G)$ or else x is G-conjugate to an involution of $S - \langle x \rangle$.*

Thompson transfer works "at the top" of G, while the Z^*-theorem focuses
on its "bottom." Since Thompson's N-group G was assumed to be simple,
these two results could be used in conjunction to force a certain amount of
G-fusion to take place within a given Sylow 2-subgroup S of G. Since
$m(S) = 2$, this placed severe restrictions on the subgroups S, ultimately
leading to a complete determination of their possible isomorphism types
(homocyclic abelian, dihedral, quasidihedral, wreathed, or of order 64 and
isomorphic to a Sylow 2-subgroup of $PSU_3(4)$).

In most of these cases, Thompson could then invoke existing classification
results to obtain the possible isomorphism types of G.

B. *Core obstruction*. The remaining obstructions deal with the $m_2(G) \geq 3$
case, where the expected answers for G were certain groups of Lie type of
characteristic 2: namely, the families $PSL_2(2^n)$ and $Sz(2^n)$, plus a few individ-
ual groups.

Borel and Tits [33] (some time later) established a basic general property of
the p-local subgroups H of a group of Lie type of characteristic p: namely, for
any such subgroup H,

$$(1) \qquad\qquad C_H\big(O_p(H)\big) \leqslant O_p(H).$$

In particular, as $O_{p'}(H)$ centralizes $O_p(H)$, it follows that

$$(2) \qquad\qquad O_{p'}(H) = 1.$$

In time, an arbitrary finite group each of whose p-local subgroups satisfied
condition (1) came to be called a group of *characteristic p type*, the groups of
Lie type of characteristic p thus being the prototypical examples of such
groups.

Thompson thus knew that as a first step in forcing the ultimate isomorphism
type of his simple N-group G, he would have to argue that G was of
characteristic 2 type. In solvable groups, it is easily shown that conditions (1)
and (2) are equivalent, so in N-groups the characteristic 2 type conclusion
reduced to the assertion

$$(3) \qquad\qquad O(H) = 1 \text{ for every 2-local subgroup } H \text{ of } G.$$

It can be easily shown that to verify (3), it suffices to prove the weaker
statement

$$(4) \qquad\qquad O(C_x) = 1 \text{ for every involution } x \text{ of } G.$$

Hence the existence of nontrivial cores in the centralizers of involutions
presented a second obstruction to Thompson—what I shall refer to here as
core obstruction.

Removal of this core obstruction is the precise formulation for N-groups for internally eliminating the $2'$-part $O(G)$ of $\mathrm{Rad}(G)$.

C. *The existence of large subgroups.* When (4) fails, Thompson proved that G contains an O-local subgroup M with the following properties:

(5a) $S \leqslant M$ for some Sylow 2-subgroup S of G; and

(5b) $N_T \leqslant M$ for every $1 \neq T \leqslant S$.

Note that (5b) is equivalent to the assertion that M intersects its distinct G-conjugates in subgroups of odd order.

Since M contains the normalizers in G of all nontrivial subgroups of S, M is a "large" subgroup of G. Thus it is $O(M)$ (nontrivial, as M is an O-local subgroup) that was Thompson's candidate for the $2'$-part of $\mathrm{Rad}(G)$, and he had to derive a contradiction from the existence of such an O-local subgroup M to complete the removal of core obstruction.

The standard term for a subgroup M of a group G satisfying the conditions of (5) is *strongly embedded*, so that Thompson's third obstruction came from the existence of strongly embedded O-local subgroups within his group G.

D. *The normal 2-rank 2 case.* Thompson's construction of M turned out to be especially difficult when a Sylow 2-subgroup S was "small," for his arguments required a certain degree of freedom within S, which was absent when $n(S) = 2$. Thus the case $m_2(G) \geqslant 3$, $n_2(G) = 2$ provided a fourth obstruction.

E. *The O_2 symplectic type case.* The bulk of Thompson's analysis occurs after he has succeeded in proving that his N-group G is of characteristic 2 type. Here we mention only one of the obstructions he encountered.

A substantial portion of that analysis focuses on a certain maximal 2-local subgroup M of G constructed by Thompson. In the general case, M is *strongly p-embedded* in G for suitable *odd* primes p (here the term strongly p-embedded is defined by direct analogy with that for the prime 2, taking S now to be a Sylow p-subgroup of G). The ensuing delicate argument concerns nontrivial elementary abelian normal 2-subgroups A of M and the possible embeddings in M of G-conjugates of subgroups of A (this has come to be called Thompson's *weak closure* analysis).

However, the argument collapses when $|A| = 2$, so that the case in which M contains no such elementary abelian normal subgroup of order $\geqslant 4$ is exceptional. However, in the latter case $O_2(M)$ cannot contain a noncyclic abelian characteristic subgroup B, otherwise the involutions of B would generate an elementary abelian normal (in fact, characteristic) subgroup A of M of order $\geqslant 4$. Hence, by Philip Hall's result, $O_2(M)$ must be of symplectic type! Furthermore, if we then denote the central involution of $O_2(M)$ by x, it follows that $M = C_x$ (since M is a maximal 2-local subgroup of G and $M \leqslant C_x$). Thus a fifth obstruction encountered by Thompson was the case in which $O_2(C_x)$ is of symplectic type for some involution x of G.

We shall now describe the five chapters of simple group theory that over time emerged from the above five obstructions.

A*. *Groups of 2-rank 2.* Work on simple groups of 2-rank 2 had begun well before Thompson's N-group theorem. Indeed, the groups $PSL_3(q)$, q odd, have quasidihedral or wreathed Sylow 2-subgroups according as $q \equiv -1 \pmod 4$

or $q \equiv 1 \pmod 4$. Thus Brauer's first classification theorems dealt with the 2-rank 2 case. Moreover, John Walter and I had begun our investigation of arbitrary simple groups with dihedral Sylow 2-subgroups [82] while Feit and Thompson were still occupied with groups of odd order.

Thus it was clear all along that the family of simple groups of 2-rank 2 was of intrinsic interest. Nevertheless, it was the N-group paper that served to put this family in perspective within the context of broader classification problems, at the same time strongly indicating that the classification of such simple groups would require its own special methods.

In the mid 1960s Alperin established a basic fusion result [1] (cf. also [74, Theorem 7.2.6]), which in essence asserted that for any group G, prime p, and Sylow p-subgroup P of G, G-fusion of any two subsets of P could be described "p-locally" via a product of conjugations occurring in the G-normalizers of a well-specified collection $\mathscr{D}$ of nonidentity subgroups of P—such a collection $\mathscr{D}$ came to be called a *conjugation family* for P (with respect to G).

As an application of his theorem, and using the fusion results described above, Alperin showed that Thompson's rank 2 N-group result held in arbitrary rank 2 simple groups [3] (cf. also [76]).

THEOREM. *If G is a simple group of 2-rank 2, then a Sylow 2-subgroup of G is either homocyclic abelian, dihedral, quasidihedral, wreathed, or of order 64 and isomorphic to a Sylow 2-subgroup of $PSU_3(4)$.*

Ultimately Walter and I handled the dihedral case, Alperin, Brauer, and I the quasidihedral and wreathed cases [2, 3], Brauer the abelian of order $\geqslant 8$ case [128], and Lyons the $PSU_3(4)$ case [102], the final theorem reading as follows.

THEOREM. *If G is a simple group of 2-rank 2, then $G \cong PSL_2(q)$, $PSL_3(q)$, $PSU_3(q)$, q odd, $PSU_3(4)$, A_7, or M_{11}.*

B*. *The signalizer functor method.* In the N-group $m_2(G) \geqslant 3$ case, Thompson's construction of the strongly embedded subgroup M in the face of core obstruction was quite cumbersome, and his method looked as though it would be very difficult to generalize. However, it was clear at the time that core obstruction would be a major problem in all classification theorems (this was certainly true in the 2-rank 2 classification theorem), for the cores of centralizers of involutions in the known simple groups have a very simple structure:

PROPOSITION. *If G is any known simple group and x an involution of G, then $O(C_x)$ is cyclic.*

Thus, if one were ever to reach Brauer form for the centralizer of some involution in a given classification theorem, one would certainly have to establish some strong results about cores of centralizers of involutions in arbitrary simple groups—not necessarily that they had to be trivial (as in the N-group case), but at least that they had a very restricted structure.

In the late 1960s, I discovered an abstract formulation of the problem of core obstruction and simultaneously developed a general method for attacking it. Underlying my approach was the notion of a *signalizer functor* and the

related, basic *solvable signalizer functor* theorem [82]. My ideas and original proof of the theorem were later greatly improved by Goldschmidt [67, 68] and others—in particular, extending the result to arbitrary primes.

Let A be an elementary abelian p-subgroup of a group G, p a prime, and suppose that for each $a \in A^{\#}$, there is associated an A-invariant subgroup $\theta(a)$ of C_a of order prime to p with the key property

$$\theta(C_a) \cap C_{a'} = \theta(C_{a'}) \cap C_a$$

for every pair $a, a' \in A^{\#}$. Then θ is called an *A-signalizer functor on G.* [Note that the abelian condition is required to insure that $A \leqslant C_a$ for each $a \in A^{\#}$.] θ is said to be *solvable* if each $\theta(a)$ is a solvable group. For $p = 2$, the odd order theorem guarantees that every A-signalizer functor θ is solvable.

Furthermore, taking $\theta(a) = 1$ for all $a \in A^{\#}$ clearly yields an A-signalizer functor—the so-called *trivial* functor. Finally, for a given A-signalizer functor θ, the subgroup $\theta(G; A)$ generated by the groups $\theta(a)$, a ranging over $A^{\#}$, is called the *closure* of θ in G.

The solvable signalizer functor theorem asserts the following:

THEOREM. *Let p be a prime and A an elementary abelian p-subgroup of rank $\geqslant 3$ of a group G. If θ is a solvable A-signalizer functor on G, then the closure $\theta(G; A)$ of θ is a p'-group.*

In particular, the theorem shows that $\theta(G; A)$ is a proper subgroup of G.

The theorem provides a further illustration of the "rank $\geqslant 3$ phenomenon." Moreover, it is, in general, false in rank 2, which helps to explain why special methods were required to study simple groups of p-rank $\leqslant 2$.

For a suitable choice of elementary abelian p-group A and A-signalizer functor θ, it is the subgroup $\theta(G; A)$ that is our internal candidate for the p'-part $O_{p'}(\mathrm{Rad}(G))$ of $\mathrm{Rad}(G)$. Thus the precise operational meaning of internally elminating $\mathrm{Rad}(G)$ is: proving the *triviality* of A-signalizer functors for certain elementary abelian p-subgroups A of G, p ranging over suitable prime divisors of $|G|$.

On the other hand, when θ is nontrivial, it is the subgroup $N_{\theta(G; A)}$ that is our internal candidate for a large subgroup of G. The essential point of the signalizer functor theorem is that it guarantees that $\theta(G; A)$ is *proper*. Since G is simple and $\theta(G; A)$ is assumed nontrivial, it thus yields the crucial conclusion that $N_{\theta(G; A)}$ is a proper subgroup of G, which is a necessary first step in establishing its largeness.

C*. *Groups with a strongly embedded subgroup.* The work of Suzuki on doubly transitive permutation groups culminated in Bender's complete classification of groups with a strongly embedded subgroup [25].

THEOREM. *If G is a simple group containing a strongly embedded subgroup, then $G \cong PSL_2(2^n)$, $PSU_3(2^n)$, or $Sz(2^n)$ for some n.*

One checks in each case that the strongly embedded subgroup M of G is not an O-local subgroup, so Bender's theorem implies that simple groups never contain large O-local subgroups, a crucial step in the internal elimination of the $2'$-part of $\mathrm{Rad}(G)$.

But the theorem is applicable more generally, insuring that a minimal counterexample to any proposed classification theorem cannot contain a strongly embedded subgroup of any kind, a fact which is also important for the internal construction of Soc(G).

We note that Bender's theorem is a characteristic 2 result, in contrast to the 2-rank 2 classification theorem, which is essentially an odd characteristic result.

Subsequently Aschbacher [4] extended Bender's theorem, classifying groups in which the corresponding subgroup M is assumed to satisfy only the weaker condition

$$N_T \leqslant M \quad \text{for all } T \leqslant S \quad \text{with } m(T) \geqslant 2.$$

The subgroup generated by the set of such N_T is called the 2-*generated core* of G (relative to the given Sylow 2-subgroup S of G).

Aschbacher's result effectively enabled one to define large subgroups in terms of proper 2-generated cores rather than strong embedding. This was a decided advantage, for direct application of the signalizer functor method (for the prime 2) produces O-local subgroups containing only the 2-generated core.

D*. *Groups of sectional 2-rank* $\leqslant 4$. As in the N-group case, removal of core obstruction in other classification problems required the normal 2-rank 2 case to be treated separately. Unfortunately, this condition does not carry over, in general, to subgroups and homomorphic images, so that one could not use induction to classify such groups.

Thompson hoped it would be possible to proceed as in the 2-rank 2 case, first forcing the possible structures of a Sylow 2-subgroup S of an arbitrary simple group G with $n_2(G) = 2$ by an analysis of 2-fusion, and then determining all simple groups with such Sylow 2-subgroups. However, the problem turned out to be less tractable than anticipated, and his student MacWilliams was forced to settle for strong partial results [85]. Among these was the following general fact:

THEOREM. *If S is a 2-group with $n(S) = 2$, then $r(S) \leqslant 4$.*

However, sectional 2-rank *is* a condition that carries over to subgroups and homomorphic images, so that it is possible to proceed inductively to classify the broader family of simple groups of sectional 2-rank $\leqslant 4$, and thereby determine all simple groups of normal 2-rank 2 as a corollary. Harada and I were the ones who carried this out [78]. Because we now had complete general information about the proper subgroup structure of a minimal counterexample G to our proposed theorem, we were able to pin down, by a combination of 2-fusion and 2-local analysis, the possible isomorphism types of Sylow 2-subgroups S of G, and then proceed to classify all such simple groups with such Sylow 2-subgroups (some of these specified Sylow 2-group cases were treated by others), thereby showing that no counterexample existed. However, in total, it was a painful undertaking, involving very detailed, technical analysis, running to some 800 journal pages.

The effort did, however, have significant further payoff, for the sectional 2-rank $\leqslant 4$ theorem was repeatedly invoked in subsequent classification theorems to dispose of unpleasant low 2-rank configurations. The result would never have been available if group theorists had somehow found an easy, direct way of settling the normal 2-rank 2 problem.

E*. *Groups of GF(2)-type.* The N-group paper also served to focus attention on the class of simple group G containing an involution x with $O_2(C_x)$ of symplectic type. This condition is, in fact, a characteristic property of many of the families of groups of Lie type defined over the *prime field GF(2)*; and for this reason such groups came to be called *GF(2)-type.*

A more exciting development was the fact that many of the sporadic groups then being discovered were also turning out to be of *GF(2)-type*—ultimately sixteen of the twenty-six sporadic groups had this property. Thus there was a strong incentive for finding all simple groups of *GF(2)-type*. The principal architect of the resulting classification theorem was Timmesfeld [118], who pushed Fischer's transposition theory far beyond the original 3-transposition case. However, several others were involved in the complete *GF(2)-type* classification theorem, notably Stephen Smith [122, 123, 124], Aschbacher [8, 9], and Janko [91].

5. Other major developments. The preceding discussion shows how the N-group theorem by itself, rather than any grand design for classifying the simple groups, provided the impetus for a number of major developments. The same remark applies to the three further important developments we now discuss, each of which arose within a specific classification context.

We shall need the definition of two additional terms. First, a quasisimple subgroup K of a group X is called a *component* of X if every X-conjugate of K distinct from K commutes elementwise with K. Thus the normal subgroup generated by all X-conjugates of K is a product of pairwise commuting quasisimple subgroups, each isomorphic to K.

Second, if S is a Sylow 2-subgroup of X, then a subgroup T of S is said to be *stongly closed in S with respect to X* provided for every $t \in T^{\#}$, any X-conjugate of t which lies in S is necessarily contained in T. [S itself is, of course, strongly closed in S with respect to X.]

A. *The Bender method.* As noted earlier, Thompson's method of removing core obstruction was quite cumbersome. It was based on certain arguments in the odd order paper that had led there to the existence of strongly p-embedded subgroups for odd primes p (when $m_p(G) \geqslant 3$). Walter and I [81] used essentially the same procedure in the dihedral case, its execution being even more complicated than for N-groups since now G could contain nonsolvable local subgroups. [A deep character-theoretic result of Brauer, fortunately, enabled one to by-pass what would surely have been an even more difficult analysis in the quasidihedral and wreathed cases.]

Once the dihedral Sylow 2-group theorem had been completed, Walter turned to groups with abelian Sylow 2-subgroups [150]. Since the 2-rank of such a group could be arbitrarily high, he hoped that in the course of determining all such simple groups, he would be able to develop effective

techniques for studying increasingly broader classes of simple groups. How-ever, Walter, too, was forced to follow the same general approach in dealing with core obstruction and his final analysis was horrendously intricate, offering little hope of generalization and further applicability.

It was out of deep frustration with these unpleasant core obstruction developments and the impending impasse towards which they were inexorably heading that I began to search for an alternative approach and was eventually led to the signalizer functor method.

About the same time, Bender, after completing the strongly embed-ded subgroup theorem, re-examined Thompson's construction of strongly p-embedded subgroups in the odd order paper, producing an ingenious sim-plification [27]. His approach was so effective that he tried next to apply it to the abelian Sylow 2-group problem [26]. His results were even more striking, for they led to a very short proof of Walter's classification theorem and, in particular, the removal of core obstruction. Soon thereafter he produced an equally dramatic simplification of the dihedral Sylow 2-group classification theorem [29].

In contrast to the signalizer functor method, which is essentially *constructive*, Bender immediately focuses on a maximal subgroup M of G containing the centralizer C_x of a suitable involution x of the group G under investigation. Using the fact that M is not strongly embedded in G, he developed general techniques—subsequently named the *Bender method*—for forcing severe re-strictions on the structure of both C_x and M, the effect of which carried him a long way towards pinning down the exact structure of C_x (in the abelian case, approximating Brauer form; in the dihedral case Brauer form does not exist, as then C_x is necessarily solvable).

After that, Goldschmidt [69] applied the Bender method to obtain an important extension of the latter's abelian Sylow 2-group theorem, classifying simple groups in which a Sylow 2-subgroup S contains a nontrivial abelian subgroup A that is strongly closed in S with respect to G (the case $S = A$ corresponding to the abelian Sylow 2-group case). Goldschmidt's result in turn became a general tool for analyzing 2-fusion in later classification theorems.

For a time, the Bender method vied in importance with signalizer functors. However, we gradually learned that, in general, the case in which $C_M(O_p(M))$ $\leqslant O_p(M)$ for some odd prime p presents seemingly intractable difficulties, so that it is now believed that its effectiveness is limited primarily to groups with certain restricted types of Sylow 2-subgroups, where one has sufficient control of 2-fusion and subgroup structure to surmount these intrinsic problems.

B. *The B-conjecture, L-balance, and the classical involution theorem.* There is a general formulation of core obstruction in terms of the structure of the centralizers of involutions in the simple group G (or, more generally, arbitrary finite groups), which Thompson called the *B-conjecture*. Throughout the 1970s, verification of this conjecture was a major focus of simple group theory.

Walter and I, using signalizer functor methods, were the first to consider the conjecture [84], always limiting ourselves, however, to groups with internal structures resembling those of groups of Lie type of odd characteristic. It was not until much later that we discovered (independently) how to extend the

method to broader situations. However, in the course of these investigations we discovered a fundamental property of centralizers of involutions in arbitrary finite groups (later extended to elements of any prime order p), called *L-balance*. *L*-balance captures a formal connection between the structures of the centralizers of two distinct commuting involutions of a group G. To comprehend the full classification proof, and in particular, the significance of the *B*-conjecture, it is essential to understand this concept.

Diagram 1 conveys its basic content. Here, x, y are commuting involutions of G, $\overline{C}_x = C_x/O(C_x)$, $\tilde{C}_y = C_y/O(C_y)$, and $\overline{K}$ is a component of $\overline{C}_x$. Since x, y commute, $y \in C_x$ and $x \in C_y$. Although the result holds more generally, we assume for clarity that y leaves $\overline{K}$ invariant and we let $\overline{I}$ be a component $C_{\overline{K}}(\bar{y})$, I its preimage in $C_x \cap C_y$, and $\tilde{I}$ the image of I in $\tilde{C}_y$.
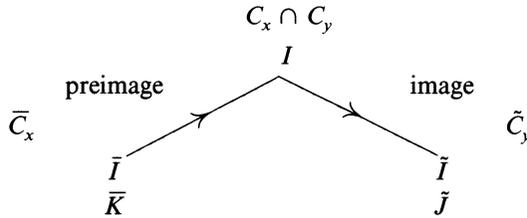
DIAGRAM 1

L-balance asserts that there exists a component $\tilde{J}$ of $\tilde{C}_y$ (in some cases the product of two components of $\tilde{C}_y$) such that

$$\tilde{I} \text{ is a component of } C_{\tilde{J}}(\tilde{x}).$$

Here is how *L*-balance is applied in practice: One starts with the component $\overline{K}$. Then, from the table of centralizers of involutions acting on known simple groups, one can determine the possible isomorphism types of $\overline{I}$ and hence also of $\tilde{I}$. Since $\tilde{I}$ is a component of $C_{\tilde{J}}(\tilde{x})$ and $\tilde{J}$ is quasisimple, one can now reverse the table and use it to determine the possibilities for $\tilde{J}$.

In this way we obtain information about the structure of $\tilde{C}_y$ from that of $\overline{C}_x$. For example, if $\overline{K} \cong PSL_4(4)$, and, say, $\bar{y}$ induces a field automorphism of $\overline{K}$ of period 2, then $\overline{I} \cong PSL_4(2)$ ($\cong A_8$), and the table yields the following possibilities for $\tilde{J}$:

$$PSL_4(2), \ PSL_4(4), \ A_{2n}, \ n \geqslant 4, \text{ or } HS.$$

This passage,

$$\overline{K} \to \overline{I} \to I \to \tilde{I} \to \tilde{J},$$

is called the *pumping-up* process, and we say that *I pumps up* to $\tilde{J}$.

Notice that the cores $O(C_x)$ and $O(C_y)$ really do act as obstructions to the process, forcing us to work inside the *factor* groups $C_x/O(C_x)$ and $C_y/O(C_y)$. The significance of the *B*-conjecture is that, once established, it removes these obstructions, thereby enabling us to perform the pumping-up process within the groups C_x and C_y themselves. In that case, K is a component of C_x, I a component of $C_K(y)$, and J is a component of C_y, so that the pumping-up

process goes directly from

$$K \rightarrow I \rightarrow J.$$

This gives the essence of the *B*-conjecture; equivalently, of the removal of core obstruction.

Shortly thereafter, Thompson tried to establish the *B*-conjecture, in general, using the Bender method. Despite making considerable progress, he soon encountered the inherent difficulty, mentioned above, in applying the method in so broad a context. However, one aspect of the problem led him to raise a very interesting general question. After some slight reformulation, it became the basis of Aschbacher's celebrated *classical involution* theorem, which provides a deep characterization of the groups of Lie type of odd characteristic in terms of their so-called "canonical $SL_2(q)$ subgroups."

An involution x of a group G is said to be *classical* provided:
(1) $\bar{C}_x = C_x/O(C_x)$ has a component $\bar{K} \cong SL_2(q)$ for some odd q; and
(2) $\bar{x}$ is the central involution of $\bar{K}$.
Aschbacher proved [7]

THEOREM. *If G is a simple group which contains a classical involution, then G is necessarily a group of Lie type of odd characteristic.*

This beautiful result turned out to be the key to the ultimate verification of the *B*-conjecture and, in addition, provided the basis for validating the Brauer principle for centralizers of involutions in Brauer form with Brauer component of Lie type of odd characteristic (apart from some low Lie rank exceptions).

C. *Fischer's theory of transpositions and the root involution theorem.* As mentioned in Chapter I, Fischer's sporadic groups $M(22)$, $M(23)$, $M(24)'$ were discovered in the process of classifying simple groups G containing a conjugacy class of 3-transpositions [56].

Given the success of this effort and the new techniques that Fischer introduced in the course of the proof, it was natural to attempt to classify groups having a conjugacy class of involutions satisfying variations of the 3-transposition condition. The first major extension was due to Aschbacher [5], who treated the case in which, for x, $y \in \mathscr{D}$,

(1) $|xy| = 1, 2,$ or odd.

Meanwhile, Fischer and his disciple Timmesfeld considered the case in which, for x, $y \in \mathscr{D}$,

(2) $|xy| = 1, 2, 4,$ or odd.

This case has an obvious division: namely, one can either postulate or not that whenever $|xy| = 4$ for x, $y \in \mathscr{D}$,

(3)   The involution $(xy)^2 \in \mathscr{D}$.

This is not an artificial condition, for in the groups of Lie type of characteristic 2 the conjugacy class of so-called "root" involutions satisfy conditions (2) and (3) (with $|xy| = 4$ for some pair x, $y \in \mathscr{D}$), except for the unitary and

symplectic groups, which instead satisfy Aschbacher's condition (1). Timmesfeld called such a conjugacy class $\mathscr{D}$ satisfying (2) and (3) a class of *root involutions*—*nondegenerate* or *degenerate* according as $|xy| = 4$ for some pair x, $y \in \mathscr{D}$ or not.

Ultimately Timmesfeld obtained the following magnificent characterization of the groups of Lie type over $GF(2^n)$ [141].

THEOREM. *If G is a simple group which contains a nondegenerate conjugacy class of root involutions, then either G is a group of Lie type of characteristic* 2 *(apart from $PSU_m(2^n)$ or $PSp_m(2^n)$), A_6, or J_2.*

The degenerate case was, of course, covered by Aschbacher's and Fischer's prior results. The root involution theorem had significant impact on the classification of the simple groups—in particular, forming the basis for Timmesfeld's general classification results on groups of $GF(2)$-type.

6. Simple groups of non–characteristic 2 type. These examples amply illustrate the step-by-step development of the classification proof. Each new theorem, whether a classification result or a new technique, became a platform on which to see a little further, and as our understanding of the simple groups grew, so, too, the scaffolding extended our horizons. Nevertheless, it was not until nearly 1980 that a complete global strategy became fully visible.

On the other hand, we knew in a general way where we were heading. One couldn't even contemplate the general classification theorem without knowing that groups of odd order were solvable. Likewise, it clearly made most sense to begin by investigating simple groups that were "small" by some measure. Later we knew, too, the central importance of eliminating core obstruction to reach Brauer form for the centralizer of some involution. And, of course, we also fully realized that no final resolution would ever be possible until all simple groups had been uncovered.

But where was it all leading? Again it was the N-group paper that pointed the way. Since Thompson's first objective (when $m_2(G) \geqslant 3$) had been to force G to be of characteristic 2 type, so too this same conclusion became an overriding objective for our minimal counterexample G to the general classification theorem.

Once the B-conjecture was proved, simple group theory reached a new plateau, for then if G was not of characteristic 2 type, yet another fundamental result of Aschbacher—known as his *component* theorem—was directly applicable to G [6].

THEOREM. *If G is a simple group of non–characteristic 2 type in which the B-conjecture holds, the centralizer of some involution x of G is in Brauer form.*

It is this theorem that *internally forces* $\mathrm{Soc}(G)$ *to be simple.*

Thus, to complete the proof that a minimal counterexample G is indeed of characteristic 2 type, we are left with the task of validating the Brauer principle for every possible Brauer component K. In other words, we must "plug in" all covering groups of known simple groups for K and in *each* case internally

construct Soc(G) and then use it to identify the isomorphism type of G (deriving instead a contradiction when there exists no simple group having a Brauer component K of a given isomorphism type).

This turned out to be a very formidable undertaking, requiring many years of effort by some 50 individuals. The complete verification of the Brauer principle is spread across some 75 separate papers, covering perhaps 2,000 journal pages.

I shall not attempt to give a complete list of solutions of the Brauer principle, but they include all the groups of Lie type of odd characteristic and alternating groups (excluding some small cases) plus most of the sporadic groups.

It will be instructive to summarize diagrammatically the basic flow chart of the proof that a minimal counterexample G to the classification of the finite simple groups is of characteristic 2 type (see Diagram 2).

Groups of non-characteristic 2 type
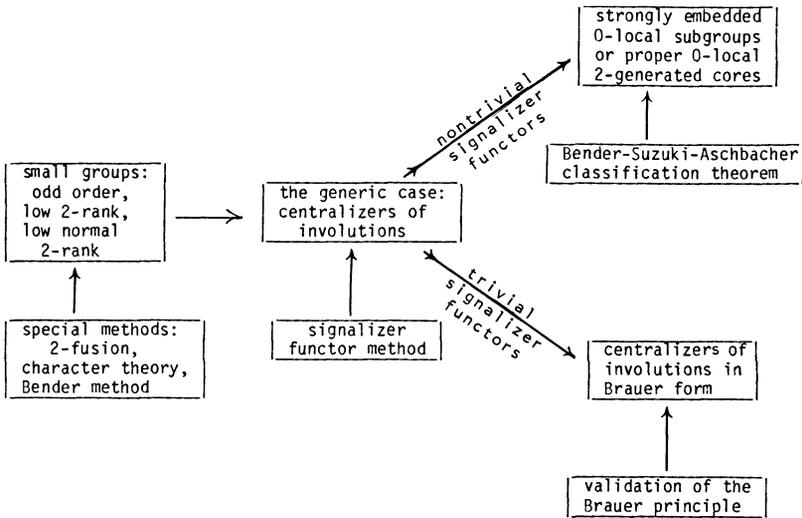
DIAGRAM 2

7. Simple groups of characteristic 2 type. To complete the classification of the simple groups, we were thus left with the task of determining those of characteristic 2 type—the anticipated answer being the groups of Lie type of characteristic 2, together with a few small groups of Lie type of odd characteristic and the remaining sporadic groups.

Here, too, Thompson's N-group paper laid out the path. However, considerable time elapsed before any major attempts to extend his result to broader classes of groups of characteristic 2 type, primarily because of the enormous effort expended on proving that a minimal counterexample was, in fact, of characteristic 2 type and on searching for new sporadic groups.

Thus it was not until 1976 that the first major extension was achieved—the classification of simple groups of characteristic 2 type with solvable 2-*local* subgroups (in place of the solvability of *all* local subgroups). The proof closely followed Thompson's N-group format. He had introduced the key general notion of the *odd 2-local rank* $e(G)$ of G as the maximum p-rank $m_{2,p}(G)$ of a 2-local subgroup H of G, taken over all such H and all odd primes p, and had realized the necessity of singling out the cases $e(G) \leqslant 2$ for special treatment.

The "generic" $e(G) \geqslant 3$ solvable 2-local case was handled by Lyons and me [79], using the signalizer functor method, but now for suitable odd primes p, with nontrivial functors on the one hand, leading to the construction of strongly p-embedded 2-local subgroups and trivial functors, on the other, to "Brauer form" for the centralizer of a suitable element of order p (the concept analogous to Brauer form for centralizers of involutions; however, in the present restricted context, they arose only for $p = 3$). The $e(G) \leqslant 1$ and $e(G) = 2$ cases were treated by Janko [131] and Fred Smith [98] respectively, while the exceptional $GF(2)$-case was handled by Lundgren [101].

Thus a pattern was emerging very similar to that for groups of non–characteristic 2 type, but focusing instead on odd primes rather than on the prime 2: namely, "small" groups ($e(G) \leqslant 2$ and $GF(2)$-type) again requiring special treatment, and the general case ($e(G) \geqslant 3$) being analyzed by means of an analogous internally distinguishing simplicity process. Undoubtedly it was this close parallel to the non–characteristic 2 case that explains why the complete determination of the simple groups of characteristic 2 type, although requiring some 3,000 journal pages, took only about six years: Aschbacher, Lyons, and I [14, 20, 80] handled the $e(G) \geqslant 3$ case, Aschbacher [13] the $e(G) \leqslant 1$ case, subsequently extended by Mason [99] to the $e(G) \leqslant 2$ case (the latter itself consuming almost 1,000 pages!), Timmesfeld and Stephen Smith [122, 123, 124, 142] (and others) treated the $GF(2)$-type case, Aschbacher [15, 16] proved that the strongly p-embedded 2-local case, p odd, led to a contradiction, and Gilman and Griess [161] (with the support of Finkelstein, Frohardt, and Solomon for Brauer components defined over $GF(2)$) substantiated the Brauer principle for odd primes. [As part of our analysis, Lyons and I showed that only Brauer components of Lie type of characteristic 2 had to be considered.]

To carry through the analysis, however, some fundamental new techniques were required. Most notable is Aschbacher's theory of "blocks" [10] (cf. also [76]) (*Aschbacher blocks* are certain types of subgroups of a group, somewhat similar to components; they bear no relation to Brauer's blocks of characters). Such blocks are intimately connected with the Baumann-Glauberman-Niles theory of "pushing up" 2-local subgroups [23, 64, 66], and their analysis culminated in the fundamental "$C(G; S)$-theorem," which determined all simple groups G of characteristic 2 type having a proper *characteristic 2-core*

$C(G; S)$. Here, by definition, if S is a Sylow 2-subgroup of G,

$$C(G; S) = \langle N_T | 1 \neq T \text{ characteristic in } S \rangle.$$

Obviously, if G has a strongly embedded subgroup, it has a proper characteristic 2-core, but the converse is by no means true. Indeed, if $C(G; S) \leq M$ for some proper subgroup M of G and M is not strongly embedded in G, it is easily shown that some maximal 2-local subgroup of G must contain an Aschbacher block (under the assumption that G is of characteristic 2 type). Moreover, the $C(G; S)$-theorem was obtained as a corollary of the classification of simple groups of characteristic 2 type in which some maximal 2-local subgroup possesses an Aschbacher block. Here is the statement of the $C(G; S)$-theorem (see Foote [53] for a discussion of the theorem).

THEOREM. *If G is a simple group of characteristic 2 type in which $C(G; S)$ is a proper subgroup, S a Sylow 2-subgroup of G, then either G has a strongly embedded subgroup (whence G is determined by Bender's theorem) or $G \cong PSL_3(2^n)$, $PSp_4(2^n)$, $PSL_2(q)$, q a Fermat or Mersenne prime or 9, $PSL_3(3)$, M_{22}, or M_{23}.*

As a corollary, one obtains a classification of all simple groups G of characteristic 2 type, in which a Sylow 2-subgroup S of G is contained in a *unique* maximal subgroup M, for then obviously $C(G; S) \leq M$, whence $C(G; S)$ is a proper subgroup of G.

Furthermore, near the end of the characteristic 2 type analysis, Goldschmidt's theory of amalgams [63], which was just then emerging as a subject of intrinsic interest and importance, played a crucial role in eliminating certain difficult residual configurations of 2-local subgroups.

Nevertheless, summarizing the proof of the classification of groups of characteristic 2 type, we obtain a diagram (Diagram 3) strikingly similar to that given in the previous section for groups of non–characteristic 2 type.

It will be convenient, for brevity, to refer henceforth to groups of non–characteristic 2 and characteristic 2 type as groups of *noneven* and *even* type, respectively. [As noted above, the noneven groups include not only the groups of Lie type of odd characteristic, but also most of the alternating and sporadic groups.] They constitute a fundamental division in the existing classification proof, providing the basis for internally distinguishing whether G is correspondingly of odd characteristic or of characteristic 2.

8. A partitioned proof. The fact that we worked during most of the "thirty-year war" with only a limited picture of the final outcome was crucial in shaping our perspective towards the classification of the finite simple groups. Since we neither knew all the simple groups nor had a strategy for completing the proof, we were unable to seriously consider a minimal counterexample to the full classification theorem.

The military metaphor is apt, for we definitely felt embattled in those years, fighting an adversary far stronger than ourselves. Indeed, Richard Brauer died in 1977 with the conviction that a complete classification, if attainable at all, was far in the future. Since we never knew what lay ahead, we were inevitably forced to view each classification theorem under consideration as necessarily an *independent* result, completely divorced from the general classification theorem itself. This was our way of "protecting our flank."

This necessity to establish only *compartmentalized* results became so deeply ingrained an attitude that it persisted for most simple group theorists until the very end, even though by 1978, when at least "evidence" for the existence of each of the sporadic groups had been obtained, our understanding of the simple groups henceforth justified consideration of a minimal counterexample to the full classification theorem.

For brevity, I shall refer here to the existing classification proof as *partitioned*.

Groups of characteristic 2 type
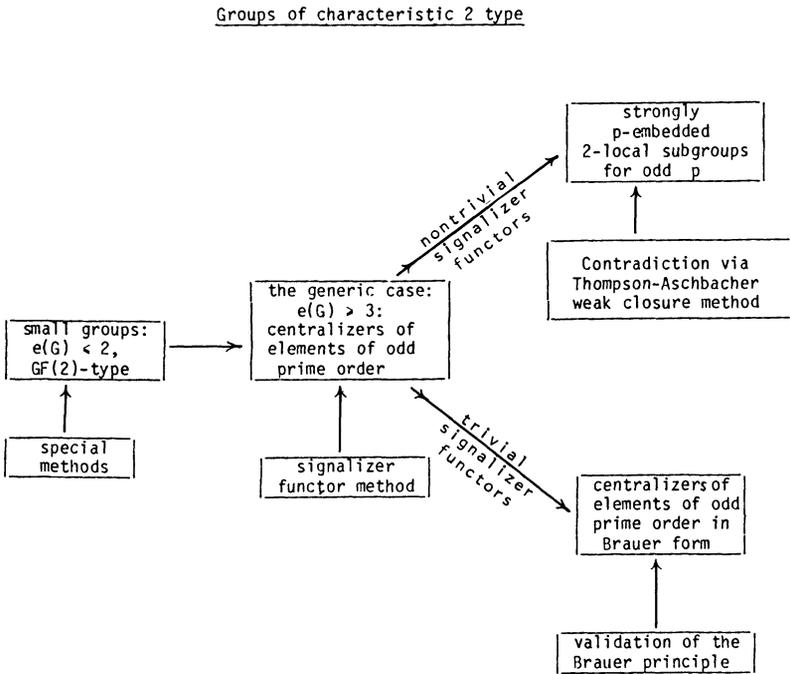
DIAGRAM 3

9. Inevitable or accidental proof? In describing the evolutionary character of the classification proof, I have stressed the groping quality of our early investigations, the specific contexts from which major developments grew, and the lack of global strategy until the final few years. Taken together, they

perhaps convey substantial arbitrariness, group theorists at various junctures choosing one path among many, filling in the open links as they plugged along, and ending up with a patchwork sort of proof.

But I also noted that many of the paths we pursued became blocked, repeatedly forcing us in the directions of our past successes, so that the links being forged did seem to possess some inherent tension.

The history of the use of character theory in the proof provides a good illustration of this forcing process. In the early years, it was the most effective available tool, and throughout most of the period many group theorists vigorously attempted to increase its range of applicability. But gradually the method's limitations came to be understood: character theory could be used to derive *structural* information about simple groups only when *tight* conditions were imposed on the structure and embedding of critical proper subgroups. If effort alone were the only measure of success, character theory would surely have played a much more dominant role throughout the entire proof! It was the underlying reality, not our preferences, that forced finite group theorists to develop other techniques for analyzing the general simple group.

My own view is that, while several options may have been available for verifying a particular portion of the long chain of the proof, its global design possesses an *inexorable* quality of profound import for understanding the finite simple groups.

First, the striking similarity between the flow charts of §§6 and 7 for the classification of the groups of noneven and even types, respectively, is too great to be entirely coincidental. In fact, almost all Thompson's work on *N*-groups, which as I have pointed out earlier represented the schematic basis for the attack on the general group of even type, was completed well before central-izers of involutions in the general group of noneven type were investigated. Conversely, since the *N*-group theorem focused on *solvable* subgroups, its proof had essentially no impact on the study of *components* of centralizers of involutions.

Thus the basic approaches for analyzing the simple groups of noneven and even types seem to have been developed independently of each other. Yet, there is the unmistakable fact that the basic outlines of their analyses are essentially identical.

Moreover, what is not generally realized is that, even though a group of odd order contains no involution, the global outline of the proof of its solvability has the same general form as those of groups of noneven and even types! Indeed, the first part of the Feit-Thompson analysis again involves the structure and embedding of the maximal subgroups of G containing a given Sylow p-subgroup of G (where now p ranges over all prime divisors of $|G|$). Bender's simplification of the original Feit-Thompson argument covers the case $m_p(G) \geqslant 3$. However, the p-rank 2 case is very difficult, requiring its own separate analysis. Thus again low-rank cases are exceptional.

This *local analysis* portion of the proof yields very precise information about the structure and embedding of every maximal subgroup of G (there are five possible *types* of maximal subgroups; a general description of their structure and embedding is given in [77, §1.2]). Considering the nature of the subsequent

analysis, it is reasonable to regard the resulting patterns of maximal subgroups as "odd order analogues of Brauer form." Indeed, the ensuing *character-theoretic* arguments eliminate all but a single such configuration of maximal subgroups and so can be viewed as the first major step in validating the Brauer principle in groups of odd order.

This final configuration involves two closely linked maximal subgroups M and N, with which there is associated a field $F = GF(p^n)$ for some odd prime power p^n, which can be interpreted as the "field of definition" of the minimal counterexample G. [Actually, a second field over $GF(q^m)$, $q \neq p$, is associated with G in an entirely parallel fashion. The field F on which the analysis focuses is distinguished by the condition $p > q$.] Thus it is also reasonable to view the pair M, N as constituting "the odd order analogue of strong Brauer form." [See Part IV, §13, for the definition of strong Brauer form.] Again the terminology is justified by the nature of the *generator-relation* analysis of the interdependent subgroups M and N that leads to a final contradiction. Indeed, this generator-relation portion of the proof ultimately forces the multiplicative group of F to have a certain crucial property, which is then shown, on purely field-theoretic grounds, to be impossible in any finite field of odd characteristic.

Many years of reflection on such considerations has led me to the conviction that no matter how any future proof of the classification theorem may be packaged, it will necessarily include the following five components as an *irreducible minimum*:

   (1) An internal criterion for G to be of even characteristic.
   (2) A separate analysis for small size groups.
   (3) Internal elimination of Rad(G) by excluding large subgroups.
   (4) Internal construction of Soc(G) by validating the Brauer principle.
   (5) Determination of the possible isomorphism types of G in terms of a presentation by generators and relations (or its equivalent).

CHAPTER III.   CONSTRUCTING A SATISFACTORY PROOF

Normally, significant new theorems are carefully checked by the experts and later read by a broader spectrum of mathematicians, in the process becoming part of the accepted body of mathematical knowledge. Yet it can be safely asserted that no one individual has read the proof of the classification theorem in its entirety. Indeed, some of the individual papers have been of such lengths that journal editors were forced to assign separate portions of them to different referees! Thus, just as the proof itself was a team effort, so, too, has been its reading.

Moreover, along the way a number of local errors have been discovered—hardly surprising, given the length and number of papers involved. These have been corrected as soon as pointed out, most often on the spot. Nevertheless, this is somewhat disconcerting since no matter how partitioned its organizational structure, the argument is, after all, in totality inductive, so that the possibility of an apparently "local" error having "global" implications is ever present.

This problem is not unique to simple group theory—we can all cite accepted "theorems" that later actually turned out to be false. What makes the classification theorem so exceptional is that the inordinate length of the proof forces the checking process to continue long after the initial announcement of the theorem.

This then is the present unsatisfactory reality. Most simple group theorists are fully convinced that the "essential" shape of the existing classification proof is correct. Perhaps the best evidence of its over-all validity is the fact that, despite some earlier predictions to the contrary, no new simple groups have been found since its announcement!

However, these comments underscore the opening statement of Graham Higman's Melbourne, Australia, lectures in January, 1981: "The single most important problem facing simple group theory is to place the classification theorem on a sound footing."

In making this comment, Higman also had in mind the added obligation of simple group theorists to provide present and future generations of mathematicians reasonable access to the theorem.

Hence, for all these reasons, I believe it is imperative that a satisfactory proof of the classification theorem be constructed.

1. Revising the proof. Unfortunately, there are aspects of the classification proof that make its systematic checking extremely difficult—in fact, in practical terms, essentially impossible.

First, as the discussion of Chapter II indicated, each investigator, with the notable exception of Bender, took up his work at the point where the last had left off: the task ahead always seemed so enormous that little incentive was left for re-examining what had already been achieved. Moreover, proofs of individual theorems were themselves of such lengths that it became standard practice to invoke any and all available results, without ever seriously investigating whether a short additional argument might, in a particular case, have avoided a given reference. Thus, the collection of ostensibly "independent" classification theorems, forming the primary links of the global proof, in reality depend critically on an elaborate network of supporting results, often even including prior such "independent" theorems.

Moreover, the final pattern of the proof is so intricately intertwined that there are only a handful of group theorists presently capable of even delineating a *detailed* flow chart of the argument.

The most striking example of this interweaving occurs in connection with completion of the proof of the B-conjecture. The proof involves some dozen difficult papers, so closely linked that a very careful reading is required to convince the reader that, taken together, they in fact establish the desired result. Furthermore, just when one is struggling towards the finish line, suddenly out comes a "blockbuster" independent theorem—its only critical usage in the entire classification proof!—to provide the coup de grace: namely, the classification of all simple groups with Sylow 2-subgroups of order $\leqslant 2^{10}$, the complete solution of which is comparable in length and technical difficulty to the sectional 2-rank $\leqslant 4$ theorem, on which the proof is modeled.

A similar phenomenon occurs in the final validation of the Brauer principle for centralizers of involutions, which, as already noted, covers a great many papers. Again a considerable effort is required just to convince oneself that all possible Brauer components have indeed been considered.

A second, equally unpleasant feature of the proof is that many of the early papers, still essential for the full classification theorem and often having very long proofs, were written without benefit of later developed techniques, so that a line-by-line checking does not seem to be the appropriate response—it would clearly be more fruitful to rethink the given theorem (with or without its original proof as guide). The quasidihedral, wreathed Sylow 2-group classification theorem—covering more than 300 journal pages—illustrates the point. Rather than reading the proof in detail, one would be far better advised to first extract the gist of the argument and then to ask, on the one hand, to what extent the Bender approach might simplify the local group-theoretic portion of the analysis, and, on the other, whether it is possible to replace Brauer's delicate block-theoretic calculations by more elementary arguments. [Indeed, the first of these questions has recently been successfully tested, while work on the second is in progress.]

Thus, whether we wish it or not, attaining Higman's "sound footing" is forcing us to carry out some type of *revision* of the existing theorem, for it is evident that by itself the checking process will never lead to a convincing, let alone satisfactory, proof.

What will such a revision entail? There are four quite distinct ways of approaching the problem, which I shall colorfully call the "Radical," "New Idea," "Conservationist," and "Major Surgery" parties, each applicable to a portion of or to the proof as a whole.

The *Radical* party considers the finite simple group theorists responsible for the present 15,000 page mess and no amount of reshaping of their internalized proof will improve the situation very much. For them, a satisfactory proof must await the discovery of some revolutionary, most likely external approach to the classification.

The *New Idea* party views our present understanding as inadequate for achieving the stated objective—it, too, submitting the present 15,000 page proof as supporting evidence. It believes that new internal ideas and techniques are required before it will be possible to find a satisfactory classification proof.

At the other end of the political spectrum are the *Conservationists* who feel that the basic structure of the proof is quite sound. Hence for them, it will be entirely sufficient to revise and otherwise separately improve the major pieces of the simple group theory puzzle to obtain a satisfactory proof.

Finally, the *Major Surgery* party, too, feels that the present global framework is sound, but believes that a complete re-examination, even though limited by the existing state of technology, but having the advantages of hindsight, can lead to a substantially reorganized, more conceptual—what one may call a "second generation"—classification proof.

Considering the long-range mathematical significance of the theorem, I feel that all four approaches should be simultaneously intensely pursued. As I shall describe in the next section, my own analysis of the existing proof has led me

into the major surgery camp. Indeed, I believe that that approach offers the best prospects for deriving as satisfactory a proof as our present knowledge will permit. Moreover, the sounder the health in which we leave the classification theorem at this time, the more accessible we can make it, the greater will be the chances of fundamental new ideas or even alternate approaches being discovered in the future.

2. A critique of the existing proof. About four years ago, as it was becoming increasingly apparent that the classification was nearing completion, I began for the first time to think about the proof we were fashioning as a single totality. No matter how forced our moves had been, given the evolutionary nature of its development, it would be too much to expect that the proof we were constructing would miraculously turn out to be the most efficient and coherent attainable by the existing techniques.

Indeed, I was soon discovering a number of puzzling features of both the global strategy and its execution. The first collection of these clustered around the way in which the groups of Lie type of characteristic 2 were being determined, and the second emerged from a comparison between our treatment of centralizers of involutions in groups of noneven type, on the one hand, and centralizers of elements of odd prime order in groups of even type, on the other. (For brevity, I shall refer here to the latter two situations as the *involution* and *odd prime* cases, respectively.)

A. *The groups of Lie type of characteristic* 2. To me, the single most surprising feature of the proof was the fact that our global strategy involved four distinct characterizations of the groups of Lie type of characteristic 2:

1. As part of the validation of the Brauer principle for centralizers of involutions.

2. In validating the Brauer principle for centralizers of elements of odd prime order.

3. In classifying the groups of $GF(2)$-type.

4. In establishing the root involution theorem.

Whatever significance I wished to assign each of these independent classification theorems, I could not escape the redundancy they collectively embodied. Certainly the groups of Lie type of characteristic 2 are somewhat special, particularly as the internal structure of the sporadic groups strongly suggests that they are linked to the groups of Lie type over the prime field $GF(2)$. But could there really be something so *intrinsically* distinguishing about these families that they should require four separate characterizations, whereas one sufficed for all the other simple groups? Or rather was this simply a product of the evolutionary, accidental growth of the proof?

B. *Centralizers of elements of prime order.* The diagrams at the end of Chapter II reveal the striking global conceptual parallel between the analysis of groups of noneven type versus groups of even type, particularly with respect to the use of the signalizer functor method in studying centralizers of involutions, in the first case, and centralizers of elements of odd prime order, in the second; i.e., between the involution and odd prime cases. Yet despite this strong

parallel, there were a number of major differences in how the two cases were being handled:

1. *Groups of normal 2-rank* 2. We have seen the historical importance of this family of simple groups, their classification leading to consideration of the inductive family of simple groups of sectional 2-rank $\leqslant 4$, whose subsequent classification constitutes one of the major chapters of the present proof. In contrast, in the odd prime case, no analogue of the sectional 2-rank $\leqslant 4$ theorem was either established or required. Indeed, a short argument enabled us to incorporate the normal p-rank 2 case within the general analysis.

2. *The classical involution theorem.* No analogue of Aschbacher's fundamental classical involution theorem was either established or required in the odd prime analysis.

3. *The B-conjecture.* For each prime p, one can formulate what one may call the B_p-*conjecture* for finite groups (the B_2-conjecture being the same as the B-conjecture), which makes precise the removal of p'-core obstruction in centralizers of elements of prime order p. Moreover, the full classification theorem implies the validity of the B_p-conjecture in arbitrary finite groups for all primes p.

As we have earlier described, verification of the B-conjecture held a dominant position throughout the analysis of the involution case. Yet, in the odd prime p analysis, only a partial form of the corresponding B_p-conjecture was ever needed.

4. *Signalizer functors.* In the proof of the classical involution theorem—which can be properly viewed as the general case of the B-conjecture—Aschbacher was able to apply the signalizer functor method very smoothly, so that core obstruction per se was not a very serious problem for him. In contrast, Lyons and I were forced to elaborate lengths to achieve the removal of p'-core obstruction in the odd prime p case.

5. *Validation of the Brauer principle.* In the involution case, validation of the Brauer principle for Brauer components K of Lie type of *odd characteristic* is derived as a direct corollary of the classical involution theorem (except for certain low Lie rank K's). However, in the odd prime case, with the corresponding K of Lie type of *characteristic* 2, validation of the Brauer principle was carried out by the ostensibly quite distinct *neighbor* method, as it has come to be called.

Actually, the distinction is not as great as appears, for a portion of Aschbacher's analysis can be interpreted as determining the isomorphism type of G by the neighbor method. However, the context in which his argument occurs is so different from that of the odd prime case that this inherent parallelism is almost completely camouflaged.

Thus, it is likewise natural to wonder how essential or accidental are the differences between the treatments of the involution and odd prime cases.

3. A unitary perspective. In my prior work on simple groups I had been as steeped in the developmental, partitioned approach to the classification as any other group theorist. But how else could I analyze the troubling questions I

have just raised without coming to think of the classification of the simple groups as a *single* theorem?

1. Can the global proof be reorganized in such a way that the groups of Lie type of characteristic 2 are characterized only *once*—in the process of validating the Brauer principle in the odd prime case for Brauer components of Lie type of characteristic 2, thereby *avoiding* (a) verification of the Brauer principle in the involution case for Brauer components of Lie type of characteristic 2, (b) separate treatment of the general $GF(2)$-type case, and (c) any reference to the root involution theorem?

2. Is it possible to handle the normal 2-rank 2 involution case in the same easy way as had been done in the odd prime case?

3. Given the striking similarity between the schematic diagrams of the involution and odd prime cases, is it possible to carry through the signalizer functor method for *all* primes simultaneously—at least in the generic cases? And if so, what would be its conceptual significance?

4. In the same vein, can the Brauer principle be validated in the involution case for Brauer components of Lie type of odd characteristic by the neighbor method, thereby merging this case with that of the corresponding odd prime case of Brauer components of Lie type of characteristic 2?

5. Pursuing this line of inquiry, can one perhaps locate other major sections of the proof amenable to consolidation and simplification? For example, can the various subcases of the 2-rank 2 analysis (dihedral, quasidihedral, wreathed, homocyclic abelian, or type $PSU_3(4)$ Sylow 2-groups) be handled simultaneously by the Bender method? Similarly, are there any portions of the proof that can be simplified by greater use of the Goldschmidt amalgam method?

As I began to think about these questions in a more systematic, but still general way, I gradually became convinced there was a high probability of obtaining affirmative answers. Thus I was being led inescapably towards a major restructuring of the classification theorem. This *unitary* perspective, as I have come to call it, held out the promise of a more coherent, much more efficient proof than presently exists. I visualized perhaps as much as a *five-fold* reduction—3,000 versus 15,000 pages. Certainly a great improvement, but still of inordinate length. However, I could see no way of producing a substantially shorter proof on the basis of techniques now available.

Sketching out the general lines of such a unitary proof is one thing, but whether the proposed grand strategy is, in fact, viable can only be decided in the process of detailed verification of its major components. The sheer magnitude of carrying this out demanded a cooperative effort of a number of finite group theorists.

Thus, first, Richard Lyons and later Ronald Solomon joined me in the task of turning the general sketch into the kind of detailed global outline that would clearly be a necessary prerequisite for achieving the desired unitary proof. The process has taken several years—reality continually forcing subtle shifts in our blueprint. Although we cannot be certain until the full proof is completed, we believe that our outline is now essentially in its final form. I shall refer to this as the *GLS* unitary approach, and shall describe its general shape in Part IV.

Our confidence in the *GLS* blueprint derives from the fact that some of its individual parts have already been completed, while work on others is in full progress. In addition to our joint efforts, Delgado and Stellmacher, Foote, and Stroth, respectively, are presently in the process of carrying through three further major steps of the program.

It should also be mentioned that improvements in other areas of the classification proof have been made that, although not motivated by the *GLS* unitary blueprint, nevertheless dovetail with it very effectively. These include the Bender-Glauberman-Sibley-Peterfalvi [65, 118, 155] work on groups of odd order, Peterfalvi's re-examination of the Suzuki-Bender work on groups with a strongly embedded subgroup [110], and Enguehard's similar effort [50] concerning the extremely difficult Thompson-Bombieri [32, 135] validation of the Brauer principle for centralizers of involutions with Brauer components $K \cong PSL_2(q)$, $q \equiv 3, 5 \,(\mathrm{mod}\, 8)$, $q > 5$.

4. A partitioned versus unitary approach. A number of simple group theorists believe that a partitioned approach will lead to a more effective classification proof, for its key independent theorems will, on the one hand, provide deeper insight into the basic structure of simple groups than could a unitary approach, and, on the other, possibly have direct application to other problems. Thus they argue that, although a partitioned proof may entail some degree of duplication, its advantages justify the extra effort.

We have examined this position with great care, for now, while our proposed second-generation proof is still taking shape and before our results are put in final form, there is yet time to modify our blueprint. However, on closer inspection, we have concluded that the differences between a partitioned and unitary proof will, in fact, be much less than their names suggest, for the 3,000 pages of a unitary proof will be as divided into separate chapters, each with its own individual theorems, as will any partitioned proof. Rather, the primary distinction between the two approaches appears to be that *a unitary perspective will push to the forefront a different set of theorems than a partitioned one*.

To illustrate, the proposed *GLS* proof anticipates a result of the following general shape:

Let G be a simple group, x an element of G of prime order p, and $\overline{K}$ a component of $\overline{C}_x = C_x/O_{p'}(C_x)$ with $\overline{K}$ of Lie type of characteristic $r \neq p$ and assume that among all such x and $\overline{K}$, the order of $\overline{K}$ is as large as possible. If the Lie rank of $\overline{K} \geqslant 3$, then one of the following holds:

(1) C_x is in Brauer form with Brauer component K isomorphic to a covering group of $\overline{K}$; or

(2) G contains a strongly p-embedded p'-local subgroup M (with $C_x \leqslant M$).

If (1) holds, then validation of the Brauer principle (by the neighbor method) will show that G itself is isomorphic to a group of Lie type of characteristic r, while if (2) holds, we should be led to a contradiction. In other words, the existence of a *single* "generic" component of Lie type of characteristic $r \neq p$ in the centralizer of an element of order p in a simple group G is enough to force G to be of Lie type of characteristic r, which clearly expresses

a deep general structural property of simple groups. In fact, the result is very much in the same spirit as the classical involution theorem, but is valid for all primes p.

Thus, without a careful comparison of the key theorems making up a unitary versus a partitioned classification proof, it is not really possible to decide which collection provides deeper insight into the structure of simple groups, and very likely such judgment will ultimately rest on personal preference.

However, I would add the following point in favor of a unitary approach: in viewing the classification as a single theorem, one is forced towards an understanding of the *indispensable core* of a classification proof, and the knowledge thereby acquired represents what I consider to be the single deepest insight into the structure of simple groups.

This last comment also bears on a "semi-unitary" approach to the classification proposed by some simple group theorists: determine the simple groups of noneven type along the lines of some improved version of the present proof, but determine those of even type from their 2-parabolic structures—thus, in effect, viewing the full classification as *two* separate theorems.

Pushing through such a 2-parabolic approach will entail a very broad development of the Goldschmidt amalgam method, covering minimal 2-parabolics of seemingly arbitrarily complex structures. The method has certainly been very effective in dealing with "tight" characteristic 2 type configurations. However, my own sense is that just as with the Bender method in the involution case, so, too, there are inherent difficulties in extending the Goldschmidt method to the general group of even type. In any event, no one has yet succeeded in producing a detailed, viable program for determining all simple groups of even type by the amalgam method.

Indeed, at present we seem to have developed only a single effective tool for analyzing the internal structure of the general simple group: *pumping up components in conjunction with the signalizer functor method.*

Many years ago, Thompson advised simple group theorists to "stay away from nilpotent groups, wherever possible." What he meant was this: Only as a last resort try to solve a given classification problem by forcing the exact structure of a Sylow 2-subgroup, for subgroups having more complex structures are easier to analyze.

There seems to be an obverse side to Thompson's dictum: namely, "Wherever possible, work with *components*, for they are the easiest subgroups to analyze!"

5. The K-group hypothesis. If G is a counterexample of least order to the complete classification theorem, then any simple group of order lower than G will obviously be isomorphic to a known simple group. In particular, this will be true for the composition factors of any *proper* subgroup H of G (for brevity, we call such a group H a K-group).

If the classification of the simple groups is being viewed as a single theorem, one begins with such a minimal counterexample G and so imposes a global K-group hypothesis on the subgroups of G at the very outset. However, as in Thompson's minimal simple group problem, the ensuing analysis seems to require this assumption almost exclusively on the local subgroups of G. Let us therefore call a group K-local if each of its local subgroups is a K-group.

Hence the unitary approach leads to the following formulation of the classification theorem:

If G is a K-local simple group, prove that G is a K-group.

Clearly this will imply that a minimal counterexample does not exist.

It is the presence of this global K-group hypothesis that prevents one from proving any results about *arbitrary* simple groups along the way—only when the complete theorem has been established do such intermediary results become general properties of simple groups. However, they are general results about K-local simple groups. Thus, a unitary proof should be visualized in the following way: the classification theorem is to be constructed from a sequence of *independent results concerning K-local simple groups*. Phrased in these terms, the distinction between a unitary and a partitioned proof becomes even more blurred, reducing, in effect, to how global or selective the K-group assumptions are.

For example, the amalgam method deals with a group $G = \langle P_i \mid 1 \leqslant i \leqslant n \rangle$, where $\{ P_1, P_2, \ldots, P_n \}$ is a parabolic system for G. Since the analysis of such parabolic systems focuses solely on these parabolics, it suffices to require the P_i themselves or at worst the subgroups of G generated by pairs, triples, etc. of the P_i to be K-groups. Hence in this chapter of simple group theory, global K-group hypotheses on the subgroups of G are unnecessary.

It should also be pointed out that, although the existing classification proof is viewed as partitioned, the groups of even type have been determined only under a global K-group hypothesis on proper subgroups, despite the fact that from the perspective of the schematic diagrams of §§6 and 7 of Part II their classification constitutes *half* the total proof.

Some group theorists attach considerable weight to the nature of the K-group hypotheses of a given theorem, but to me this seems a very minor issue for the classification of the simple groups as a whole.

There is, however, a more significant side to the issue. Essentially all of the independent theorems of a partitioned proof will themselves be proved by induction, but a considerable effort may well be required to complete that induction. If I may cite the classical involution theorem again, Aschbacher's paper divides into thirty-one sections. However, it is not until section 29 that one reaches the general "connected" case! All prior sections (excluding two that give a general description of the finite groups of Lie type of odd characteristic) are devoted to special cases: largely low-rank configurations plus the general orthogonal case (which from the classical involution viewpoint is somewhat exceptional) and a general "characteristic 2 type" case.

Had Aschbacher been operating under a global K-group hypothesis, a portion of his wonderful argument could have been abstracted to yield a short self-contained proof covering the "generic" component case (including large orthogonal groups), thus highlighting the essence of the theorem.

6. Noninductive results. Given the classification statement, with its twenty-six singular sporadic solutions, it is hard to even visualize the general nature of a complete proof of the theorem that at some point would not require one to investigate a minimal counterexample in close detail. Thus it seems inevitable

that some *theory of K-groups*, analogous to Philip Hall's classical theory of solvable groups (cf. [**74**, Chapter 6]) which was so central to the odd order and N-group theorems, will underlie any proposed proof of the classification theorem, as is certainly the case with the present proof.

However, this observation does not preclude the possibility that some deep structural properties of simple groups can be established noninductively—i.e., without recourse to any K-group assumptions on proper subgroups. Representation theory, character theory, and transfer are examples of noninductive techniques, but, unfortunately, by themselves they provide only superficial structural information.

Brauer and Fowler [**36**] a long time ago obtained the following lovely, noninductive *qualitative* result about simple groups G by elementary counting arguments on the number of involutions of G.

THEOREM. *There are at most a finite number of simple groups in which the centralizer of some involution has a specified order.*

However, this result, too, yields no specific structural information.

I have often been asked at what point in the present proof can we assert that there exists at most a finite number of sporadic groups. Since our proof is inductive, the unfortunate answer is the following: only as a corollary of the complete classification theorem. Clearly an independent proof that all simple groups of *sufficiently high order* have a "regular" internal structure would have profound import. The proof of such a result would of necessity be noninductive.

An attempt, so far unsuccessful, has been made to prove the Sims conjecture independently of the classification theorem. Likewise such a result would have deep impact on the study of simple groups.

Because of the dominant position that the internally distinguishing simplicity process has in my way of thinking about simple groups, I have considerable doubts that such deep noninductive structural results can ever be established. Only the future will determine the soundness of this judgment. However, in the meantime we have no choice but to live with the fact that induction is the single effective general method we have for studying simple groups, allowing us to dissect their *internal* structures on the basis of K-group properties of their subgroups.

CHAPTER IV.   OUTLINE OF A UNITARY PROOF

In this final chapter, I shall describe in some detail the outline of the proposed unitary proof of the classification theorem which Lyons, Solomon, and I have been developing, the general shape of which grew out of an attempt to respond to the various criticisms of the existing proof discussed in Chapter III.

However, determination of its more exact form has required a considerable number of preliminary investigations of the major components of our program. Our objective has been to devise a comprehensive organizational plan leading to as simplified a proof as possible. The difficulty in obtaining such a formulation has been twofold: first, there has been the question of determining

the most effective approach to a particular phase of the classification; second, there has been the problem of deciding where one phase, whose proofs are to utilize one set of group-theoretic techniques, is to end and the next phase, dependent on a different set of methods, is to begin. The various preliminary analyses were undertaken to enable us to make the required decisions.

We must first define several of the key terms which underlie our proposed unitary proof. Only then will we be able to present the outline itself.

As I have said, work on certain portions of the global plan has been completed, on some portions work is in progress, while other portions remain to be systematically investigated. The extent of success so far makes us optimistic about the final outcome.

1. The over-all plan. The *GLS* unitary outline involves a rather complex division into a number of cases, each requiring its own special analysis. It will be easier to follow the technical details if I first describe its general features schematically.

The first step is to internally distinguish the *even characteristic* groups. This will form the basis of our definition of a group of *even* type (a weaker notion than that of a group of characteristic 2 type). In the groups of noneven type, we primarily study centralizers of involutions, while in groups of even type primarily centralizers of elements of odd prime order.

Both types of simple groups will be divided into *nongeneric* and *generic* subclasses. These subclasses in turn require some further divisions, determined by the set of techniques on which the analysis of a particular subcase is to be based. Moreover, these various subcases are to be sequentially ordered in such a way that *each simple K-group arises as a solution of one and only one case*.

The analysis of each case in turn splits into *three* stages. The aim of *stage 1* is to reach Brauer form, which we shall later define precisely. The obstruction to achieving this objective will be the existence of large subgroups, which will form the basis of our definition of a *uniqueness* subgroup—one variety in groups of noneven type and a second in groups of even type.

The aim of *stage 2* is to reach what we shall call *strong Brauer form*. When this stage is reached, the simple group G under investigation has an internal structure very closely approximating that of some simple K-group G^* (in a few instances only that of an *almost* simple K-group G^*).

Finally, the goal of *stage 3* is to show that the multiplication table of G is uniquely determined. Since G^* satisfies the same internal conditions as G, it will then follow that $G \cong G^*$. [Since G is assumed simple, this will also rule out the nonsimple possibilities for G^*.]

Unfortunately, just as was true in Aschbacher's classical involution proof, by far the greater portion of the analysis will be concerned with the determination of the nongeneric simple groups (which include all twenty-six sporadic groups).

2. Groups of even type. The fundamental division in the existing classification proof is between groups of *non–characteristic* 2 type (noneven type), which are studied by analyzing centralizers of involutions, and groups of *characteristic* 2 type (even type), which are studied by analyzing centralizers of elements of odd prime order.

However, as we have described in Chapter III, the final step in proving that a minimal counterexample G to the classification is of characteristic 2 type was an extremely elaborate undertaking, requiring validation of the Brauer principle for all possible Brauer components K. Moreover, the case K of Lie type of characteristic 2 was especially difficult and the process of carrying it out produced one of the four distinct characterizations of the groups of Lie type in the course of the proof of the full classification theorem.

It is therefore natural to ask whether it might be possible to avoid these very difficult Brauer component cases by slightly shifting the point at which one drops the centralizer-of-involution perspective and begins instead to focus on odd primes. In other words, does there exist a slightly weaker notion of a group of even type than a group of characteristic 2 type, to which the present odd prime analysis can be extended without great difficulty?

We have taken an affirmative position regarding this question and have been led to a new definition of a group of even type, which plays a central role in our unitary outline. The idea is to broaden the notion of a group of even type sufficiently that not only will it include the groups of Lie type of characteristic 2 (and other simple groups of characteristic 2 type), but also all sporadic groups of $GF(2)$-type.

If one examines each of the simple G of the latter type, one finds that the following two general conditions are satisfied:

$\qquad$ (1) $\quad O(C_x) = 1$ for every involution x of G; and

$(*)\quad$ (2) $\quad$ If K is a component of C_x for some involution x of G,

$\qquad\qquad$ then $K/Z(K)$ is either of characteristic 2 type or of $GF(2)$-type.

In particular, $(*)$ holds if G is of characteristic 2 type (in which case C_x has no components for any involution x of G). The conditions of $(*)$ thus provide a generalization of a group of characteristic 2 type, sufficiently broad to include all the groups of $GF(2)$-type. In our actual analysis, we replace the condition on $K/Z(K)$ in (2) by a more intrinsic definition that leads to a slightly larger family of groups. The reason for doing this is that the latter definition, when extended to odd primes p, yields a family of groups which arises in a natural way in the course of the analysis of groups of even type.

However, to keep this outline as simple as possible, we shall not present the general definition, but instead shall limit ourselves to a list of the simple groups that arise for the prime 2. Here then is the key definition we shall use.

DEFINITION. A simple K-group K is said to be of *weak characteristic 2 type* —for brevity, a W_2-*group*—provided K is of one of the following isomorphism types:

(1) A group of Lie type of characteristic 2;

(2) $PSL_2(q)$, q a Fermat or Mersenne prime or 9 (for brevity, $q \in \text{FM9}$);

(3) $PSL_3(3)$ ($\cong G_2(2)'$), $PSU_3(3)$, $PSL_4(3)$, $PSU_4(3)$, $PSp_4(3)$ ($\cong PSU_4(2)$), $G_2(3)$, or $PSO_8^+(3)$;

(4) A_5 ($\cong PSL_2(5) \cong PSL_2(4)$), A_6 ($\cong PSL_2(9)$), A_8 ($\cong PSL_4(2)$), A_9, or A_{10}, or

(5) M_{11}, M_{12}, M_{22}, M_{23}, M_{24}, J_2, J_3, J_4, HS, He, Suz, Ru, .2, .1, $M(22)$, $M(24)'$, F_5, F_3, F_2, or F_1.

More generally, a quasisimple group K with $O(K) = 1$ is called a W_2-group if $K/Z(K)$ is a W_2-group.

Schur multipliers of all the simple groups have been determined (see [**76**, Table 4.1]), so a complete list of W_2-groups is easily obtained.

Now we define a group of even type. A simple group G (of even order) is said to be of *even* type provided:

(1) $O(C_x) = 1$ for every involution x of G; and

(2) If K is a component of C_x for some involution x of G, then K is a W_2-group.

Henceforth the term *even* type (and its complement *noneven* type) will refer to a group satisfying the above conditions.

The implication of this change is that, to prove that a minimal counterexample G to the classification of the simple groups is of even type, *one will no longer need to validate the Brauer principle for Brauer components K with K a W_2-group except in certain special cases.* [These exceptions include the cases (a) $K \cong L_2(q)$ or $SL_2(q)$, $q \in$ FM9, and (b) $K \cong L_2(2^n)$, $L_3(2^n)$, or $PSp_4(2^n)$, $n \geqslant 2$, and $|G : N_K|$ odd.]

If this definition of a group of even type provides an effective replacement for a group of characteristic 2 type, as appears to be the case, it will by itself yield a major simplification in the classification proof, for the most difficult cases in validating the Brauer principle arise in connection with W_2-type Brauer components.

3. Small groups. We now give the precise definition of a small simple group G. There will be four kinds of smallness, two for the prime 2 and two for groups of even type.

As before, G is 2-*small* if $m_2(G) \leqslant 2$. These are divided into two kinds according as G has even or odd order.

Now let G be of even type. We denote by $\mathcal{M}(G)$ the set of 2-local subgroups of odd index in G. There will likewise be two kinds of "odd-small" simple groups, one in which $\mathcal{M}(G)$ has only one conjugacy class of maximal elements containing a Sylow 2-normalizer and another in which the elements of $\mathcal{M}(G)$ have low p-rank for all odd p.

We say that G is of *small parabolic* type if $\mathcal{M}(G)$ possesses a unique conjugacy class of maximal elements containing the normalizer in G of a Sylow 2-subgroup of G.

[The determination of the small parabolic simple groups of even type requires an extension to groups of even type of the present $C(G; S)$ theorem concerning groups of characteristic 2 type.]

Next, define the set $\sigma(G)$ of odd primes p by the following condition:

$p \in \sigma(G)$ provided either

$p \geqslant 5$ and $m_p(M) \geqslant 3$ for some $M \in \mathcal{M}(G)$, or

$p = 3$ and $m_3(M) \geqslant 4$ for some $M \in \mathcal{M}(G)$.

We say that G is of *small odd rank* if $\sigma(G)$ is empty and G is not of small parabolic type.

The definition differs in two ways from the existing notion of a "quasithin" group: first, the rank condition is restricted to 2-locals of odd index rather than being imposed on all 2-locals; and second, the case $m_3(M) = 3$ for some $M \in \mathcal{M}(G)$ is included among the small odd rank groups, whereas it was excluded from the earlier definition.

In our proposed unitary approach, the small odd rank simple groups of even type are to be analyzed by the Goldschmidt amalgam and parabolic system methods, which focus on 2-locals of odd index, so conditions on the elements of $\mathcal{M}(G)$ would seem to be sufficient. Concerning the second modification, Aschbacher's existing treatment of simple groups G of characteristic 2 type with $e(G) = 3$ and $m_{2,p}(G) \leqslant 2$ for all $p \geqslant 5$ [14] is extremely difficult, and there is good reason to believe that the parabolic approach to this additional case will yield a shorter proof.

A slight variation of these last conditions arises in connection with elimination of "uniqueness" 2-local subgroups in simple groups of even type, which is the basis of the following definition.

We include among the groups of *small odd rank* those groups G with $O_2(G) = 1$ (but not necessarily simple) such that $G = \langle P_1, P_2 \rangle$ for proper subgroups P_1, P_2 of G satisfying the following conditions:

$(*)$

(1) $P_1 \cap P_2$ contains a Sylow 2-subgroup of G;

(2) $C_{P_i}(O_2(P_i)) \leqslant O_2(P_i)$, $i = 1, 2$; and

(3) $m_p(P_i) \leqslant 2$ for all $p \geqslant 5$ and $m_3(P_i) \leqslant 3$, $i = 1, 2$.

[We have $P_i \leqslant N_G(O_2(P_i))$; however, we do not require P_i itself to be a 2-local subgroup, $i = 1, 2$.]

The parabolic methods use only the fact that $O_2(G) = 1$, rather than simplicity itself, and also do not require the generating parabolic subgroups to be local subgroups, so the conditions here, "localized" at P_1 and P_2, are essentially the same as those for simple groups G of even type with $\sigma(G)$ empty.

The justification for the definition is that the thrust of the analysis of a simple group G of even type having a "large" 2-local subgroup is to produce a subgroup G_1 of G satisfying all the conditions of $(*)$, with the corresponding subgroup P_1 contained in an element M of $\mathcal{M}(G)$, with M a uniqueness subgroup (see §4 for the definition). The goal of the amalgam and parabolic system methods is to force the possible isomorphism types of the parabolic P_1 (as well as of the corresponding parabolic P_2). If this can be achieved, the final step of the analysis will then be to show that the structure of P_1 is incompatible with that of the uniqueness 2-local M in which it is contained (by assumption, P_1 contains a Sylow 2-subgroup of M), thus eliminating the possibility of large 2-subgroups in simple groups of even type.

4. Uniqueness subgroups. The large subgroup problem arises in several distinct portions of the classification: in internally eliminating (a) the odd part of $\mathrm{Rad}(G)$ in groups of noneven type and (b) the 2-part of $\mathrm{Rad}(G)$ in groups

of even type, and also in internally constructing Soc(G) in both groups of noneven type and of even type. As usual, we shall consider the cases $p = 2$ and p odd, G of even type separately.

In the course of classifying groups with a proper 2-generated core, Aschbacher [4] established a basic sufficient condition for a subgroup M of a group G to be strongly embedded, which is often easier to verify than the fact that M contains the 2-generated core of G. This result forms the basis for the first type of uniqueness subgroup. To state it, we need the following preliminary definition.

If M is a subgroup of a group G and x an involution of M, we say that M *controls* the G-fusion of x in G provided whenever $x^g \in M$ for some $g \in G$, there is $h \in M$ such that $x^g = x^h$.

THEOREM. *Let G be a simple group, M a proper subgroup of G, and x an involution of M, and assume the following conditions hold*:
   (i) *M controls the G-fusion of x in M*;
   (ii) *$C_x \leqslant M$; and*
   (iii) *If for $g \in G$, $x^g \in C_M(x)$ with $x^g \neq x$, then $C_{xx^g} \leqslant M$.*
Under these conditions, M is strongly embedded in G.

[Conditions (1) and (2) together imply that M contains a Sylow 2-subgroup of G.]

For brevity, we call a subgroup M of G satisfying the conditions of the theorem *pre-strongly embedded* in G.

In this terminology, a proper subgroup M of G will be called a 2-*uniqueness* subgroup provided M is either pre-strongly embedded in G or M contains the 2-generated core of G.

In practice, we construct such 2-uniqueness subgroups M of three different types: (a) as an O-local subgroup, (b) as a 2-local subgroup, and (c) as the normalizer of a quasisimple subgroup of G. The first two types arise in connection with the internal elimination of Rad(G) and the third with the internal construction of Soc(G).

The notion of a uniqueness subgroup for odd primes arises only in simple groups G of even type in connection with internally eliminating the 2-part of Rad(G).

We need the following additional terminology. For $M \in \mathcal{M}(G)$, $\sigma(M)$ will denote the set of odd primes p such that either $m_p(M) \geqslant 3$, $p \geqslant 5$, or $m_3(M) \geqslant 4$, $p = 3$. Thus $\sigma(M) \subseteq \sigma(G)$.

Furthermore, for any odd prime p, and Sylow p-subgroup P of G, the subgroup

$$\langle N_Q \mid Q \leqslant P, m(Q) \geqslant 2 \rangle$$

is called the 2-*generated p-core* of G. This is the direct analogue of the 2-generated core for $p = 2$, and, as in that case, it is determined only up to conjugacy by the choice of P.

Now we can define the desired term. We call an element $M \in \mathscr{M}(G)$ a $\sigma(G)$-*uniqueness* subgroup provided the following conditions hold:

(1) $\sigma(M)$ is nonempty;

(2) M contains the 2-generated p-core of G for each $p \in \sigma(M)$; and

(3) If H is any 2-local subgroup of G such that $m_p(H \cap M) \geqslant 2$ for $p \in \sigma(M)$, then $H \leqslant M$.

Thus M is large in the sense that it contains many p-local subgroups of G for $p \in \sigma(M)$ and also many 2-local subgroups of G.

5. The sets $\mathscr{L}_p(G)$. Our analysis of K-local simple groups G that are not small focuses almost exclusively on the sets $\mathscr{L}(x)$ of components of $C_x/O_{p'}(x)$ as x ranges over the elements of G of prime order p for suitable p. We therefore introduce the fundamental sets of quasisimple groups

$$\mathscr{L}_p(G) = \{\, \mathscr{L}(x) \,|\, x \text{ ranging over the elements of } G \text{ of order } p \,\}.$$

There exist important criteria for $\mathscr{L}_p(G)$ to be nonempty, both when G is not 2-small and $p = 2$ and when G is of even type and $p \in \sigma(G)$ (in which case G is not odd-small). These criteria turn out to be directly related to the question of the existence of local uniqueness subgroups.

Thus, the results to be stated in this and the next section are proved under the appropriate one of the following assumptions:

(1) $p = 2$, $m_2(G) \geqslant 3$, and G does not contain an O-local uniqueness subgroup; or

(2) G is of even type with $\sigma(G)$ nonempty, $p \in \sigma(G)$, and no element of $\mathscr{M}(G)$ is a uniqueness subgroup.

The so-called "balance" theorem [**76**, Theorem 4.54] asserts:

THEOREM. *If G is not of characteristic 2 type, then $\mathscr{L}_2(G)$ is nonempty.*

The theorem thus shows that $\mathscr{L}_2(G)$ is nonempty in any simple group G of noneven type that is not of small 2-rank. This is a basic distinction between the 2-rank $\leqslant 2$ and 2-rank $\geqslant 3$ cases and provides a further reason why the groups of small 2-rank require separate treatment.

For groups of even type, we have the following similar result whose proof involves a slight modification of a theorem of Klinger and G. Mason concerning groups of characteristic 2 type [**80**]:

THEOREM. *If G is of even type, then $\mathscr{L}_p(G)$ is nonempty for each $p \in \sigma(G)$.*

The Klinger-Mason argument gives a somewhat sharper conclusion, which we shall need for our subsequent outline of the analysis of groups of even type. To state it, we denote by $\mathscr{B}^*(p)$ for any $p \in \sigma(G)$ the set of elementary abelian p-subgroups of G of maximal rank subject to lying in a 2-local subgroup of G. In view of the definition of $\sigma(G)$, for any $B \in \mathscr{B}^*(p)$, we have $m(B) \geqslant 3$ if $p \geqslant 5$ and $m(B) \geqslant 4$ if $p = 3$.

THEOREM. *If G is a simple group of even type, then for each $p \in \sigma(G)$, there is $B \in \mathscr{B}^*(p)$ and $b \in B^{\#}$ such that $\mathscr{L}(b)$ contains an element K with K not of Lie type of characteristic p.*

Their analysis focuses on elements $B \in \mathscr{B}^*(p)$ and 2-groups T of maximal order invariant under B. Under the assumption that the theorem is false, their argument first forces T to be of symplectic type for any such B and T. Then, applying the result with B and T chosen so that T/T' is of maximal order, it yields a contradiction.

6. Generic and nongeneric groups. The key to our reorganization of the classification proof is a division of the elements of $\mathscr{L}_p(G)$ into those that are *p-generic* and those that are not (here again G is a K-local simple group which is not small and either $p = 2$ or G is of even type and $p \in \sigma(G)$). When $\mathscr{L}_p(G)$ contains such a p-generic component, there exists a rather uniform method for studying the centralizers of elements of order p in G, valid for arbitrary primes, while in the contrary case, as with the small simple groups, special methods are again required.

Essentially, the p-generic groups are the groups of Lie type of characteristic relatively prime to p, apart from those of low Lie rank, the alternating groups of suitably high degree, and for small values of p, certain sporadic groups. However, unfortunately, the known simple groups seem to be too perverse to permit a simple conceptual definition that produces a perfect split between those types of components that are to be treated by the general method and those requiring special methods. The following definition, expressed by a single easily formulated condition, comes very close—only a few groups fall on the wrong side of the "divide."

First we call an element x of order p in a group X *half-p-central* provided a Sylow 2-subgroup of $C_X(x)$ is of index $\leqslant p$ in a Sylow p-subgroup of X (*p-central* if $C_X(x)$ contains a Sylow p-subgroup of X).

We now define a simple K-group K to be *p-generic* provided $C_K(x)$ has a component for some half-p-central element x of K. Furthermore, a quasisimple K-group K with $O_{p'}(K) = 1$ is called *p-generic* if $K/Z(K)$ is p-generic.

Here are the quasisimple groups that end up on the wrong side of the definition of 2-generic. First, the groups

$$Sp_4(3) \text{ and } SU_4(3)$$

are non-2-generic, but are to be treated by the 2-generic methods when they occur as elements of $\mathscr{L}_2(G)$. [The groups $PSp_4(3)$, $PSU_4(3)$, and $2PSU_4(3)$ are also non-2-generic, but they will be treated by the special non-2-generic methods.]

In the other direction, the groups

$$2A_9, 2A_{10}, 2A_{11}, \text{ and } M(23)$$

are 2-generic, but the first three can be handled in the same way as the group $2A_8$ when they arise as elements of $\mathscr{L}_2(G)$, and it is more efficient to include $M(23)$ within the non-2-generic analysis.

It will therefore be convenient to call the groups $Sp_4(3)$ and $SU_4(3)$ *ambiguous* non-2-generic groups and the groups $2A_9$, $2A_{10}$, $2A_{11}$, and $M(23)$ *ambiguous* 2-generic groups.

It is to be understood that the term 2-generic will henceforth include the ambiguous non-2-generic groups, but not the ambiguous 2-generic groups, with an analogous meaning for the term non-2-generic.

Similar imperfections occur for $p = 3$, but none arise for $p \geqslant 5$.

We also note the overlap between simple W_2-groups and simple 2-generic groups: namely, the groups

$$L_4(3), \ PSO_8^+(3), \ A_9, \ A_{10}, \text{ and } He.$$

Finally, we shall say that G is of *p-generic type* provided some element of $\mathcal{L}_p(G)$ is *p-generic*. Otherwise G will be called of *non-p-generic type*. [It is the groups of non-2-generic type that we referred to as of "intermediate" type at the end of Chapter I.]

7. The major case division. Using all the above terminology, Diagram 4 gives the basic case decomposition of our proposed unitary proof of the classification theorem. The arrows indicate the order in which the $p = 2$ and even type cases are to be considered, while the broken arrow indicates the place in the $p = 2$ analysis in which the even type alternative emerges.
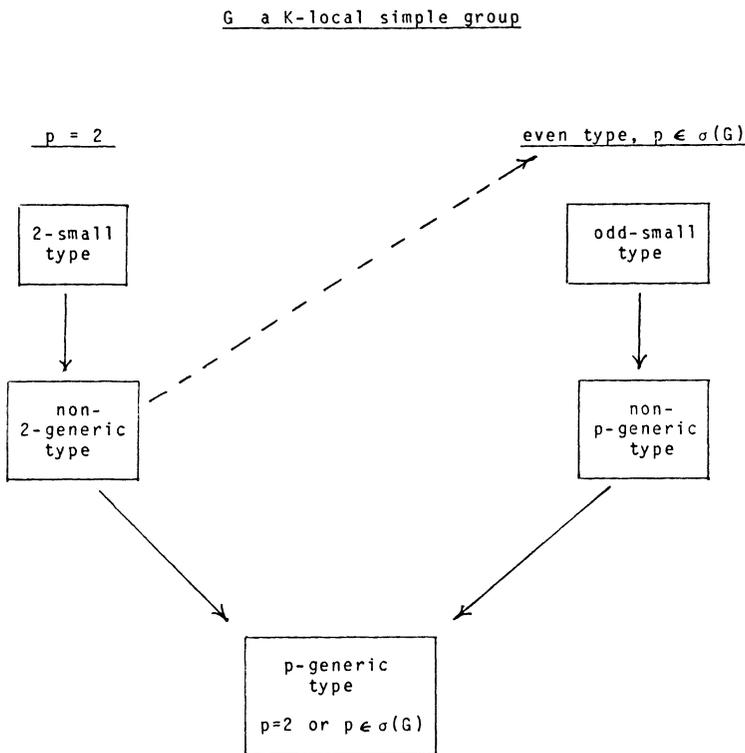
DIAGRAM 4. Major case decomposition of the classification.

However, we first emphasize that many of these cases must themselves be divided into subcases corresponding to the specific method of analysis to be used. For example, the 2-small groups divide into those of odd order and those of even order, since in the first case we must study *all* maximal subgroups of G, while in the second only those containing a Sylow 2-subgroup of G. Likewise the odd-small groups separate into those of small parabolic type, whose analysis involves Aschbacher's theory of so-called "χ-blocks" and an extension of the $C(G; T)$ theorem to groups of even type, and those of small odd rank, whose analysis is based upon the Goldschmidt amalgam method.

As a further illustration, the case in which G is of non-2-generic type divides into two subcases according as some element of $\mathscr{L}_2(G)$ is isomorphic to one of the groups

$$SL_2(q), \, q \text{ odd}, \quad 2A_n, 7 \leqslant n \leqslant 11, \quad \text{or } (X)L_3(4), \, X = 4, 4 \times 2, \text{ or } 4 \times 4,$$

or no element of $\mathscr{L}_2(G)$ is of this form. In the first case, core obstruction is handled by the Bender method, whereas in the second the key tool for eliminating core obstruction is the signalizer functor method.

G a K-local simple group, p = 2

Case	Solutions
2-small type $\longrightarrow$	$PSL_2(q), PSL_3(q), PSU_3(q),$ q odd, $PSU_3(4)$, A_7, M_{11}
non-2-generic type $\longrightarrow$	Even type or $G_2(q), PSL_4(q), PSU_4(q), PSp_4(q),$ $^2G_2(q)$, q odd, q > 3, $^3D_4(q)$, q odd, $A_9, A_{10}, A_{11}, J_1, Mc, ON, Ly, .3, M(23)$
2-generic type $\longrightarrow$	Lie type, odd char. not already listed and A_n, n ⩾ 13

TABLE 3

G a K-local simple group of even type, $p \in \sigma(G)$

Case		Solutions

odd-small type $\longrightarrow$

$PSL_2(q), PSU_3(q), Sz(q), q = 2^n, n > 3;$

$PSL_3(q), PSp_4(q),$

$G_2(q), {}^2F_4(q), q=2^n, n > 2; PSU_4(q),$

${}^3D_4(q), q=2^n, n > 2; PSL_4(2) (\cong A_8),$

$PSL_5(2), PSL_6(2), PSp_6(2), {}^2F_4(2)',$

$PSU_5(4), PSU_5(8), PSL_4(4), PSL_7(2), PSU_5(2),$

$PSU_6(2), PSp_8(2), PSp_6(4), PSO_8(2),$

$F_4(2), G_2(3), PSL_4(3), PSU_4(3), A_{10}, A_{12}, M_{12},$

$M_{22}, M_{23}, M_{24}, J_2, J_3, J_4, HS, He, Ru, .2, M(22)$

Suz, F_5, F_3

non-p-generic $\longrightarrow$

$M(24)', .1, F_2, F_1$

p-generic type $\longrightarrow$

Lie type, char. 2
not already listed

TABLE 4

Also, the subcase of the non-p-generic type case, $p \in \sigma(G)$, that leads to the four sporadic groups $M(24)'$, .1, F_2, F_1 is very interesting. It is here that we need the notion of a group of weak characteristic p-type—for brevity a W_p-group (a concept we have not defined here). Under the assumption that G is of even type, $3 \in \sigma(G)$, and every element of $\mathscr{L}_3(G)$ is a W_3-group, we use the Klinger-Mason methods to argue that the centralizer of a 2-central involution of G is isomorphic to that in one of the above four sporadic groups.

Finally, in the case G of p-generic type, $p = 2$ or $p \in \sigma(G)$, it is only the subcases in which some element of $\mathscr{L}_p(G)$ is a p-generic group of Lie type of characteristic prime to p or an alternating group that are treated for all primes simultaneously. Even in these subcases, however, certain parts of the analysis require us to distinguish further subcases for separate treatment.

Although the reader should be cognizant of the existence of this important finer subdivision of the proof throughout the ensuing discussion, Diagram 4

entirely suffices to give a picture of the global strategy of the proposed *GLS* classification proof.

To establish the classification theorem, it will be necessary (and sufficient) to determine the simple groups satisfying the assumptions of each case of the above decomposition. Tables 3 and 4 list the expected case-by-case solutions (including the even type solution in the non-2-generic case).

However, the tables require a word of explanation. First, there is some slight overlap among the separate cases. For example, the groups $PSL_2(q)$, $q \in$ FM 9, $PSL_3(3)$, $PSU_3(3)$, $PSU_3(4)$, and M_{11} are both 2-small and of even type. We prefer to include these solutions within the 2-small analysis. Similarly, the groups A_9, A_{10}, A_{12}, J_1, J_2, Mc, Ly, ON, and $M(23)$ are both non-2-generic and of even type. The way the analysis is organized, it is preferable to treat all but the cases leading to A_{12} and J_2 within the non-2-generic case. On the other hand, A_{12} and J_2 emerge most naturally from the small odd rank even type case.

One should also keep in mind that the term 2-generic includes the ambiguous non-2-generic groups, while the term non-2-generic includes the ambiguous 2-generic groups.

Finally, in the 2-generic type case, when we say "the groups of Lie type not already listed," we mean to exclude as well the groups $G_2(3)$, $PSL_4(3)$, $PSU_4(3)$, and $PSp_4(3)$, which are of odd-small even type and which will emerge as solutions from the Goldschmidt amalgam method.

8. The working hypothesis. In the remainder of Part IV we make a number of comments about various aspects of our proposed unitary project; in particular, we describe the steps involved in validating the Brauer principle. We begin with the general "ground rules" under which we plan to operate.

A. *An efficient proof*. The overriding aim of our global strategy is to produce as self-contained, elementary, and efficient a proof of the classification as presently available methods will allow. This principle requires amplification, for our present intention is to utilize results from a number of basic chapters of simple group theory. More specifically, we plan to operate under the assumption that the following results are available to us:

1. General techniques of finite group theory
2. Existence and uniqueness of the simple groups
3. Construction of the sporadic groups from the centralizers of their involutions
4. Solvability of groups of odd order
5. The classification of the groups of Lie type and Lie rank 1 as doubly transitive groups, and the Suzuki-Bender-Aschbacher theory of groups with a strongly embedded subgroup.

Each of these five fundamental topics would appear to be an essential component of any attempted classification of the simple groups. Perhaps, if our proposed unitary program is completed, we shall then consider the feasibility of developing a simplified treatment of one or more of them (if that has not already been done), but at present that prospect seems far off.

Most of the needed material for the first topic appears in the following books: Gorenstein [74] and a recent text of Aschbacher [18] for the foundations of local group theory; Isaacs [90] and Goldschmidt [72] for basic character theory (the latter book includes a proof of Glauberman's Z^*-theorem, which depends on some of Brauer's block theory results).

The algebraic group perspective, as developed by Steinberg [104, 105] provides a very satisfactory treatment of existence of the groups of Lie type and uniqueness for those of Lie rank ≥ 2. However, uniqueness for those of Lie rank 1 requires a (difficult) separate verification.

On the other hand, existence and uniqueness of the sporadic groups and their construction from the centralizers of their involutions is spread across a very large number of papers. A systematic treatment of this major chapter of simple group theory would be a very valuable contribution to the subject.

The solutions of certain centralizer-of-involution problems in simple groups of low 2-rank involve the same kind of character-theoretic analysis as the construction of some of the small sporadic groups. We shall therefore include these among our assumed results.

In other words, for the present our intent is to limit reorganization of the classification theorem to *that portion of the internal structure of the simple group G under investigation which can be achieved by local group-theoretic analysis* (and general techniques of finite group theory). In simple groups G of p-generic type, these methods suffice for the internal construction of $\mathrm{Soc}(G)$. However, in many remaining cases they yield only centralizer-of-involution information.

Thus, as we have indicated, the present status of these five major areas of simple group theory can be described as follows: many parts are in definitive form, in others revision is in progress, and others remain to be restudied.

But our point here is this: by a *self-contained* proof, it is to be understood that we mean *subject to the assumption that the principal results in each of these five areas of simple group theory are valid*. It also means that *no other major result about simple groups can be used in the unitary development without including its proof*.

Thus it is likewise to be understood that the terms *efficient* and *elementary* are to be construed as relative to the above constraints.

B. *The theory of K-groups.* A significant problem facing any contemplated revision of the classification theorem is how to handle the vast variety of properties of K-groups that will be needed for the proof. Some of these express very general properties of arbitrary K-groups or simple K-groups (often the precise result includes certain exceptional cases), while others give detailed information about a single group or family of groups (usually small K-groups).

For simple K-groups, they range from properties of their automorphism groups, central covering groups, tables of centralizers of elements of prime order, generation by such centralizers, conjugacy classes of elements of prime order, and p-rank, Sylow p-subgroup, and p-local structure for various primes p, as well as certain results concerning p-modular representations for appropriate primes p.

A prerequisite for achieving a satisfactory internal classification of the simple groups will be the development of a coherent *general theory* and *special*

theory of *K*-groups, bringing together in a systematic manner results which are presently rather chaotically scattered through the literature. Moreover, a major objective for an efficient proof should be to minimize the number of *K*-group properties required in the analysis.

For the groups of Lie type, the desired properties are established either using the theory of rational points of endomorphisms of algebraic groups or, in the case of the classical groups, from the geometry associated with their natural representations. For the alternating groups, they are usually proved by direct calculation. On the other hand, each sporadic group must be treated individually; in fact, many of the needed properties are established in the course of constructing the sporadic groups as solutions of appropriate centralizer-of-involution problems.

In implementing our proposed unitary proof, we have settled on the following procedure: In each chapter we shall list *without proof* those *K*-group properties needed for the results to be proved there. Then, after all the chapters have been written, we shall systematically verify the total resulting list of *K*-group properties. This "deferred" approach, although unsatisfactory in that it postpones completion of the principal results of the various chapters, has the advantage of avoiding duplication and will enable us to treat the theory of *K*-groups systematically as separate chapters of simple group theory.

9. *k*-balance and signalizer functors. The signalizer functors used in the classification proof are all defined in terms of "*k*-balance" for suitable *k*. We describe these briefly.

Let G be a simple group, p a prime, and A an elementary abelian p-subgroup of G. For each subgroup D of A, set

$$\Delta_G(D) = \bigcap_{d \in D^\#} O_{p'}(C_d).$$

For any positive integer k, if $m(A) \geqslant k$, we say that G is *k-balanced with respect to A* provided for each $D \leqslant A$ with $m(D) \geqslant k$ and each $a \in A^\#$,

$$\Delta_G(D) \cap C_a \leqslant O_{p'}(C_a),$$

and we say that G is *weakly k-balanced with respect to A* provided

$$[\Delta_G(D) \cap C_a, A] \leqslant O_{p'}(C_a).$$

Furthermore, if $m(A) \geqslant k + 1$, we say that G is $(k + \frac{1}{2})$-balanced with respect to A provided G is both $(k + 1)$-balanced and weakly k-balanced with respect to A.

Assume now that $m(A) \geqslant k + 2$ and that G is either k-balanced or $(k + \frac{1}{2})$-balanced with respect to A; for each $a \in A^\#$, correspondingly set

$$\theta(a) = \langle \Delta_G(D) \cap C_a \,|\, D \leqslant A, \, m(D) \geqslant k \rangle; \text{ or}$$

$$\theta(a) = \langle \Delta_G(D) \cap C_a, [\Delta_G(E) \cap C_a, D] \,|\, \rangle$$

$$\langle E \leqslant D \leqslant A, \, m(D) \geqslant k + 1, \, m(E) \geqslant k \rangle.$$

Under these conditions, one has the following basic result.

THEOREM. *If $\theta(a)$ is a K-group for each $a \in A^{\#}$, then θ is an A-signalizer functor on G.*

If each $\theta(a)$ is solvable, the result follows from the solvable signalizer functor theorem [**17, 63, 67, 68**]. In the general case it is a consequence of McBride's nonsolvable signalizer functor theorem [**108**], which extends the solvable theorem precisely to the case in which each $\theta(a)$ is a K-group. The signalizer functor theorem itself asserts:

THEOREM. *Under the same assumptions, the closure $\theta(G; A) = \langle \theta(a) | a \in A^{\#} \rangle$ is a p'-group.*

Corresponding to the two cases, we call θ the *k-balanced* or $(k + \frac{1}{2})$-*balanced* A-signalizer functor on G.

Walter and I introduced the notion of k-balance for integral values of k and, in particular, worked out the general theory of 1-balanced groups [**83, 84**]. Furthermore, 2-balanced functors (for $p = 2$) were used in many classification theorems. Subsequently, Goldschmidt [**70**] and Aschbacher [**7**] defined signalizer functors involving commutator conditions, which included the 3/2-balanced case. In particular, Aschbacher's proof of the classical involution theorem utilizes 3/2-balanced functors. After that, Lyons and I, adapting Aschbacher's construction of 3/2-balanced functors, defined 5/2-balanced functors, which we used in our analysis of centralizers of elements of odd prime order p in groups G of characteristic 2 type for $p \in \sigma(G)$ and $m_p(G) \geqslant 4$ [**71**]. In the case $p = 2$, Walter [**149**] has used yet another type of signalizer functor that had been earlier suggested by Goldschmidt, in which the subgroups $O_{2'}(C_d) = O(C_d)$ are replaced by the subgroups $O_r(C_d)$ for a suitable fixed odd prime r.

A key feature of the proposed unitary strategy is to use 3/2-*balanced signalizer functors* with respect to suitable elementary abelian p-subgroups in studying centralizers of elements of order p in K-local simple groups G of p-generic type for both $p = 2$ and $p \in \sigma(G)$. Our work to date indicates that this will yield a short, uniform treatment of the signalizer functor portions of the analysis of such simple groups. On the other hand, when G is of non-p-generic type, the analysis appears to require k-balanced functors for $k = 1, 2, 5/2$, and 3, again for suitable elementary abelian p-subgroups of G.

Because it was available, McBride's theorem was used in the existing classification proof for $p \in \sigma(G)$ in studying simple groups G of characteristic 2 type. However, by considering an inductive ordering on the primes in $\sigma(G)$ in terms of the rank of elements of $\mathscr{B}^*(p)$, it may be possible with some additional argument to avoid the use of nonsolvable signalizer functors in analyzing groups of even type; but the matter has not yet been resolved.

Finally, as we have observed earlier, it is the triviality of an appropriate signalizer functor that yields the internal elimination of $\text{Rad}(G)$. On the other hand, its nontriviality represents the first step in the construction of a uniqueness subgroup of G. Indeed, if G is k or $(k + \frac{1}{2})$-balanced with respect to the elementary abelian p-subgroup A of G of rank $\geqslant k + 2$, k a positive integer,

and $\theta(G; A)$ is a nontrivial p'-group, then as G is simple, $N_{\theta(G;A)}$ is a proper subgroup of G.

Moreover, if B is any subgroup of A of rank $k + 1$, an elementary argument, using the specified balance assumption on A, implies that

$$\theta(G; B) = \langle \theta(b) \mid b \in B^{\#} \rangle = \theta(G; A).$$

But elements of N_B clearly permute the $\theta(B)$ among each other, because $b \in B^{\#}$ by definition of θ and consequently $\theta(G; B) = \theta(G; A)$ is invariant under N_B. Hence we have

PROPOSITION. *If B is any subgroup of A of rank $k + 1$, then $N_B \leqslant N_{\theta(G;A)}$.*

Thus the signalizer functor theorem enables one to produce a proper subgroup $M = N_{\theta(G;A)}$ of G which contains the normalizer in G of certain subgroups of A. This is the first step in the process of showing that M is, in fact, strongly (or pre-strongly) p-embedded in G, which asserts that M contains the normalizers in G of *all* (or *most*) nontrivial p-subgroups of M.

10. Internally eliminating Rad(G). Internal elimination of Rad(G) in those cases that depend upon the Bender and/or signalizer functor methods gets to the heart of the study of the centralizers of prime order in simple groups. Table 5 lists the various such cases and the corresponding methods that are used to construct uniqueness subgroups under the assumption that the odd part or the 2-part of Rad(G) is nontrivial.

G a K-local simple group

Case		Methods
2-small, |G| even	$\longrightarrow$	Bender method + 2-fusion
G of non-p-generic type	$\longrightarrow$	Bender method, k-balanced signalizer functors, k = 1, 3/2, 2, 5/2, 3 + 2-fusion in the case p = 2
G of p-generic type	$\longrightarrow$	3/2-balanced signalizer functors

TABLE 5. Internal elimination of Rad(G) in cases involving Bender or signalizer functor methods.

I would like to discuss two subcases to highlight some of the important features of the analysis.

A. *The 2-generic Lie type case.* I shall outline the argument in the p-generic case in some detail. The subcase in which $p = 2$ and $\mathscr{L}_2(G)$ has a 2-generic component $\overline{K}$ of Lie type (of suitably high Lie rank) has some simplifying features not present in the other p-generic cases, so I shall limit the discussion

to this subcase. As we shall see, the argument involves a closely linked interplay between 3/2-balanced signalizer functors, L-balance, and particular properties of simple K-groups, including the B-property.

By definition of $\mathscr{L}_2(G)$, there is an involution $x \in G$ such that $\overline{K}$ is a component of $\overline{C}_x = C_x/O(C_x)$. Let K_0 be the preimage of $\overline{K}$ in C_x and let K be a *minimal normal* subgroup of K_0 that maps on $\overline{K}$. [K is said to be a 2-*component* of C_x. K is clearly perfect, and is, in fact, uniquely determined.]

By definition of K_0 and K, we have $K_0 = KO(C_x)$ and one easily proves the following result:

LEMMA 1. *K is quasisimple if and only if K centralizes $O(C_x)$.*

One of the main assertions of the internal elimination of Rad(G) is that K is, in fact, quasisimple; and it is a proof of this result we intend to outline. Assuming false, it follows from Lemma 1 that

(1) $$[K, O(C_x)] \neq 1.$$

We use (1) to construct a 2-uniqueness subgroup in G (in fact, a strongly embedded O-local subgroup), in which case the Aschbacher-Bender-Suzuki theorem will yield the desired contradiction.

Since $\overline{K}$ is of Lie type of odd characteristic of sufficiently high degree, $\overline{K}$ contains an E_8-subgroup $\overline{A}$ (i.e., $\overline{A} \cong Z_2 \times Z_2 \times Z_2$) with the following properties:

(1) $C_{\overline{K}}(\overline{A})$ contains a product $\overline{I}$ of components, each of which is nonsimple of Lie type of odd characteristic; and

(2) $\overline{A} \leqslant \overline{I}$ (whence $\overline{A} \leqslant Z(\overline{I})$).

Note that as the Schur multiplier of any simple K-group has 2-rank $\leqslant 2$, $\overline{A} \not\leqslant Z(\overline{K})$, so $\overline{I} < \overline{K}$.

Since K maps on $\overline{K}$, K contains an E_8-subgroup A mapping on $\overline{A}$. We argue first that G is 3/2-balanced with respect to A and then that the 3/2-balanced A-signalizer functor θ is nontrivial. It is here that the special choice of A enters. Indeed, let I_0 be the preimage of $\overline{I}$ in $C_K(A)$ (note that $C_K(A)$ maps onto $C_{\overline{K}}(\overline{A})$) and let I be a minimal normal subgroup of I_0 which maps on $\overline{I}$, so that I is a product of 2-components of $C_K(A)$ and

(2) $$A \leqslant Z(I).$$

We use (2), L-balance, and properties of the groups of Lie type of odd characteristic to prove

LEMMA 2. *G is 3/2-balanced with respect to A.*

PROOF. Here is the gist of the argument. Fix any $a \in A^{\#}$ and set $\tilde{C} = C_a/O(C_a)$. Then $I \leqslant C_a$ as $A \leqslant Z(I)$, and it follows by L-balance that $\tilde{I}$ is contained in a product $\tilde{J}$ of components of $\tilde{C}$. We choose $\tilde{J}$ with the minimal number of components, in which case clearly I centralizes no component of $\tilde{J}$.

Let $\tilde{H}$ be a component of $\tilde{J}$ and set $N^* = N_{\tilde{C}}(\tilde{H})/C_{\tilde{C}}(\tilde{H})$, so that N^* is isomorphic to a subgroup of the automorphism group of the simple group $H^* \cong \tilde{H}/Z(\tilde{H})$. Since $\tilde{A} \leqslant \tilde{J}$, $\tilde{A}$ induces inner automorphisms on $\tilde{H}$. Also $\tilde{a}$ clearly centralizes $\tilde{H}$. Hence the image A^* of $\tilde{A}$ in H^* is isomorphic to $Z_2 \times Z_2$, Z_2, or 1.

There exists a completely formal argument which reduces verification of 3/2-balance with respect to A to the following assertion:

For each $a \in A^{\#}$ and each such component $\tilde{H}$ of $\tilde{C}$ for which $A^* \cong Z_2 \times Z_2$, H^* is *locally 3/2-balanced* with respect to A^*—that is,

(a) $\bigcap_{a^* \in (A^*)^{\#}} O(C_{N^*}(a^*)) = 1$; and

(b) $[O(C_{N^*}(a^*)), A^*] = 1$ for each $a^* \in (A^*)^{\#}$.

But we can determine the possible isomorphism types of H^*. Indeed, the image I^* of $\tilde{I}$ in N^* is a product of components of $C_{H^*}(A^*)$, each nonsimple of Lie type of odd characteristic with $A^* \leqslant Z(I^*)$. Moreover, H^* is a K-group, as G is a K-local simple group. One checks now from the known structure of the centralizers of involutions of simple K-groups that H^* must be of Lie type of odd characteristic and Lie rank $\geqslant 2$.

Finally, it is a fact that every simple group of Lie type of odd characteristic and Lie rank $\geqslant 2$ is locally 3/2-balanced with respect to each of its four-subgroups, so the lemma follows.

The nontriviality of the 3/2-balanced A-signalizer functor θ depends, in particular, on properties of the action of A on $O(C_x)$. It will be convenient to describe these in general terms, as they will be critical at later stages of the argument as well.

A group X "acted on" by an E_8-group B will be said to be *well-generated* for B provided:

$$X = \langle C_X(D) \mid D \text{ ranging over all four-subgroups of } B \rangle.$$

We need the following result:

LEMMA 3. *If X is a group with $X/O(X)$ a product of $m \geqslant 0$ quasisimple components, each of Lie type of odd characteristic and Lie rank $\geqslant 2$, (excluding a few groups over $GF(3)$), then X is well generated for every E_8-group acting on X. In particular, this is the case if X is of odd order (in which case $m = 0$).*

Another elementary fact about solvable groups of odd order acted on by a 2-group B is the following:

(3) $$[X, B] = [[X, B], B].$$

With this information we can prove

LEMMA 4. *θ is nontrivial.*

PROOF. Again we present only the main points of the argument. Since $\overline{K}$ is quasisimple with $\overline{A} \leqslant \overline{K}$ and $\overline{A} \nleqslant Z(\overline{K})$, it follows from (1) that A does not centralize $O(C_x)$. But $O(C_x)$ is well generated for A, so there is $a \in A^{\#}$ such that

$$Y = \left[C_{O(C_x)}(a), A \right] \neq 1.$$

We need only show that $Y \leqslant O(C_a)$, for then $Y \leqslant [O(C_a), A] \leqslant \theta(a)$ by definition of θ, in which case θ is nontrivial. Furthermore, again putting $\tilde{C} = C_a/O(C_a)$, the desired conclusion will follow provided

(4) $$\tilde{A} \text{ centralizes } \tilde{Y}.$$

Let $\tilde{J}$, $\tilde{H}$, N^*, and H^* be as in Lemma 2. By (3), $\tilde{Y} = [\tilde{Y}, \tilde{A}]$; and as $\tilde{A} \leqslant \tilde{I} \leqslant \tilde{J}$, it follows that $\tilde{Y} \leqslant \tilde{J}$. But $Y \leqslant O(C_x)$, so $\tilde{Y} \leqslant O(C_{\tilde{J}}(\tilde{x}))$. Again, an entirely formal argument reduces the verification of (4) to the following assertion for each $\tilde{x}$-invariant component $\tilde{H}$ of $\tilde{J}$ (whence $x^* \in N^*$):

(5) A^* centralizes $O(C_{H^*}(x^*))$.

On the other hand, by definition of $\tilde{I}$, I^* is a product of components of $C_{H^*}(A^*\langle x^* \rangle)$, whence by L-balance I^* is contained in a product L^* of components of $C_{H^*}(x^*)$. But H^* is a K-group and the B-property holds in all simple K-groups. In particular, it implies that L^* centralizes $O(C_{H^*}(x^*))$. However, $A^* \leqslant I^* \leqslant L^*$, so A centralizes $O(C_{H^*}(x^*))$, as required.

Thus the closure $\theta(G; A)$ of θ is nontrivial and is of odd order by the solvable signalizer functor theorem. We let M be a maximal O-local subgroup of G containing $N_{\theta(G;\ A)}$. Since G is 3/2-balanced with respect to A and $N_{\theta(G;\ A)} \leqslant M$, we conclude that

(6) $N_D \leqslant M$ for each four-subgroup D of A.

But $\overline{K} = K/O(K)$ is of Lie type of odd characteristic and not one of the exceptions of Lemma 3, so K is well generated for A. It follows therefore from (6) that

LEMMA 5. *We have $K\langle x \rangle \leqslant M$.*

Setting $M^* = M/O(M)$, K^* is thus a component of $C_{M^*}(x^*)$, and we can identify its pump-up J^* in M^* as the product of one or two components of M^* of Lie type of odd characteristic (and Lie rank $\geqslant 2$). We let L^* be the product of all such components of M^*, so that L^* is normal in M^*, and we let L be the corresponding product of 2-components of M, so that likewise L is normal in M.

Furthermore, as $O(C_x)$ is well generated for A, $O(C_x) \leqslant M$, whence $[K^*, O(C_x)^*] \leqslant L^*$. But by the B-property in L^*, $[K^*, O(C_x)^*] = 1$. Since $[K, O(C_x)] \neq 1$, we thus conclude

LEMMA 6. *We have $O(L) \neq 1$.*

We are now almost in a position to show that M is strongly embedded in G. It will be convenient to say that an E_8-subgroup B of M is *M-closed* provided

$$C_E \leqslant M \text{ for every four-subgroup } E \text{ of } B.$$

In particular, A is M-closed by (6).

The following result will be critical.

LEMMA 7. *Let B be an M-closed E_8-subgroup of M and let $g \in G$. If $B^g \leqslant M$ and $B \leqslant M^g$, then $g \in M$.*

PROOF. Indeed, given the structure of L, L is well generated for B^g. But B^g is M^g-closed, so $L \leqslant M^g$. Similarly $L^g \leqslant M$. Using L-balance and the definition of L^*, one concludes now that, in fact, $L = L^g$. Hence $O(L) = O(L^g) = (O(L))^g$, whence $g \in N_{O(L)}$. But $O(L)$ is normal in M, as L is normal in

M and $O(L)$ is characteristic in L. Thus $M \leqslant N_{O(L)}$. Since $O(L) \neq 1$ by Lemma 6, and M is a maximal O-local subgroup, it follows that $M = N_{O(L)}$, whence $g \in M$, as asserted.

Lemma 7 has the following key consequence.

LEMMA 8. *If B is an M-closed E_8-subgroup of M and $B \leqslant Y \leqslant M$, then $N_Y \leqslant M$.*

PROOF. Indeed, if $g \in N_Y$, then $Y = Y^g \leqslant M^g$, so $B \leqslant M^g$. Likewise $B^g \leqslant Y^g = Y \leqslant M$, so $g \in M$ by the previous lemma. Since g was arbitrary, we conclude that $N_Y \leqslant M$.

As a first application, we have

LEMMA 9. *M contains a Sylow 2-subgroup S of G.*

PROOF. Expand A to a Sylow 2-subgroup T of M and T to a Sylow 2-subgroup S of G. Since A is M-closed, $N_T \leqslant M$ by Lemma 8. In particular, $N_S(T) \leqslant M$. But as T is Sylow in M, this forces $T = N_S(T)$. Since S is nilpotent, it follows that $T = S$, proving the lemma.

We describe the remaining steps without proofs; the arguments are similar in spirit to the preceding. First, Lemmas 8 and 9 imply that $N_S \leqslant M$. Hence to establish the strong embedding of M, it suffices to prove

(7) $C_t \leqslant M$ for every involution t of S.

To establish (7) for a given involution t, it is probably necessary only to show that C_t contains an M-closed E_8-subgroup B with $B \leqslant L$. However, there are then a few delicate configurations to analyze, which can be entirely avoided by imposing the following stronger conditions:

LEMMA 10. *Let t be an involution of S such that C_t contains an M-closed E_8-subgroup B of L and an involution u with the following properties*:
  (i) $C_u \leqslant M$; *and*
  (ii) $u^* \in J^*$ *for some component J^* of $C_{L^*}(t^*)$ with J^* of Lie type of odd characteristic other than $PSL_2(q)$.*
Under these conditions $C_t \leqslant M$.

The exclusion of $PSL_2(q)$, q odd, avoids certain unpleasant potential pump-ups (when $q = 5, 7$, or 9). The point of conditions (i) and (ii) is that they enable one to prove that M contains the normal closure Y of B in C_t (i.e., the smallest normal subgroup of C_t containing B). Since B is M-closed, it follows then from Lemma 8 that $C_t \leqslant N_Y \leqslant M$.

This immediately yields

LEMMA 11. *We have $C_a \leqslant M$ for every $a \in A^\#$.*

[A little extra work is required if $\overline{K}$ is defined over $GF(3)$, as $SL_2(3)$ is a solvable group.]

For brevity, call an E_8-subgroup B of M *strongly M-closed* if $C_b \leqslant M$ for every $b \in B^\#$ (in which case clearly B is M-closed). Thus A itself is strongly M-closed.

The idea now is, beginning with A, to produce other strongly M-closed E_8-subgroups B of M. For example, suppose B centralizes A and for each $b \in B^\#$, there exists an involution u of M (depending on b) such that condition (ii) of Lemma 10 holds with b, A in place of t, B. Then since $A \leqslant C_b$ and $C_u \leqslant M$, it follows from Lemma 11 that each $C_b \leqslant M$, whence B is strongly M-closed.

Using the structure of L, and repeating this argument successively, one can establish the following result:

LEMMA 12. *Every E_8-subgroup of L is strongly M-closed. In particular, $C_t \leqslant M$ for every involution t of L.*

The second assertion is a consequence of the first and the fact that $m_2(C_L(t)) \geqslant 3$ for every involution t of L.

Finally, if t is an arbitrary involution of S, then likewise $m_2(C_L(t)) \geqslant 3$, so C_t contains a strongly M-closed E_8-subgroup B of L. Furthermore, $C_{L^*}(t^*)$ contains a component J^* of Lie type of odd characteristic other than $PSL_2(q)$, and so $C_L(t)$ contains an involution u with $u^* \in J^*$, and by the previous lemma, $C_u \leqslant M$. Hence $C_t \leqslant M$ by Lemma 10 and so M is strongly embedded in G, as required.

REMARK. The p-generic cases in which $\overline{K}$ is an alternating group (with $p = 2$) or in which p is odd follow a similar pattern of proof. There is again an appropriate E_{p^3}-subgroup $\overline{A}$ of $\overline{K}$. However, in these cases there does not exist a product $\overline{I}$ of components of $C_{\overline{K}}(\overline{A})$ satisfying the critical inclusion $\overline{A} \leqslant Z(\overline{I})$. As a consequence, verification of $3/2$-balance as well as certain other steps in the proof require additional argument.

B. *Elimination of uniqueness elements of $\mathscr{M}(G)$.* Our second subcase will explain how the amalgam–parabolic system method enters into the elimination of uniqueness elements of $\mathscr{M}(G)$ when G is of even type with $\sigma(G) \neq \varnothing$.

The definition of such uniqueness subgroups is given in §4 of this chapter. To describe the situation we wish to consider, we need a preliminary definition.

Let $N \in \mathscr{M}(G)$ and let S be a Sylow 2-subgroup of N, so that S is a Sylow 2-subgroup of G by definition of $\mathscr{M}(G)$. Set

$$\mathscr{S}(N; S) = \{ K \mid S \leqslant K \leqslant N \text{ and } C_K(O_2(K)) \leqslant O_2(K) \}.$$

[If $C_N(O_2(N)) \leqslant O_2(N)$, then, of course, $N \in \mathscr{S}(N; S)$.]

We assume the following conditions are satisfied:

(1) If M is any maximal element of $\mathscr{M}(G)$, and S a Sylow 2-subgroup of M, then $M = \langle K \mid K \in \mathscr{S}(M; S) \rangle$;

(2) If $\sigma(M) \neq \varnothing$ and M is as in (1), then M is a uniqueness subgroup and for each $p \in \sigma(M)$, $m_p(K) \geqslant 2$ for some $K \in \mathscr{S}(M; S)$; and

(3) G is not of small parabolic type.

In particular, conditions (1) and (2) automatically hold if $C_M(O_2(M)) \leqslant O_2(M)$ for every maximal element M of $\mathscr{M}(G)$.

Since $\sigma(G) \neq \varnothing$, there exists some $M \in \mathscr{M}(G)$ with $\sigma(M) \neq \varnothing$, and without loss we can assume that M is a maximal 2-local subgroup. Since $M \in \mathscr{M}(G)$, M contains a Sylow 2-subgroup S of G. Furthermore, as G is not of small

parabolic type, it follows by Sylow's theorem that M is not the unique maximal 2-local subgroup of G containing S. Thus there is a maximal element $N \in \mathscr{M}(G)$ with $S \leqslant N$ and $N \neq M$.

We consider subgroups M_1, N_1 of M, N, respectively, defined as follows:

(1) Choose $M_1 \in \mathscr{S}(M; S)$ minimal subject to $m_p(M_1) \geqslant 2$ for some $p \in \sigma(M)$;

(2) *Case* 1. If $\sigma(N) \neq \varnothing$ (in which case also N is a uniqueness subgroup), choose $N_1 \in \mathscr{S}(N; S)$ minimal subject to $m_q(N_1) \geqslant 2$ for some $q \in \sigma(N)$;

Case 2. If $\sigma(N) = \varnothing$, choose N_1 minimal subject to $N_1 \not\leqslant M$.

Since $N = \langle K \mid K \in \mathscr{S}(N; S) \rangle$ and $N \neq M$ with N, M maximal 2-local subgroups of G, clearly such a subgroup N_1 of N exists in Case 2.

We then have the following key fact.

LEMMA. *If* $G_1 = \langle M_1, N_1 \rangle$, *then* $O_2(G_1) = 1$.

PROOF. Suppose false and let H be a 2-local subgroup of G containing $N_{O_2(G_1)}$. Then $\langle M_1, N_1 \rangle \leqslant H$. But $m_p(H \cap M) \geqslant m_p(M_1) \geqslant 2$, so by condition (3) in the definition of a $\sigma(G)$-uniqueness subgroup, $H \leqslant M$ and consequently $N_1 \leqslant M$. In particular, Case 2 does not hold, so $\sigma(N) \neq \varnothing$ and $m_q(N_1) \geqslant 2$ for some $q \in \sigma(N)$. Thus $m_q(N \cap M) \geqslant m_q(N_1) \geqslant 2$, so again by condition (3) (with N, M, q in the roles of M, H, p, respectively), we conclude that $M \leqslant N$, whence $M = N$ (again as each is a maximal 2-local), contrary to the choice of N. This completes the proof.

Since $O_2(G_1) = 1$, and $C_{M_1}(O_2(M_1)) \leqslant O_2(M_1)$, $C_{N_1}(O_2(N_1)) \leqslant O_2(N_1)$ (by definition of $\mathscr{S}(M; S)$, $\mathscr{S}(N; S)$), the amalgam–parabolic system method is applicable to G_1. Moreover, because of our minimal choice of M_1, N_1, there is considerable constraint on the p-structure of M_1, and in Case 1 on the q-structure of N_1, while in Case 2 $M_1 \cap N_1$ is necessarily the unique maximal subgroup of N_1 containing S. In addition, condition (3) in the definition of a $\sigma(G)$-uniqueness subgroup is very effective in the analysis.

As we have noted earlier, the aim of the analysis is to reduce the problem to showing that G_1 is of small odd rank (i.e. to proving that $m_r(M_1) \leqslant 2$, $m_r(N_1) \leqslant 2$ for all primes $r \geqslant 5$ and that $m_3(M_1) \leqslant 3$, $m_3(N_1) \leqslant 3$) and then to determine the possibilities for the "parabolics" M_1 and N_1, thereby forcing the structure of S. Finally it is shown that the uniqueness subgroup M cannot possess a Sylow 2-subgroup of the shape of S. [Stroth has carried this out in the case that M_1, N_1 are solvable [106] and is presently at work on the general case.]

11. Brauer form. As we have observed earlier, internal elimination of Rad(G) yields a first approximation of the centralizer of some element of order p in G with that of a corresponding centralizer in some known simple group G^* and represents the first of three stages in the internal construction of Soc(G).

We shall describe the results obtained at the first stage, which we have referred to as Brauer form, in the p-generic case and shall add some brief comments about the shape of Brauer form in the remaining cases. The exact definition of Brauer form in the p-generic case involves the notion of the "neighbor" of a component of a suitable element of order p.

However, to define this notion, we need some preliminary general terminology. First, if $x \in G$ is of order p and K is a component of C_x of order divisible by p (whence the image of K in $C_x/O_{p'}(C_x)$ is an element of $\mathscr{L}_p(G)$), then K is said to be *terminal* in G (for the prime p) provided K is a component of C_y for every element y of order p in C_K.

By Sylow's theorem, it suffices to impose the condition on a Sylow p-subgroup P of C_K with $x \in P$. Moreover, if $m_p(C_K) = 1$, then x is the unique element of order p in P and, as K is a component of C_x, K is automatically terminal in this case, which will explain why the concept of terminality was not needed in our earlier examples of Brauer form (Part II, §3).

For $p = 2$, terminality is very close to Aschbacher's notion of a *standard* component, which is defined by the conditions:

(1) If $H = C_K$, then $|H \cap H^g|$ is odd for all $g \in G - N_K$;

(2) K commutes elementwise with none of its G-conjugates; and

(3) $N_K = N_H$.

The following result of Gilman [**60**], which formalizes a portion of Aschbacher's argument in [**6**], gives the connection between terminal and standard components.

THEOREM. *If x is an involution of G and K is a component of C_x with K terminal in G and $m_2(K) \geqslant 2$, then K is standard in G.*

The case $m_2(K) = 1$ is exceptional. In that case K has quaternion Sylow 2-subgroups and $K/O(K) \cong SL_2(q)$, q odd, $q \geqslant 5$, or $2A_7$, the double cover of A_7.

In view of conditions (1) and (3), Aschbacher defines a subgroup X of even order in an arbitrary group G to be *tightly embedded* if $|X \cap X^g|$ is odd for all $g \in G - N_X$ ($|H|$ is even in (1), as $x \in H$).

Tightly embedded subgroups play a central role in the existing classification proof. For example, the general form of the classical involution theorem includes a determination of all simple groups containing a tightly embedded subgroup with quaternion Sylow 2-subgroups. Moreover, properties of tightly embedded subgroups of both arbitrary groups and of almost simple K-groups are critical for validation of the Brauer principle for centralizers of involutions in Brauer form.

Terminality is the basic notion in our proposed unitary approach. Indeed, our investigations to date indicate that very few properties of tightly embedded subgroups are needed for the internal construction of $\mathrm{Soc}(G)$ by the "neighbor" method.

Next, consider an arbitrary quasisimple K-group L and set $\overline{L} = L/Z(L)$. We say that a quasisimple subgroup I of L is *subterminal* in L for the prime p if I is a component of $C_L(u)$ for some element u of order p in L and $\overline{I}$ is terminal in $\overline{L}$.

If L is simple, then, of course, a subterminal component is terminal in L. We have the following fact.

PROPOSITION. *If L is p-generic, then L contains a subterminal component for the prime p.*

Now let K be a terminal component in G for the prime p with K p-generic, so that by the proposition K contains a subterminal component for p. If I is any such subterminal component of K and y is any element of C_I of order p that induces a nontrivial inner automorphism on K, then the pump-up J of I in C_y (see Chapter II, §5) is called a *neighbor* of K in G.

Since y induces a nontrivial inner automorphism on K and I is subterminal, it follows that I is a component of $C_K(y)$ and that y centralizes some element z of order p in C_K. Also, as K is terminal, K is a component of C_z. But now as y centralizes z, we see from the definition that I does, in fact, have a pump-up in C_y.

Furthermore, by the exact assertion of L-balance, J is z-invariant, $\bar{J} = J/O_{p'}(J)$ is either quasisimple or the product of p components cycled by z, and in either case $\bar{I}$ is a component of $C_{\bar{J}}(y)$.

If J itself is a product of (one or p) quasisimple components, we say that J is *semisimple*.

We are at last in a position to give the exact definition of Brauer form in the p-generic case. Let x be an element of prime order p in G, where either $p = 2$ or G is of even type with $p \in \sigma(G)$. We say that C_x is in *Brauer form* with *Brauer component* K provided the following hold:

(1) K is p-generic;

(2) K is terminal in G; and

(3) Every neighbor of K is semisimple.

The significance of the neighbor is that for a suitable choice of x and K it is ultimately shown that $\mathrm{Soc}(G) = \langle K, J \rangle$, where J is an appropriate neighbor of K. Indeed, the second stage in its construction proves the existence of a *quasisimple* neighbor J of K which contains I properly, so that K is a proper subgroup of $\langle K, J \rangle$. Finally, it is argued that $G = \langle K, J \rangle$.

On the other hand, in the non-p-generic case the notion of neighbor does not enter into the definition of Brauer form. Instead, it is defined by the condition that for $x \in G$ of order p, C_x contains a component K with K terminal in G and non-p-generic; but some constraint is usually placed on the element x. In particular, when G is of even type with $p \in \sigma(G)$, x is restricted to be in an element $B \in B_p^*(G)$. [It is shown that B, x, and K satisfying these conditions always exist in this case.] However, when G is 2-small and the centralizer of some involution x of G has a component K, K is necessarily terminal in G; in this case any such x, K define Brauer form.

Not all simple groups, however, possess elements of order p whose centralizers have components. In particular, since components are nonsolvable, this is the case if the centralizer of every element of order p is solvable. Hence such groups (for example, N-groups) do not possess Brauer components. Likewise it is true for all 2-small simple groups other than $PSL_3(q)$, $PSU_3(q)$, $q > 5$ (for $p = 2$).

In such cases Brauer form must therefore be defined in terms of conditions on the *solvable* group C_x, x an involution of G. However, as these are rather technical, we omit the precise definition. The point is that Brauer form in these cases expresses the conclusions concerning the structure of C_x implied by successful application of the Bender method.

On the other hand, when G is odd-small of even type, elimination of the 2-part of $\text{Rad}(G)$ involves an analysis of the 2-local subgroups of G containing a Sylow 2-subgroup (i.e., the elements of $\mathcal{M}(G)$) by the Aschbacher block method in the small parabolic case and by the Goldschmidt amalgam–parabolic subgroup method in the small odd rank case. In these cases, Brauer form expresses the conclusions which these respective analyses yield. In particular, it involves assertions about the structure of 2-local subgroups rather than of centralizers of elements of odd prime order. Again we omit the precise definitions.

It should also be emphasized that many of the possible Brauer components, in fact, correspond to no simple K-group. Such components are eliminated during later stages of validation of the Brauer principle.

12. Strong Brauer form. In the second stage of the internal construction of $\text{Soc}(G)$, which we refer to as *strong Brauer form*, the critical subgroups of G are shown to have a much tighter structure than that expressed by Brauer form itself, one which more closely approximates that of the corresponding configurations in simple K-group G^*, and we denote the resemblance by the symbol $G \approx G^*$. We note also that most of the configurations for which there exists no such G^* are eliminated during this stage.

In the preceding section, we have already intimated the definition of $G \approx G^*$ in the p-generic case, but we formalize it now. Note that the possibilities for G^* in this case consist of (a) the general groups of Lie type of odd characteristic plus the alternating groups of sufficiently high degree when $p = 2$, and (b) the general groups of Lie type over $GF(2^n)$ (with either p dividing the order of a Cartan subgroup of G^* or $p = 3$ and $n = 1$) when $p \in \sigma(G)$.

First of all, there may exist more than one element x of order p in G for which C_x is in Brauer form with Brauer component K. We begin by ordering the set of all such x, K, choosing them so that

(1) $|K|$ is maximal; and

(2) Subject to (1), $m_p(C_K)$ is maximal.

Moreover, it is only for such a maximal x, K in Brauer form that we establish strong Brauer form.

In this case, let J be a neighbor of K in G, so that J is the pump-up in C_u of a subterminal component I of K for a suitable element u of order p in K. We call J *proper* if J is quasisimple and $I < J$.

With this terminology, we say that G is in *strong Brauer form G^** and write $G \approx G^*$, G^* one of the simple K-groups described above, provided C_{x^*} is in Brauer form with Brauer component K^* for a suitable element x^* of G^* of order p and the following conditions hold:

$$(*) \qquad \begin{array}{ll} (1) & K \cong K^*; \text{ and} \\ (2) & K, K^* \text{ possess proper neighbors } J, J^* \text{ in } G, G^*, \\ & \text{respectively, with } J \cong J^*. \end{array}$$

We refer to the group G^* as the *approximate* of G.

Thus strong Brauer form in the p-generic case is expressed by a precise resemblance of the centralizers of *two* elements x, u of order p in G with those of suitable corresponding elements x^*, u^* of its approximate group G^*. We refer to it as strong *neighbor* Brauer form.

We note that the cases in which no approximate G^* of G exists—such as when $K \cong 2A_n$, $n \geqslant 12$ (for $p = 2$), or when the pair K and proper neighbor J of K do not occur in any known simple group (for $p = 2$ or for $p \in \sigma(G)$)—are eliminated primarily by an analysis of p-fusion.

However, there is an important such case that is not removed until the next stage of the analysis. Indeed, the case $K \cong A_n$ has two potential solutions: first, the expected solution $G^* \cong A_{n+4}$, in which $C_K \cong Z_2 \times Z_2$, $I \cong A_{n-4}$, and the proper neighbor $J \cong A_n$; and second, one which leads to the *nonsimple* solution $G^* \cong \Sigma_{n+2}$, in which $C_x \cong Z_2 \times \Sigma_n$ (and $C_K \cong Z_2$), and we take $u \in C_x$ to act as a transposition on K (whence $u \notin K$ in this one case). This time $I \cong A_{n-2}$ and the pump-up J of I in C_u is likewise isomorphic to A_n. We shall write $G \approx \Sigma_{n+2}$ in this case and shall call Σ_{n+2} the approximate of G.

When G is of non-2-generic type or of 2-rank 2 and C_x is in Brauer form with Brauer component K, the definition of strong Brauer form differs from that in the 2-generic case: in place of the existence of a proper neighbor are the following conditions:

(∗∗)

    (1)   A Sylow 2-subgroup S of G is isomorphic to a Sylow 2-subgroup S^* of its approximate G^*;

    (2)   Two involutions t_1, $t_2 \in S$ are conjugate in G if and only if the corresponding involutions t_1^*, t_2^* of S^* are conjugate in G^* (i.e., G and G^* have identical *involution fusion patterns*); and

    (3)   If t is any involution of S, C_t closely resembles C_{t^*}, where t^* is the corresponding involution of S^*.

Here the set of approximates G^* consists of the groups of Lie type of odd characteristic and Lie rank $\leqslant 2$ (apart from the groups $PSL_2(q)$ and most of those defined over $GF(3)$); A_n, $n = 9, 10, 11$; J_1, Mc, Ly, ON, $.3$, and $M(23)$.

The meaning of "closely resembles" in (3) varies according as G^* is or is not of Lie type of odd characteristic. Indeed, in the latter cases we, in fact, require C_t to be isomorphic to C_{t^*}; while in the former cases, if L^* is a component of C_{t^*} (whence $L^* \cong L_2(r^*)$ or $SL_2(r^*)$ for suitable odd r^*), we require only that C_t have a component $L \cong L_2(r)$ or $SL_2(r)$, respectively, with L and L^* having isomorphic Sylow 2-subgroups. It is only at the next stage of the analysis that the equality $r = r^*$ is achieved.

Likewise we must carry along two further possibilities related to the cases $G^* \cong Mc$ or Ly (in which cases $C_{t^*} \cong 2A_8$ or $2A_{11}$, respectively, for every involution t^* of G^*). Indeed, the possibilities $C_t \cong 2A_9$ or $2A_{10}$ for each involution t of S (whence correspondingly S is isomorphic to a Sylow 2-subgroup of Mc or Ly, respectively) are not eliminated until later in the analysis.

We refer to the conditions of (**) as strong *centralizer-of-involution* Brauer form.

For $p = 2$, it remains to describe strong Brauer form when G is small and the approximate G^* of G is isomorphic to $PSL_2(q)$, q odd; $U_3(4)$, A_7, or M_{11}, in which cases C_{t^*} is *solvable* for every involution t^* of S^*. Apart from the $U_3(4)$ case, the definition is again expressed by (**) with $C_t \cong C_{t^*}$ when $G^* \cong A_7$ or M_{11}. However, when $G^* \cong PSL_2(q)$, q odd, (whence S and S^* are dihedral), we require only the following:

(1) $O(C_t)$ is abelian and inverted by every involution of $C_t - \langle t \rangle$; and

(2) $N_Y \leqslant C_t$ for every nontrivial subgroup Y of $O(C_t)$.

[In $PSL_2(q)$, $O(C_{t^*})$ itself is cyclic of prescribed order depending on q, but these conditions are not verified for $O(C_t)$ until the subsequent analysis.]

On the other hand, it is preferable to define the term $G \approx PSU_3(4)$ as part of the case that G^* is any group of Lie type over $GF(q)$, $q = 2^n$, of Lie rank 1: namely, the families $PSL_2(q)$, $PSU_3(q)$, and $Sz(q)$. Moreover, we note that these are precisely the possible solutions for G when G is of small parabolic even type.

We shall write $G \approx G^*$ for one of the groups G^* in these three families provided the following condition holds:

$$(***) \qquad \begin{array}{l} N_S \text{ is the unique maximal 2-local subgroup} \\ \text{of } G \text{ containing the Sylow 2-subgroup } S \text{ of } G. \end{array}$$

In this case, we say G is of strong *rank 1 parabolic* Brauer form. Again the exact structure of S is not determined until the next stage.

We are left to define strong Brauer form when G is of even type and either odd-small or of non-p-generic type for $p \in \sigma(G)$. In the first case, the table of solutions for G is very extensive, including the groups of Lie type of characteristic 2 and Lie rank 2; a large number of individual groups of Lie type defined over $GF(q)$, $q = 2$, 3, 4, or 8, of Lie rank 3 or 4; and no less than fifteen sporadic groups.

When G^* is one of the above groups of Lie type of characteristic 2 and Lie rank 2, we write $G \approx G^*$ provided the maximal 2-local subgroups of G containing S closely approximate those containing S^*: more precisely, provided there exist two maximal elements M_1, M_2 of $\mathcal{M}(G)$ containing S and two maximal elements M_1^*, M_2^* of $\mathcal{M}(G^*)$ containing S^* such that

$$(****) \qquad \begin{array}{ll} (1) & O^{2'}(M_i) = O^{2'}(M_i^*), \ i = 1, 2; \text{ and} \\ (2) & O^{2'}(M_1 \cap M_2) = O^{2'}(M_1^* \cap M_2^*). \end{array}$$

In this case, we say that G is of strong *rank 2 parabolic* Brauer form.

We note that this case includes the nonsimple possibility $G^* \cong {}^2F_4(2)$; it, too, is not removed until the next stage.

On the other hand, when G^* is of Lie type above of Lie rank 3 or 4, or sporadic, or ${}^2F_4(2)'$, strong Brauer form for G differs dramatically from Brauer form. Indeed, in the first case the 2-local information provided by the conclusions of the Goldschmidt amalgam method are used to determine the structure of the centralizers of suitable elements of order 3 in G and to show

that, in fact, G satisfies the conditions of strong neighbor Brauer form for $p = 3$ with respect to its approximate G^*. Similarly, in the sporadic and $^2F_4(2)'$ cases the given information is used to determine the involution fusion pattern of G and the structure of the centralizers of involutions of G, so that in these cases it is strong centralizer-of-involution Brauer form that defines the relationship $G \approx G^*$.

Hence, in these cases, 2-local information is used to determine, on the one hand, centralizers of elements of order 3 and, on the other, centralizers of involutions.

Strong Brauer form in the remaining non-p-generic type case with $p \in \sigma(G)$ likewise involves a shift of primes. Here the possibilities for G^* are the four sporadic groups $M(24)'$, .1, F_2, and F_1.

Recall that in this case Brauer form requires $x \in B$ for some $B \in \mathscr{B}^*(p)$. The analysis depends on a further elaboration of the Klinger-Mason method [96], involving an interesting interplay between centralizers of elements of order p and centralizers of involutions. All cases with $p > 3$ and all but four cases with $p = 3$ lead to contradictions. In the latter four cases the analysis forces the exact structure of the centralizers of involutions, including determination of a Sylow 2-subgroup S of G as well as the involution fusion pattern, identical to those in any one of the groups $G^* = M(24)'$, .1, F_2, F_1. Hence in these cases the term $G \approx G^*$ is again expressed in terms of strong centralizer-of-involution Brauer form.

We have presented this amount of detail about strong Brauer form because it is at this stage of the analysis that the internal structure of our abstract K-local simple group G is first shown to closely approximate that of some known simple group G^*. Moreover, as we have seen, this approximation is expressed by one of four sets of conditions, which define the respective terms:

($*$) Strong neighbor Brauer form.
($**$) Strong centralizer-of-involution Brauer form.
($***$) Strong rank 1 parabolic Brauer form.
($****$) Strong rank 2 parabolic Brauer form.

13. Internal definition of $\mathrm{Soc}(G)$. The next major step is to describe $\mathrm{Soc}(G)$ internally and to verify that $G = \mathrm{Soc}(G)$. It turns out that there are two essentially distinct types of such descriptions of $\mathrm{Soc}(G)$, corresponding to the cases in which (A) G is of strong neighbor or rank 2 parabolic Brauer form, or (B) G is of strong centralizer-of-involution or rank 1 parabolic Brauer form.

(A) *Strong neighbor or rank 2 parabolic Brauer form.* In the first case, let x, K, u, J express strong neighbor Brauer form, and in the second, let M_1, M_2 be the two maximal 2-local subgroups of G containing the Sylow 2-subgroup S of G and expressing strong rank 2 parabolic Brauer form. Correspondingly set

$$\mathrm{Soc}(G) = \langle K, J \rangle \quad \text{or} \quad O^{2'}(\langle M_1, M_2 \rangle).$$

Our task now is to prove that the group $G_0 = \mathrm{Soc}(G)$, as just defined, is equal to G. In the next section, we shall explain the uniqueness of the multiplication table of G_0, a result which is valid whether or not $G_0 = G$, and which will imply that $G_0 \cong G^*$, where G^* is the approximate of G. However,

for expository purposes, it is preferable to assume, when $G_0 < G$, that G_0 is a K-group (even though G_0 is not a local subgroup), in which case the conclusion $G_0 \cong G^*$ can be checked directly.

Since G is simple and $G_0 < G$, $G_1 = N_{G_0}$ is likewise a proper subgroup of G. The goal is to verify that G_1 is pre-strongly embedded in G (for the prime 2), in which case G will be determined by the Aschbacher-Bender-Suzuki theorem, yielding an immediate contradiction, and thus forcing the desired conclusion $G_0 = \mathrm{Soc}(G) = G$. The argument is entirely local group-theoretic. Because we know the isomorphism type of G_0, it is not difficult to establish sufficient information about the 2-structure of G in either case to establish the required properties of G_1.

(B) *Strong centralizer-of-involution or rank* 1 *parabolic Brauer form.* In contrast to the previous cases, the goal of the analysis here is to prove that G has the same *order* as its approximate G^*. Once this is achieved, we *define*

$$G = \mathrm{Soc}(G).$$

When G has more than one conjugacy class of involutions, $|G|$ is computed from the so-called "Thompson order formula," which is an expression for $|G|$ involving only the structure of centralizers of involutions and the involution fusion pattern of G (see [76, Theorem 2.43]). The formula itself is established by elementary Brauer-Fowler type arguments.

The determination of $|G|$ is considerably more complicated when G has only one class of involutions, requiring a combination of local analysis, involution counts, and in many cases character theory as well as a theorem of Frobenius which gives a *congruence* for $|G|$. In a few cases, to turn this congruence into an equality, one needs certain deep character-theoretic results of Brauer concerning groups whose orders are divisible by primes r to the first power only, which depend on an analysis of the r-blocks of characters of G.

When G^* (under (B)) is of Lie type over $GF(2)$, alternating, or sporadic, the analysis yields directly that $|G| = |G^*|$. However, in the remaining Lie type cases, it yields only that G has the same order as some subgroup of $\mathrm{Aut}(G^*)$ containing G^* as a subgroup of odd index, and it requires additional transfer arguments to force $|G| = |G^*|$.

Furthermore, we note that it is in this portion of the analysis that the cases in which G has no approximate G^* are eliminated (in particular, the cases $C_x \cong 2A_n$, $n = 9$ or 10).

Hence, at this stage of the classification proof, we have reached an almost complete description of the internal structure of $G = \mathrm{Soc}(G)$, very closely approximating that of a suitable (almost) simple K-group G^*. It remains only to prove that G is, in fact, isomorphic to G^*.

14. The identification of G. At this point, we have the following internal data concerning G:

$$G = \mathrm{Soc}(G);$$

and one of the following holds:

A. $G = \langle K, J \rangle$ or $O^{2'}(\langle M_1, M_2 \rangle)$ (with G^* of Lie type of Lie rank $\geqslant 2$ or an alternating group of degree $\geqslant 12$); or

B. $|G|$ is completely specified.

The final step in the internal analysis is to show that the multiplication table of G is *uniquely* determined by the above conditions. Since the approximate G^* of G satisfies these conditions, it will then follow that $G \cong G^*$. This will be achieved by forcing G to have either a uniquely determined presentation by generators and relations, primitive permutation representation, or matrix representation over a suitable field, or to be a transitive automorphism group of a suitable graph.

It is the case $G = \langle K, J \rangle$ which, of course, yields the p-generic simple groups. The arguments are completely local group-theoretic. We separate the Lie type and alternating component cases.

Case A_1. G^* of Lie type of Lie rank ≥ 3. Here the goal is to show that G satisfies the same so-called *Curtis-Phan* relations as G^* [45, 111] (see also [77, pp. 218–220]). We describe these, but omit technicalities.

Let Δ be the Dynkin diagram of G^*. By definition, Δ is the set of vertices of a suitable connected graph. Let G^* be defined over $GF(q)$. The Curtis-Phan relations assert that G is generated by a collection of $SL_2(q)$ subgroups L_i, indexed by Δ, with the following properties;

(1) If i, j are not adjacent in Δ, then $[L_i, L_j] = 1$; and

(2) If i, j are adjacent in Δ, then $\langle L_i, L_j \rangle$ is isomorphic to a covering group of $PSL_3(q)$ or $PSU_3(q)$ and $[Z(L_i), Z(L_j)] = 1$.

These conditions imply that G satisfies the conditions of the Steinberg presentation for G^*. [More precisely, it is the "universal covering group" $\tilde{G}^*$ of G^* that possesses the Steinberg presentation, and the Curtis-Phan theorem asserts that G is necessarily isomorphic to a homomorphic image of $\tilde{G}^*$. Since G^* is the unique (nonabelian) simple homomorphic image of G^*, it will then follow that $G \cong G^*$.]

Case A_2. $G^* \cong A_n$ or Σ_n, $n \geq 12$. Correspondingly $K \cong A_{n-4}$ or A_{n-2}, J is G-conjugate to K, and $C_K \cong Z_2 \times Z_2$ or Z_2. In the A_n case, the G-conjugates of C_K are called *root* four-subgroups of G. In this case one constructs a graph Γ whose vertices are the root four-subgroups of G with two vertices connected by an edge if and only if they commute elementwise. Clearly G acts by conjugation as a transitive automorphism group of Γ. In this case, the aim is to show that Γ is *uniquely* determined—equivalently, that Γ is graph-isomorphic to the root four-subgroup graph Γ^* of G^*.

Aschbacher [11] has shown, for all $n \geq 10$, that $\text{Aut}(\Gamma^*) \cong \Sigma_n$. Since G is simple with $G \leq \text{Aut}(\Gamma)$, once the uniqueness of Γ is established, it will follow that $G \cong G^*$.

Now consider the Σ_n case. Here the goal is to force G to satisfy the conditions of the classical presentation for Σ_n: namely, G is generated by involutions x_i, $1 \leq i \leq n - 1$, subject to the relations $|x_i x_j| = 1$, 2, or 3 according as $i = j$, $|j - i| \geq 2$, or $|j - i| = 1$. It is, in fact, shown that the desired involutions x_i can be chosen to lie in $N_K \cup N_J$. The argument thus yields that $G \cong G^*$. Since Σ_n is not simple, this case therefore leads to a contradiction and so is excluded.

Now consider the remaining possibility under A:

Case A_3. G^* of Lie type of characteristic 2 and Lie rank 2 (with $G = O^{2'}(\langle M_1, M_2 \rangle)$). Again the argument is completely local group-theoretic. Here

the goal is to show that G is a (B, N)-pair of the same shape as G^*. Let U be a Sylow 2-subgroup of G, and set $B = N_U$. One must prove the existence of a subgroup N of G with the following properties:

(1) $H = B \cap N$ is normal in N with H of odd order and $B = HU$;

(2) $N/U = W$ is a group generated by two reflections;

(3) $G = BNB$ and the multiplication table of double cosets ByB, $By'B$ for y, $y' \in N$ is completely specified; and

(4) The isomorphism types of B and N are completely determined, as is the action of N on U.

Together, conditions (3) and (4) show that the multiplication table of G is uniquely determined. Thus it will follow that $G \cong G^*$. But now if $G^* \cong {}^2F_4(2)$, then G is not simple, so that possibility will be excluded.

Now we briefly discuss the various possibilities for G^* under B. Here the arguments are considerably more elaborate, some subcases requiring a complete determination of the character table of G, including in a few instances information about r-blocks of characters for suitable primes r, others requiring considerable generator-relation computations. Most subcases also require considerable local analysis.

We separate our comments into seven subcases.

Case B_1. $G^* \cong PSL_2(p^n)$, p an arbitrary prime. Again let $B = N_U$, U a Sylow p-subgroup of G. Here the goal is to prove that G is a (B, N)-pair of rank 1 with a uniquely determined structure for B and N, in which case the multiplication table of G will be uniquely determined.

Case B_2. $G^* \cong PSU_3(p^n)$, $Sz(2^n)$ or ${}^2G_2(3^n)$. Setting $p = 2$ or 3 in the latter two cases, the goal is the same as that of the previous case, but is now vastly more complicated. One must first force an initial set of possibilities for the structure of B by local methods, and must then prove that only one of these can occur within G. The latter argument seems to be inherently extremely difficult, necessitating very delicate generator-relation calculations, rather reminiscent of the final chapter of the odd order theorem.

Case B_3. G^* is one of the groups under B of Lie type and Lie rank $\geqslant 2$. Again let p be the characteristic of G^*, U a Sylow p-subgroup of G, and $B = N_U$. Once again the goal is to prove that G is a (B, N)-pair of Lie rank 2 (except in the cases $PSL_5(2)$, $PSU_6(2)$, $PSO_8^+(2)$, which have respective Lie ranks 3, 3, 4) with B and N having a uniquely determined structure and unique action of N on U, in which case it will follow that $G \cong G^*$.

Case B_4. G^* is the Tits group ${}^2F_4(2)'$. This group is not a (B, N) pair, but as Tits has shown [119], it has a very nice presentation. The goal of the analysis is to force G to satisfy the conditions of his presentation.

Case B_5. $G^* \cong A_n$, $n = 7, 9, 10$ or 11. Here the goal is to construct a subgroup G_0 of index n in G, in which case the permutation representation of G on the cosets of G_0 will yield a homomorphism of G into Σ_n. Since G is simple, this will force $G \cong A_n$.

Case B_6. G^* sporadic with $G^* \not\cong J_1$, J_4, F_5, F_3, or F_1. Apart from the listed groups and the three Conway groups, the remaining eighteen sporadic groups were originally constructed as primitive permutation groups. Moreover, it can be shown that each of them is uniquely determined by their degrees n, the

structure of a one-point stabilizer X, and the various one-point stabilizers of X on the remaining $n - 1$ points.

Although the Conway groups .3, .2, and .1 were constructed from the Leech lattice, they have also been shown to be uniquely determined as suitable primitive permutation groups.

Thus the goal in these twenty-one cases is to force G to possess a primitive permutation representation of the same degree and with the same one-point and two-point stabilizers as G^*, in which case it will follow that G must be isomorphic to G^*.

Case B_7. $G^* \cong J_1$, J_4, F_5, F_3, or F_1. In these cases, the aim is to show that G has an irreducible representation Φ as a group of linear transformations of respective degrees 7, 112, 133, 248, or 196,883 over the respective fields $GF(11)$, $GF(2)$, $\mathbf{C}$, $GF(3)$, or $\mathbf{C}$, and then that $\Phi(G)$ is uniquely determined.

In the case of F_3, Φ maps G into the subgroup $E_8(3)$ of $GL_{248}(3)$, and the uniqueness of $\Phi(G)$ is established as a subgroup of $E_8(3)$. On the other hand, in the case of F_1, if χ denotes the character of Φ, χ is self-dual (i.e., χ is real-valued) and χ occurs with multiplicity 1 in the decomposition of both its symmetric square and symmetric cube. These conditions imply that the vector space V afforded by χ admits a $\Phi(G)$-invariant commutative, nonassociative algebra structure B, equipped with an associative form (B is the Griess-Norton algebra refered to in Chapter I, §1), and $\Phi(G)$ is uniquely determined as a linear group acting on this algebra.

15. Concluding comments. The preceding outline should give a picture of the general shape of the proposed *GLS* unitary classification proof. There is a single generic case, leading quite uniformly to the large Lie type or alternating groups, plus a variety of small and intermediate cases, leading to the sporadic groups and the remaining Lie type and alternating groups, each such case requiring its own special treatment. Moreover, it is the analyses of these nongeneric cases that unfortunately make the total length of the classification proof so great, and explains why we consider a 3,000-page proof to be about the minimum we could hope to achieve with the present methods.

Finally, we ask the question: Is there a good conceptual way of interpreting the proof in the generic case?

This case is defined in terms of the centralizer of an element x of suitable prime order p in quasisimple Brauer form with p-generic Brauer component K. Consider first the case that K (and ultimately G) is of Lie type and let L_1, $L_2, \ldots, L_m$ be the canonical $SL_2(q)$ subgroups of G satisfying the Curtis-Phan relations. Then it is easily checked that for a suitable choice of the L_i, the element x can be taken in L_1 with p dividing $q \pm 1$. Since elements of order prime to q in a group of Lie type over $GF(q)$ are *semisimple* in the Lie sense, it is reasonable, in keeping with our terminology, to refer to such an element x as a *semisimple Brauer* element.

On the other hand, when K (and G) are alternating groups, it is shown in the course of the proof that the element x is a *root involution* of G—i.e., x is contained in a root four-subgroup of G.

Hence, if we add root involutions in alternating groups to the definition of

semisimple Brauer elements, we see that we can view the classification of the generic simple groups in the following way:

The classification proof for the generic simple group G consists of the following two steps:

(1) *An internal search for a semisimple Brauer element x of G.*

(2) *Construction of a presentation for G from the specified structure of the centralizer of x.*

The nongeneric cases arise precisely when either G does not contain such a semisimple Brauer element or when such an element exists, but its centralizer does not contain an appropriate generic Brauer component.

BIBLIOGRAPHY

1. J. L. Alperin, *Sylow intersections and fusion*, J. Algebra **6** (1967), 222–241.

2. J. L. Alperin, R. Brauer, and D. Gorenstein, *Finite groups with quasi-dihedral and wreathed Sylow 2-subgroups*, Trans. Amer. Math. Soc. **151** (1970), 1–261.

3. _____, *Finite simple groups of 2-rank two*, Scripta Math. **29** (1973), 191–214.

4. M. Aschbacher, *Finite groups with a proper 2-generated core*, Trans. Amer. Math. Soc. **197** (1974), 87–223.

5. _____, *Finite groups generated by odd transpositions. I, II, III, IV*, Math Z. **127** (1972), 45–56; J. Algebra **26** (1973), 451–459; **26** (1973), 460–478; **26** (1973), 479–491.

6. _____, *On finite groups of component type*, Illinois J. Math. **19** (1975), 78–115.

7. _____, *A characterization of Chevalley groups over fields of odd characteristic*, Ann. of Math. (2)**106** (1977), 353–468.

8. _____, *On finite groups in which the generalized Fitting group of the centralizer of some involution is symplectic but not extra-special*, Comm. Algebra **4** (1976), 595–616.

9. _____, *On finite groups in which the generalized Fitting group of the centralizer of some involution is extra-special*, Illinois J. Math. **21** (1977), 347–364.

10. _____, *A factorization theorem for 2-constrained groups*, Proc. London Math. Soc. **43** (1981), 450–477.

11. _____, *Standard components of alternating type centralized by a 4-group* (preprint).

12. _____, *On the maximal subgroups of the finite classical groups*, Invent. Math. **76** (1984), 469–514.

13. _____, *Thin finite simple groups*, J. Algebra **54** (1978), 50–152.

14. _____, *Finite groups of rank 3. I, II*, Invent. Math. **63** (1981), 357–402; **71** (1983), 51–162.

15. _____, *Weak closure in finite groups of even characteristic*, J. Algebra **70** (1981), 561–627.

16. _____, *The uniqueness case for finite groups*, Ann. of Math. (2)**117** (1983), 383–551.

17. _____, *Subgroup structure of finite groups*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 35–45.

18. _____, *Finite groups*, Cambridge Univ. Press, New York-Cambridge (to appear).

19. _____, *Overgroups of Sylow subgroups in sporadic groups*, Mem. Amer. Math. Soc. (to appear).

20. M. Aschbacher, D. Gorenstein, and R. Lyons, *The embedding of 2-locals in finite groups of characteristic 2 type*, Ann. of Math. (2)114 (1981), 335–455.

21. M. Aschbacher and L. Scott, *Maximal subgroups of finite groups*, J. Algebra 92 (1985).

22. M. Atiyah, *An interview with Michael Atiyah*, Math. Intelligencer 6 (1984), 9–19.

23. B. Baumann, *Endliche Gruppen mit einer 2-zentralen Involution, deren Zentralisator 2-abgeschlossen ist*, Illinois J. Math. 22 (1978), 240–261.

24. G. Belyĭ, *On Galois extensions of a maximal cyclotomic field*, Izv. Akad. Nauk SSSR Ser. Mat. 43 (1979), 267–276 = Math. USSR-Izv. 14 (1980), 247–256.

25. H. Bender, *Transitive Gruppen gerader Ordnung, in denen jede Involution genau einen Punkt festlasst*, J. Algebra 17 (1971), 527–554.

26. _____, *On groups with abelian Sylow 2-subgroups*, Math. Z. 117 (1970), 164–176.

27. _____, *On the uniqueness theorem*, Illinois J. Math. 14 (1970), 376–384.

28. _____, *Goldschmidt's 2-signalizer functor theorem*, Israel J. Math. 22 (1975), 208–213.

29. _____, *Finite groups with dihedral Sylow 2-subgroups*, J. Algebra 70 (1981), 216–228.

30. R. Blattner, *Induced and produced representations of Lie algebras*, Trans. Amer. Math. Soc. 144 (1969), 457–474.

31. R. Block and R. Wilson, *The restricted simple Lie algebras are of classical or Cartan type*, Proc. Nat. Acad. Sci. U.S.A. 81 (1984), 5271–5274.

32. E. Bombieri, *Thompson's problem ($\sigma^2 = 3$)*, Invent. Math. 58 (1980), 77–100.

33. A. Borel and J. Tits, *Elements unipotents et sousgroupes paraboliques de groupes reductifs. I*, Invent. Math. 12 (1971), 97–104.

34. R. Brauer, *On the structure of groups of finite order*, Proc. Internat. Congr. Math., vol. 1, Noordhoff, Groningen, North-Holland, Amsterdam, 1954, pp. 209–217.

35. _____, *Some applications of the theory of characters of finite groups. II*. J. Algebra 1 (1964), 307–334.

36. R. Brauer and K. Fowler, *On groups of even order*, Ann. of Math. (2) 62 (1955), 565–583.

37. R. Brauer and M. Suzuki, *On finite groups of even order whose 2-Sylow group is a quaternion group*, Proc. Nat. Acad. Sci. U.S.A. 45 (1959), 1757–1759.

38. F. Buekenhout, *Diagrams for geometries and groups*, J. Combinatorial Theory Ser. A 27 (1979), 121–151.

39. P. Cameron, S. Praeger, J. Saxl, and G. Seitz, *The Sims conjecture* (to appear).

40. G. Cherlin, L. Harrington, and A. Lachlan, $\aleph_0$-*categorical*, $\aleph_0$-*stable structures* (to appear).

41. G. Cherlin and A. Lachlan, *Stable finitely homogeneous structures* (to appear).

42. C. Chevalley, *Sur certains groupes simples*, Tôhoku Math. J. 7 (1955), 14–66.

43. J. Conway, *A characterization of the Leech lattice*, Invent. Math. 7 (1969), 137–142.

44. J. Conway and S. Norton, *Monstrous moonshine*, Bull. London Math. Soc. 11 (1979), 308–339.

45. C. Curtis, *Central extensions of groups of Lie type*, J. Reine Angew. Math. 220 (1965), 174–185.

46. A. Delgado and B. Stellmacher, *Weak BN-pairs of rank 2* (preprint).

47. E. Dynkin, *Maximal subgroups of classical groups*, Trudy Moskov. Mat. Obshch. 1 (1952), 39–166 = Amer. Math. Soc. Transl. (2)6 (1957), 245–378.

48. _____, *Semisimple subalgebras of semisimple Lie algebras*, Mat. Sb. (N.S.) 30(72) (1952), 349–462 = Amer. Math. Soc. Transl. (2)6 (1957), 111–244.

49. F. Dyson, *Unfashionable pursuits*, Math. Intelligencer 5 (1983), 47–54.

50. M. Enguehard, *Caractérisation des groupes de Ree* (preprint).

51. B. Fein, W. Kantor, and M. Schacher, *Relative Brauer groups. II*, J. Reine Angew. Math. 328 (1981), 39–57.

52. W. Feit, *Rigidity of $PSL_2(p^2)$, $p \equiv \pm 2 \pmod 5$, $p \neq 2$*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 351–356.

53. W. Feit and P. Fong, *Rational rigidity of $G_2(p)$ for any $p > 5$*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 323–326.

54. W. Feit, P. Fong, and B. Srinivasan, *Rigidity of $SL_n(q)$ and certain subgroups for $(n, q - 1) = 1$ and $n > 2$*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 303–308.

55. W. Feit and J. Thompson, *Solvability of groups of odd order*, Pacific J. Math. 13 (1963), 775–1029.

56. B. Fischer, *Finite groups generated by 3-transpositions*, Invent. Math. **13** (1971), 232–246.

57. P. Fong, *Characters arising in the monster-modular connection*, Proc. Sympos. Pure Math., vol. 37, Amer. Math. Soc., Providence, R. I., 1980, pp. 557–560.

58. R. Foote, *Aschbacher blocks*, Proc. Sympos. Pure Math., vol. 37, Amer. Math. Soc., Providence, R.I., 1980, pp. 37–42.

59. I. Frenkel, J. Lepowsky and A. Meurman, *A natural representation of the Fischer-Griess monster with the modular function J as character*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), 3256–3260.

60. R. Gilman, *Components of finite groups*, Comm. Algebra **4** (1976), 1133–1198.

61. R. Gilman and R. Griess, *Finite groups with standard components of Lie type over fields of characteristic 2*, J. Algebra **80** (1983), 383–516.

62. G. Glauberman, *Central elements in core-free groups*, J. Algebra **4** (1966), 403–420.

63. _____, *On solvable signalizer functors in finite groups*, Proc. London Math. Soc. **33** (1976), 1–27.

64. _____, *Factorizations in local subgroups of finite groups*, CBMS Regional Conf. Ser. in Math., no. 33, Amer. Math. Soc., Providence, R.I., 1977.

65. G. Glauberman, *Local analysis in the odd order paper*, Proc. Sympos. Pure Math., vol. 37, Amer. Math. Soc., Providence, R.I., 1980, p. 137.

66. G. Glauberman and R. Niles, *A pair of characteristic subgroups for pushing up*, Proc. London Math. Soc. (3) **46** (1983), 411–453.

67. D. Goldschmidt, *Solvable signalizer functors on finite groups*, J. Algebra **21** (1972), 137–148.

68. _____, *2-signalizer functors on finite groups*, J. Algebra **21** (1972), 321–340.

69. _____, *2-fusion in finite groups*, Ann. of Math. (2) **99** (1974), 70–117.

70. _____, *Strongly closed 2-subgroups of finite groups*, Ann. of Math. (2) **102** (1975), 475–489.

71. _____, *Automorphisms of trivalent graphs*, Ann. of Math. (2) **111** (1980), 377–406.

72. _____, *Lectures on character theory*, Publish or Perish Press, Berkeley, 1980.

73. D. Gorenstein, *On the centralizers of involutions in finite groups*, J. Algebra **11** (1969), 243–277.

74. _____, *Finite groups*, Harper and Row, New York, 1968; 2nd ed., Chelsea, New York, 1980.

75. _____, *The classification of the finite simple groups*, Bull. Amer. Math. Soc. **1** (1979), 43–199.

76. _____, *Finite simple groups*, Plenum, New York, 1982.

77. _____, *The classification of finite simple groups*, Plenum, New York, 1983.

78. D. Gorenstein and K. Harada, *Finite groups whose 2-subgroups are generated by at most 4 elements*, Mem. Amer. Math. Soc. No. 147 (1974), 1–464.

79. D. Gorenstein and R. Lyons, *Nonsolvable finite groups with solvable 2-local subgroups*, J. Algebra **38** (1976), 453–522.

80. _____, *The local structure of finite groups of characteristic 2 type*, Mem. Amer. Math. Soc. No. **42** (1983), 1–731.

81. D. Gorenstein and J. Walter, *On the maximal subgroups of finite simple groups*, J. Algebra **1** (1964), 168–231.

82. _____, *The characterization of finite groups with dihedral Sylow 2-subgroups*, J. Algebra **2** (1964), 85–151, 218–270, 354–393.

83. _____, *Centralizers of involutions in balanced groups*, J. Algebra **20** (1972), 284–319.

84. _____, *Balance and generation in finite groups*, J. Algebra **33** (1975), 224–287.

85. R. Griess, *The friendly giant*, Invent. Math. **69** (1982), 1–102.

86. G. Higman, *Suzuki 2-groups*, Illinois J. Math. **7** (1963), 79–96.

87. G. Hoyden, *Realisierung der Jankogrouppen J_1 and J_2 als Galoisgruppen uber*, Algebra J. (to appear).

88. G. Hoyden and B. Matzat, *Realisierung sporadischer einfacher gruppen als Galoisgruppen ober Kreisteilungskorpern* (to appear).

89. D. Hunt, *Rational rigidity and the sporadic groups* J. Algebra (to appear).

90. M. Isaacs, *Character theory of finite groups*, Academic Press, New York, 1976.

91. Z. Janko, *A new finite simple group of order 86, 775, 571, 046, 077, 562, 880 which possesses M_{24} and the full cover of M_{22} as subgroups*, J. Algebra **42** (1976), 564–496.

92. _____, *Nonsolvable finite groups all of whose 2-local subgroups are solvable*. I, J. Algebra **21** (1972), 458–517.

93. V. Kac., *On the classification of simple Lie algebras over a field of nonzero characteristic*, Izv. Akad. Nauk SSSR Ser. Mat. **34** (1970), 385–408; English transl. Math. USSR-Izv. **4** (1970), 391–413.

94. _____, *Description of filtered Lie algebras with which graded Lie algebras of Cartan type are associated*, Izv. Akad. Nauk SSSR Ser. Mat. **38** (1974), 800–834; English transl. Math. USSR-Izv. **8** (1974), 801–835.

95. W. Kantor, *Polynomial-time algorithms for finding elements of prime order and Sylow subgroups* (submitted).

96. K. Klinger and G. Mason, *Centralizers of p-groups in groups of characteristic 2, p-type*, J. Algebra **37** (1975), 362–375.

97. A. Kostrikin and I. Shafarevich, *Graded Lie algebras of finite characteristic*, Izv. Akad. Nauk SSSR Ser. Mat. **33** (1969), 748–764; English transl. Math. USSR-Izv. 3 (1969), 237–304.

98. J. Leech, *Some sphere packings in higher space*, Canad. J. Math. **16** (1964), 657–682.

99. _____, *Notes on sphere packings*, Canad. J. Math. **19** (1967), 251–267.

100. E. Luks, *Isomorphism of graphs of bounded valence can be tested in polynomial time*, J. Comp. System. Sci. **25** (1982), 42–65.

101. J. Lundgren, *On finite simple groups all of whose 2-local subgroups are solvable*, J. Algebra **27** (1973), 491–515.

102. R. Lyons, *A characterization of PSU(3, 4)*, Trans. Amer. Math. Soc. **164** (1972), 371–387.

103. A. MacWilliams, *On 2-groups with no normal abelian subgroups of rank 3 and their occurrence as Sylow 2-subgroups of finite simple groups*, Trans. Amer. Math. Soc. **150** (1970), 345–408.

104. G. Mason, *The classification of quasithin simple groups* (preprint).

105. _____, *Modular forms and the theory of Thompson series*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, 391–408.

106. B. Matzat, *Zum einbettungsproblem der algebraischen zahlentheorie mit nicht abelschem kern* (to appear)

107. B. Matzat and A. Zeh, *Realisierung der mathieugruppen M_{11} and M_{12} als Galoisgruppen uber*, J. Number Theory (to appear).

108. P. McBride, *Nonsolvable signalizer functors on finite groups* (preprint).

109. E. Noether, *Gleichungen mit Vorgeschriebener Gruppe*, Math. Ann. **78** (1918), 221–229.

110. T. Peterfalvi, *Le Théorème de Bender-Suzuki*. I, II (preprints).

111. K. Phan, *On groups generated by SU(3, q)'s*. I, II, J. Austral. Math. Soc. **23** (1977), 66–76; 129–146.

112. R. Ree, *A family of simple groups associated with the simple Lie algebra F_4*, Amer. J. Math. **83** (1961), 401–420.

113. _____, *A family of simple groups associated with the simple Lie algebra G_2*, Amer. J. Math. **83** (1961), 432–462.

114. M. Ronan and S. Smith, *2-local geometries for some sporadic groups*, Proc. Sympos. Pure Math., vol. 37, Amer. Math. Soc., Providence, R.I., 1980, pp. 283–289.

115. G. Seitz, *The root subgroups for maximal tori in finite groups of Lie type*, Pacific J. Math. **106** (1983), 153–244.

116. _____, *Overgroups of irreducible linear groups*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 95–106.

117. I. Shafarevich, *Construction of fields of algebraic numbers with given solvable Galois group*, Izv. Akad. Nauk SSSR Ser. Mat. **18** (1954), 525–578.

118. D. Sibley, *Groups of odd order revisited* (in preparation).

119. C. Sims, *Graphs and finite permutation groups*. I, II, Math. Z. **95** (1967), 76–86; **103** (1968), 276–281.

120. _____, *Some group theoretic algorithms*, Lecture Notes in Math., vol. 697, Springer-Verlag, Berlin and New York, 1978, pp. 108–124.

121. F. Smith, *Finite simple groups all of whose 2-local subgroups are solvable*, J. Algebra **34** (1975), 481–520.

122. S. Smith, *Large extraspecial subgroups of widths 4 and 6*, J. Algebra **58** (1979), 251–280.

123. _____, *A characterization of orthogonal groups over GF(2)*, J. Algebra **62** (1980), 39–60.

124. _____, *A characterization of finite Chevalley and twisted groups of type E over GF(2)*, J. Algebra **62** (1980), 101–117.

125. _____, *On the head character of the monster simple group*, Proc. Montreal Finite Group Theory Conference, 1982 (John McKay, ed.) (to appear).

126. R. Steinberg, *Variations on a theme of Chevalley*, Pacific J. Math. **9** (1959), 875–891.

127. _____, *Generateurs, relations, et revêtements de groupes algébriques*, Colloque sur la theorie des groupes algébriques, C.B.R.M., Brussels, 1962, pp. 113–127.

128. _____, *Lectures on Chevalley groups*, Lecture Notes, Yale Univ. 1967-1968.

129. _____, *Endomorphisms of linear algebraic groups*, Mem. Amer. Math. Soc. No. 80 (1968).

130. G. Stroth, *Solvable uniqueness subgroups* (to appear).

131. M. Suzuki, *Finite groups with nilpotent centralizers*, Trans. Amer. Math. Soc. **99** (1961), 425–470.

132. _____, *On a class of double transitive groups*. I, II, Ann. of Math. (2)**75** (1962),

133. _____, *On characterizations of linear groups*. III, Nagoya J. Math. **21** (1962), 159–183.

134. J. Thompson, *Nonsolvable finite groups all of whose local subgroups are solvable*. I-VI, Bull. Amer. Math. Soc. **74** (1968), 383–437; Pacific J. Math. **33** (1970), 451–536; **39** (1971), 483–534; **48** (1973), 511–592; **50** (1974), 215–297; **51** (1974), 573–630.

135. _____, *Toward a characterization of $E_2^*(q)$*. I, II, III, J. Algebra **7** (1967), 406–414; **20** (1972), 610–621; **49** (1977), 163–166.

136. _____, *Some numerology between the Fischer-Griess monster and the elliptic modular function*, Bull. London Math. Soc. **11** (1979), 352–353.

137. _____, *Some finite groups which appear as $G(L/K)$ where $K \supseteq Q(\mu_m)$*, J. Algebra **89** (1984), 437–499.

138. _____, *PSL_3 and Galois groups over Q*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 309–320.

139. _____, *Rational rigidity of $G_2(5)$*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 321–322.

140. _____, *Primitive roots and rigidity*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 327–350.

141. F. Timmesfeld, *Groups generated by root involutions*. I, J. Algebra **33** (1975), 75–135.

142. _____, *Finite simple groups in which the generalized Fitting group of the centralizer of some involution is extra-special*, Ann. of Math. (2)**107** (1978), 297–369.

143. F. Timmesfeld, *Tits geometries and parabolic systems in finite groups*. I, II, Math. Z. **184** (1983), 377–396; 449–487.

144. J. Tits, *Algebraic and abstract simple groups*, Ann. of Math. (2)**80** (1964), 313–329.

145. _____, *Buildings of spherical type and finite (B, N)-pairs*, Lecture Notes in Math., vol. 386, Springer-Verlag, Berlin and New York, 1974.

146. _____, *A local approach to buildings*, The Geometric Vein, (Coxeter-Festschrift), Springer-Verlag, Berlin and New York, 1981, pp. 519–547.

147. _____, *Affine buildings of rank $\geqslant 4$* (unpublished).

148. W. Tutte, *A family of cubical graphs*, Proc. Cambridge Philos. Soc. **43** (1947), 459–474.

149. _____, *On the symmetry of cubic graphs*, Canad. J. Math. **11** (1959), 621–624.

150. J. Walter, *The characterization of finite groups with abelian Sylow 2-subgroups*, Ann. of Math. (2)**89** (1969), 405–514.

151. _____, *Classical groups as Galois groups*, Proc. Rutgers Group Theory Year 1983–1984, Cambridge Univ. Press, New York-Cambridge, 1984, pp. 385–390.

152. _____, *The B-conjecture; 2-components in finite simple groups* (to appear).

153. W. Wong, *Determination of a class of primitive permutation groups*, Math. Z. **99** (1967), 235–246.

154. B. Zilber, *Strongly minimal countably categorical theories*, Siberian Math. J. **21** (1980), 98–112. (Russian)

155. T. Peterfalvi, *Simplification du chapitre VI de l'article de Feit et Thompson sur les groupes d'ordre impair*, C. R. Acad. Sci. Paris Ser. I **229** (1984), 531–534.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, NEW BRUNSWICK, NEW JERSEY 08903