CHAPTER 0

Introduction

In this introductory chapter we give an overview of the strategy behind the proof of the theorem classifying the finite simple groups, while for the most part attempting to avoid the fine structure of the proof. At various places we make oversimplifications to avoid technicalities; this means that some statements are not quite true in special cases, but hopefully this approach conveys the flavor of the proof more clearly than a more involved, technically correct discussion.

To be a bit more specific, we describe how, for a suitable choice of prime p, each finite simple group G is determined up to isomorphism by its p-local subgroups. This makes it possible to classify the simple groups in terms of their local structure.

We explain why the prime 2 plays a special role and is usually the optimal choice for p. We sketch a proof of the Dichotomy Theorem 0.3.10, which partitions the simple groups into groups of Gorenstein-Walter type and characteristic 2 type, according to two possible 2-local structures. We also describe the partition of each of these two types of simple groups into large and small groups of the given type, leading to a four-part subdivision of the proof of the Classification.

0.1. The Classification Theorem

We begin with a statement of the Classification Theorem:

THEOREM 0.1.1 (Classification Theorem). *Each finite simple group is isomorphic to one of:*

(1) A group of prime order.

(2) An alternating group on a set of order at least 5*.*

(3) A finite simple group of Lie type.

(4) One of 26 *sporadic simple groups.* □

In Section 0.4 we discuss each of the four classes of groups in more detail. But first we begin with a brief discussion of local group theory, and the role it plays in the proof of the Classification. We assume the reader is familiar with the most basic notation and terminology from finite group theory. For those who are not, some notation and terminology is defined at the start of Section A.1 in Appendix A.

Throughout this chapter, G is a finite group and p is a prime. Further $Syl_p(G)$ denotes the set of all Sylow p-subgroups of G. Pick $S \in Syl_p(G)$, and write $\mathcal{P}$ for the set of nontrivial subgroups of S.

A subgroup of G is called *p-local* if it is the normalizer $N_G(P)$ of some nontrivial p-subgroup P of G. We speak simply of a *local subgroup* if we do not wish to specify the prime p. The study of finite groups from the point of view of their local subgroups is called local group theory.

The local theory is the most important tool in the proof of the Classification of the finite simple groups. In particular the proof of the Classification is based on the following two principles:

Principle I: Recognition via local subgroups. If G is simple and the p-local structure of G is sufficiently rich, then G is determined up to isomorphism by the local subgroups $(N_G(P) : P \in \mathcal{P})$; and even by a subfamily $(N_G(Q) : Q \in \mathcal{Q})$ for some suitable small subset $\mathcal{Q}$ of $\mathcal{P}$.

Principle II: Restricted structure of local subgroups. If G is simple then the structure of the locals $(N_G(P) : P \in \mathcal{P})$ is highly restricted. Indeed up to a suitable notion of isomorphism, there are only a small number of choices for the family $(N_G(\mathcal{P}) : P \in \mathcal{P})$.

To oversimplify a bit, the Classification amounts to translating the two Principles into precise statements, and writing down proofs of those statements, at least in enough special cases to include all simple groups. Once this is accomplished, given a finite simple group G, Principle II tells us that for some prime p, G has the same p-local structure as one of the groups $\overline{G}$ listed in the Classification Theorem. Then Principle I says that G is determined up to isomorphism by its p-locals, so G is isomorphic to $\overline{G}$.

0.2. Principle I: Recognition via local subgroups

To make Principle I more precise, we will use the language of topological combinatorics:

NOTATION 0.2.1 (Graphs and complexes associated to p-subgroups). Given a positive integer k, define $\Lambda_k = \Lambda_k^p(G)$ to be the commuting graph[1] on the set of elementary abelian p-subgroups E of G with p-rank $m_p(E) \geq k$. Define $\Lambda_k^p(G)^\circ$ to be the subgraph of Λ_k consisting of those E with $m_p\big(C_G(E)\big) > k$. Notice that if $m_p(G) \leq 1$, then $\Lambda_1^p(G)$ is totally disconnected, and $\Lambda_1^p(G)^\circ$ is the empty graph. Thus these graphs are of no interest unless $m_p(G) \geq 2$.

For a subset $\mathcal{Q}$ of $\mathcal{P}$, form the corresponding set $\mathcal{N} = \mathcal{N}_{\mathcal{Q}} := \{N_G(Q) : Q \in \mathcal{Q}\}$ of local subgroups, and regard $\mathcal{N}$ as a poset (partially ordered set), ordered by inclusion.

The corresponding *amalgam* $\mathcal{A}(\mathcal{N})$ (cf. the discussion in the relevant subsection of Section 3.3, e.g. 3.3.14) roughly encodes the elements of members of $\mathcal{N}$ as generators for any abstract group $\overline{G}$ satisfying the relations defining the members of $\mathcal{N}$, plus the relations induced by the inclusions among the members of $\mathcal{N}$.

Also let $\Sigma(\mathcal{N})$ denote the "space of cosets" $\coprod_{N \in \mathcal{N}} G/N$, regarded as a poset via $Nx \leq My$ iff $N \leq M$ with $Nx \subseteq My$. Write $\mathcal{O}(\mathcal{N})$ for the order complex of the poset $\Sigma(\mathcal{N})$. That is, $\mathcal{O}(\mathcal{N})$ is the abstract simplicial complex whose simplices are the chains in $\Sigma(\mathcal{N})$. Observe that G acts as a group of automorphisms of $\Sigma(\mathcal{N})$ and $\mathcal{O}(\mathcal{N})$ via right multiplication. ♢

We remark (cf. the discussion after 0.4.1) that if G is a finite simple group of Lie type defined over a field of characteristic p, and $\mathcal{N}$ is the set of all p-local subgroups of G, then $\mathcal{O}(\mathcal{N})$ is the spherical Tits building associated with G. Further

[1] i.e. where edges are defined by the corresponding subgroups centralizing each other; cf. Definition 1.3.8 for further discussion of commuting graphs. Also the p-rank $m_p(X)$ is the maximal rank of an elementary abelian p-subgroup of X; cf. Definition A.1.6.

if G has BN-rank at least 3, then G is determined as a distinguished subgroup of the automorphism group of $\mathcal{O}(\mathcal{N})$. Thus Principle I holds in this important case.

Recognition via simple connectivity. For the purposes of Principle I, the p-local structure of G is "sufficiently rich" if (again oversimplifying for expository purposes) Λ_1 is connected, and there exists $\mathcal{Q} \subseteq \mathcal{P}$ such that the simplicial complex $\mathcal{O}(\mathcal{N})$ is simply connected. The condition that Λ_1 is connected is equivalent to $G = \langle N_G(P) : P \in \mathcal{P} \rangle$, while $\mathcal{O}(\mathcal{N})$ is connected if and only if $G = \langle \mathcal{N} \rangle$. Thus the condition that Λ_1 is connected is implied by the condition that $\mathcal{O}(\mathcal{N})$ is simply connected, but we include that condition anyway for emphasis. The condition that $\mathcal{O}(\mathcal{N})$ is simply connected implies that G is the free amalgamated product of the amalgam $\mathcal{A}(\mathcal{N})$; that is, that G is the largest group $\overline{G}$ generated by the elements of members of $\mathcal{N}$, subject to defining relations for those members, and relations from the inclusions among members of $\mathcal{N}$. Hence G is indeed determined up to isomorphism by the amalgam $\mathcal{A}(\mathcal{N})$ of local subgroups, so that Principle I holds.

Recognition via a single subgroup (such as an involution centralizer). Now in practice the amalgam $\mathcal{A}(\mathcal{N})$ is a fairly complicated and unwieldy object. Thus we often focus on one carefully chosen local $H := N_G(Q)$, where we have $N_S(Q) \in Syl_p(H)$; and show that $\mathcal{A}(\mathcal{N})$ can be retrieved from H—together with the embedding of H in G imposed by the constraint that H is a local subgroup, and the G-conjugacy pattern of suitable elements of $N_S(Q)$. Hence G also is often determined up to isomorphism by one local H, its embedding in G, and *fusion* (i.e. G-conjugacy) in the Sylow p-subgroup $N_S(Q)$ of H.

Let us look at an example from the early days of the modern effort to classify the simple groups. We can hope to have some freedom in choosing a prime p, subject to the constraint that the p-local structure is rich. In practice the 2-local structure is usually rich, so we can take $p = 2$. Moreover most often it is best to choose $Q = \langle t \rangle$ with t of order 2 (an *involution*), so that $H = N_G(Q)$ is the involution centralizer $C_G(t)$.

This is precisely the approach for studying finite simple groups proposed by Richard Brauer in his 1954 address [**Bra57**] to the International Congress of Mathematicians in Amsterdam, where he suggested that simple groups should be characterized in terms of the centralizers of involutions. Moreover his program was anchored and motivated by his seminal result with Fowler:

THEOREM 0.2.2 (Brauer-Fowler [**BF55**]). *Let H be a finite group. Then there are at most a finite number of finite simple groups G possessing an involution t such that $C_G(t) \cong H$.*

Now an obvious necessary condition for the 2-local structure of a nonabelian simple group G to be rich is that 2 be a prime divisor of the order of G. An old conjecture of Burnside from the beginning of the twentieth century postulates that each nonabelian finite simple group *is* of even order. Some years after Brauer proposed his program, Feit and Thompson proved Burnside's conjecture in their monumental work [**FT63**]. The Feit-Thompson Odd Order Theorem (which we discuss later as 1.2.1), and Thompson's thesis (establishing an old conjecture of Frobenius), were two of the early major papers in local group theory.

The case where $\Lambda_1(G)$ is disconnected. In order to classify finite simple groups in terms of their p-local structure, we also need to prove some results in the case where that structure is not sufficiently rich in the above sense; in particular, we need to first determine those groups G such that p divides the order of G, but $\Lambda_1^p(G)$ is disconnected.

As noted earlier, this is always the case when G has p-rank 1. However, an early theorem of Brauer-Suzuki (see 1.4.2), in conjunction with the Odd Order Theorem, guarantees that every non-abelian simple group G has 2-rank at least 2.

But there are also disconnected examples with 2-rank at least 2: for example if $G = A_5$ (the alternating group of degree 5), then $\Lambda_1(G)$ has five connected components, each of which is a complete graph on three vertices. Furthermore given the knowledge of finite groups supplied by the Classification, it can be checked that, with certain known exceptions such as the example just mentioned, if the p-rank $m_p(G)$ of a simple group G is at least 2, then $\Lambda_1(G)$ is connected. (Cf. [**GLS98**, Sec 7.6].)

But of course we need to *prove* a result of this type—*without assuming the Classification*—as part of a proof of the Classification Theorem via p-local structure. Fortunately such a result was proved for the prime 2 by Suzuki and Bender;

THEOREM 0.2.3 (Bender-Suzuki [**Suz62, Suz64, Ben71**]). *Let G be a finite simple group of even order such that $\Lambda_1^2(G)$ is disconnected. Then $G \cong SL_2(2^n)$, $PSU_3(2^n)$, or one of the simple Suzuki groups $Sz(2^{2n+1})$; i.e., G is a simple group of Lie type, of characteristic 2 and BN-rank 1.*[2]

We will discuss this result in more detail later at 1.3.5. And we will discuss the groups of Lie type further in Section 0.4 in this chapter.

However it should be emphasized that when p is odd, the only known determination of groups G such that $\Lambda_1^p(G)$ is disconnected, depends upon the Classification. In effect this means that in classifying simple groups, one must always keep the prime 2 in the picture. In our discussion of groups of characteristic 2 type, after 2.2.1 and at various later points, we will see that the absence of an analogue of the Bender-Suzuki Theorem 0.2.3 for odd primes leads to great difficulties in the treatment of those groups.

Existence and uniqueness problems. Finally observe that if we are to characterize simple groups in terms of some p-local condition, denoted say by $\mathcal{C}$, then we must address the corresponding problems of existence and uniqueness:

Existence Problem. There exists *at least* one simple group satisfying $\mathcal{C}$.

Uniqueness Problem. There exists *at most* one simple group satisfying $\mathcal{C}$.

Existence. For most of the finite simple groups, the Existence Problem is almost trivial: Namely, with the exception of some of the sporadic groups and groups of BN-rank 1, each finite simple group G is essentially the group of automorphisms of some highly symmetric and natural object X (cf. our discussion in later Section 0.4). Indeed for the alternating group A_n, X is just the set $\{1, 2, \ldots, n\}$; and for the classical linear groups, X is a vector space or a space with a form. Thus the existence of X gives an easy proof of the existence of G. The representation of G on X makes it relatively easy to check that G satisfies a suitable p-local defining condition $\mathcal{C}$.

[2]The groups of rank 1 and even characteristic are often called *Bender groups*.

The representation of G on X is also the primary tool for proving the many properties of G which are used in an inductive context in the proof of the Classification. These properties and others in addition provide the basis for applying the Classification—to prove numerous results both in finite group theory and other areas of mathematics.

Uniqueness. Similarly the representation of G on X is usually the basis for a fairly simple treatment of the Uniqueness Problem. For example, we can choose our set $\mathcal{N}$ of local subgroups so that the complex $\mathcal{O}(\mathcal{N})$ is isomorphic to a geometry visible in X, with the members of $\mathcal{N}$ the stabilizers of simplices in that geometry. The natural description of X and its symmetry makes it relatively easy to prove that $\mathcal{O}(\mathcal{N})$ is simply connected, hence establishing the uniqueness of G.

Unfortunately some of the sporadic groups have no known representation as a group of automorphisms of some nice highly symmetric object. For these groups, either or both of the Existence and Uniqueness Problems can present difficulties.

0.3. Principle II: Restricted structure of local subgroups

As discussed in Section 0.1, the proof of the Classification amounts to translating the two Principles listed there into precise statements, and providing proofs of those statements.

In Section 0.2 we gave a fairly precise translation of Principle I, and some indication of how to implement that Principle. In this section we discuss Principle II—although we won't get very far in restricting locals in simple groups. More significant restrictions must wait until Section 0.5, following our discussion of the finite simple groups in Section 0.4.

The generalized Fitting subgroup. Our eventual objective is to see what distinguishes local subgroups of simple groups, from local subgroups of the general finite group. We first need some tools for discussing the structure of general finite groups. (Cf. Appendix A starting at Definition A.1.17.)

The first and most important tool is the generalized Fitting subgroup $F^*(H)$ of a finite group H. (We write H in place of G, since often we will be considering the situation where G is a finite simple group, and H is some local subgroup of G.) We define $F^*(H)$ precisely in Definition A.1.19; in this paragraph, we give a more informal introduction. First a *component* of H is a subnormal quasisimple subgroup of H. For the purposes of this section, we can think of these components roughly as certain nonabelian simple normal subgroups of H; their product is denoted by $E(H)$, and we can think of this product as roughly being a direct product. Then $F^*(H)$ is the product of $E(H)$ with the classical Fitting subgroup $F(H)$ of H (i.e. the largest normal nilpotent subgroup of H). We write $O_p(H)$ for the largest normal p-subgroup of H, and it is easy to see that $F(H)$ is the direct product of the subgroups $O_p(H)$, as p varies over the prime divisors of the order of H.

The definition of $F^*(H) = E(H)F(H)$ is due to Bender, building on earlier work of Wielandt, and of Gorenstein and Walter. The fundamental property of $F^*(H)$ is:

THEOREM 0.3.1 (Bender). $C_H\big(F^*(H)\big) = Z\big(F^*(H)\big)$.

PROOF. See e.g. 31.13 in [**Asc00**] □

Notice for example that $F^*(H) = 1$ implies $H = 1$; more generally, $F^*(H)$ roughly "organizes" the structure of H: For as $F^*(H)$ is normal in H, the conjugation map c takes H into $Aut\big(F^*(H)\big)$; and we see by the self-centralizing property 0.3.1 above that $\ker(c) = Z\big(F^*(H)\big)$. The latter group is abelian, and typically we can expect it to be only a small part of H; thus $c(H)$ should be a "nearly" faithful representation of H as a group of automorphisms of $F^*(H)$. Furthermore $F^*(H)$ has a natural product structure—and its automorphisms correspond essentially to those of the factors, along with permutations of isomorphic components. Thus $F^*(H)$ is a relatively uncomplicated group, which controls the structure of the possibly highly complex group H.

The generalized Fitting group in an involution centralizer. Moreover as we continue through this section, we will see that for P a p-subgroup of G, the embedding of $F^*\big(N_G(P)\big)$ in G is fairly restricted.

Indeed first we will examine two examples of very natural groups G, which produce contrasting structures for $F^*\big(C_G(t)\big)$ in the centralizer of an involution t. Then later we will see that those are essentially the *only* two possible structures.

EXAMPLE 0.3.2 (Involution centralizers in matrix groups of odd characteristic). First let G be the "nearly-simple" group $GL_n(q)$, i.e. the group of all invertible $n \times n$ matrices with entries from a finite field of order q. In this example we take q to be a power of an *odd* prime.

Then every involution t of G is conjugate to a diagonal matrix with -1 and 1 as the eigenvalues. (The choice of q odd is implicitly used, since we distinguish -1 from 1.) That is, for some $1 \leq k < n$, we have the block-diagonal form:

$$t = \left(\begin{array}{c|c} -I_k & 0 \\ \hline 0 & +I_{n-k} \end{array}\right).$$

By standard linear algebra, the centralizer $H := C_G(t)$ of t preserves the distinct eigenspaces of t. Hence the centralizer also has block-diagonal form, namely a product

$$H = \left(\begin{array}{c|c} GL_k(q) & 0 \\ \hline 0 & GL_{n-k}(q) \end{array}\right) = \{\left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right) : A \in GL_k(q),\ B \in GL_{n-k}(q)\}.$$

Thus $H = L_1 \times L_2$ is the direct product of two nearly-simple factors isomorphic to $GL_k(q)$ and $GL_{n-k}(q)$. More accurately, $E(H)$ is the product of two (quasisimple) components $E_1 \cong SL_k(q)$ and $E_2 \cong SL_{n-k}(q)$.[3] So $F^*(H)$ has some components and $|H : E(H)|$ is small. $\diamondsuit$

Next we consider a related example yielding a strikingly different structure:

EXAMPLE 0.3.3 (Involution centralizers in matrix groups of characteristic 2). This time consider the linear group $G := GL_n(2)$ over the field $\mathbb{F}_2$ of characteristic 2. The diagonal matrix of the previous Example 0.3.2 is now just the identity, since $-1 = 1$ in characteristic 2; so to get an involution, we instead take t to be a transvection (i.e., the matrix $t - I_n$ has rank 1—cf. Section A.6). By conjugating

[3]Even this statement is sometimes inaccurate—for example, quasisimplicity fails if $k = 1$ or $n - k = 1$. But it holds for "typical" values of k (and n and q).

in G, we can arrange for t to have the following convenient block-triangular form:

$$t := \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & I_{n-2} & 0 \\ \hline 1 & 0 & 1 \end{array}\right).$$

Again let $H := C_G(t)$ be the centralizer in G of t; it has the corresponding block-triangular form

$$H = \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline * & SL_{n-2}(2) & 0 \\ \hline * & * & 1 \end{array}\right),$$

where $*$ indicates arbitrary elements of $\mathbb{F}_2$. Indeed H is even a semidirect product $H = UL$, where

$$U := \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline * & I_{n-2} & 0 \\ \hline * & * & 1 \end{array}\right), \text{ and } L := \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & SL_{n-2}(2) & 0 \\ \hline 0 & 0 & 1 \end{array}\right).$$

In this example, $U = F(H) = O_2(H)$ is the largest normal nilpotent subgroup of H. Further H has no components at all (in particular the group L is *not* normal). Thus $E(H) = 1$, so that in fact we have $F^*(H) = F(H) = U$; and in particular

$$F^*(H) = O_2(H) \tag{0.3.4}$$

is a 2-group. (This last feature also occurs if we let G be the nearly-simple group $PGL_n(2^m)$ for any $m \geq 1$—and even if we choose t to be *any* involution in G.) Notice furthermore that H is much larger than $F^*(H)$ in this example, so the full structure of H is not so well predicted by the structure of $F^*(H)$. Nevertheless, it is still true that $H/Z\big(F^*(H)\big)$ embeds into $Aut\big(F^*(H)\big)$. Indeed H/U embeds into $GL(\overline{U})$, where $\overline{U}$ is the $\mathbb{F}_2$-vector space $U/\langle t \rangle$. ♢

The Dichotomy Theorem distinguishing odd and even cases. Next, motivated by the two examples of linear groups discussed above, we define two classes of finite groups in terms of the local structure of the groups (rather than in terms of linear representations).

The even case. First, as suggested by $GL_n(2)$ in Example 0.3.3, the p-local subgroups in each simple group G of Lie type over a finite field of characteristic p satisfy condition (0.3.4), except with 2 replaced by p. (Indeed we state this property later as 0.4.1.3.)[4] This suggests the following terminology:

DEFINITION 0.3.5 (Characteristic p type). A finite group G is said to be of *characteristic p type* if $F^*(H) = O_p(H)$ for each p-local subgroup H of G. ♢

The groups in our *even* case are the groups of characteristic 2 type. In particular the groups $GL_n(2)$ in 0.3.3 are of characteristic 2 type, so they are even groups.

[4]In some of the more recent literature, our terminology of "characteristic p type" in 0.3.5 is replaced by the alternative terminology *local characteristic p*; and groups H satisfying the condition in (0.3.4) are said to be of *characteristic p*. We have retained the more traditional terminology of 0.3.5, to clearly distinguish this more general class of groups from the class of groups of Lie type in characteristic p.

The odd case. The groups in our "odd case" have involution centralizers similar to those in $GL_n(q)$ with q odd, in Example 0.3.2, in that some centralizer has components. But for our proof below of the Dichotomy Theorem 0.3.10, we will need a certain generalization of the notion of component, which we now introduce after a few preliminaries.

DEFINITION 0.3.6. For p a prime, $O_{p'}(G)$ denotes the largest normal subgroup of G of order coprime to p. And roughly dually, $O^{p'}(G)$ denotes the smallest normal subgroup H of G such that G/H is a p'-group. Equivalently, $O^{p'}(G)$ is the normal subgroup of G generated by all of the p-elements of G. ♢

In our study of the structure of p-local subgroups H of G, the groups $O_{p'}(G)$ and $O_{p'}(H)$ will often loom as possible "obstructions" to various desirable properties.

For example consider the components in $E(\overline{H})$, where $\overline{H} := H/O_{p'}(H)$: Here for $q \neq p$ we have $O_q(\overline{H}) \leq O_{p'}(\overline{H}) = 1$ by definition of $\overline{H}$, so that $F(\overline{H}) = O_p(\overline{H})$; hence $E(\overline{H})$ is the obstruction to the condition $F^*(\overline{H}) = O_p(\overline{H})$ (as in (0.3.4))—*for the quotient* $\overline{H}$ (cf. B.1.9). This motivates the study of:

DEFINITION 0.3.7 (The p-layer and p-components). Define $O_{p',E}(H)$ to be the preimage in H of $E\big(H/O_{p'}(H)\big)$, and set $L_{p'}(H) := O^{p'}\big(O_{p',E}(H)\big)$. We call $L_{p'}(H)$ the *p-layer* of H.

A subnormal subgroup $L = O^{p'}(L)$ of H which surjects onto a component of $E\big(H/O_{p'}(H)\big)$ is called a *p-component* of H. Hence $O_{p'}(L)$, when it is not central in L, is the obstruction to L being an actual component (i.e. quasisimple) of H. ♢

We will see that p-components arise naturally in arguments later in this section,[5] although the p-components in centralizers H in our examples of simple groups G—notably for $G = GL_n(q)$ (q odd) in Example 0.3.2—are actual components. So for expository purposes in this section, where usually $p = 2$, the reader should think of $L_{2'}(H)$ as being roughly the same as $E(H)$. But the statement and proof of the Dichotomy Theorem require the following more general notion:

DEFINITION 0.3.8 (component type). A group G is said to be of *component type* if $L_{2'}\big(C_G(t)\big) \neq 1$ for some involution t in G. ♢

Thus Example 0.3.2 shows that for q odd, $GL_n(q)$ is of component type for all $n \geq 3$—with the exception of $GL_3(3)$ because of the solvability of $SL_2(3)$. Indeed small exceptions such as $GL_2(q)$ and $GL_3(3)$ motivate the following additional definition (which we have named after two of the major contributors to the classification of the corresponding groups):

DEFINITION 0.3.9 (Gorenstein-Walter type). We say that a finite group G is of *Gorenstein-Walter type* (or *GW type*) if either (1) G has 2-rank $m_2(G) \leq 2$, or (2) G is of component type. ♢

The groups in our *odd* case are the groups of Gorenstein-Walter type.

[5] More precisely, it is the *absence* of p-components which makes possible results such as (0.3.21).

The Dichotomy Theorem. Recall that Principle II says that the structure of local subgroups of finite simple groups should be restricted. Our first major result in that spirit shows that the 2-local structure of a simple group satisfies the conditions in either Definition 0.3.5 or Definition 0.3.9, and hence its 2-locals resemble those in Example 0.3.2 or Example 0.3.3:

THEOREM 0.3.10 (Dichotomy Theorem). *Assume that G is of even order and satisfies*
(a) $G = O^2(G)$, and
(b) $O_{2'}(G) = 1$.
Then G is either of Gorenstein-Walter type or of characteristic 2 type.

The Dichotomy Theorem begins to distinguish the 2-local structure of simple groups from that of the general finite group, supplying some progress toward making Principle II precise. Specifically, if G is an arbitrary finite group which is not of component type, and if H is an involution centralizer in G, then by definition H has no 2-components and hence no components, so that $F^*(H) = F(H)$—but a priori, this only says that $F(H)$ is nilpotent, not a 2-group. The contribution of the Dichotomy Theorem is to show, when G is also simple and of 2-rank at least 3, that $F^*(H)$ must in fact be a 2-group (as in (0.3.4)); and indeed this holds for any 2-local subgroup H.

After a brief discussion in Section 0.4 of the finite simple groups, we will further pursue Principle II by giving a somewhat more detailed idea of the structure we expect in 2-locals in simple groups of Gorenstein-Walter type, and in groups of characteristic 2 type.

A proof of the Dichotomy Theorem. But in the remainder of this section, we first sketch a fairly complete proof of the Dichotomy Theorem. Our aims are:

- to demonstrate that the basic ideas underlying the dichotomy are actually fairly elementary (admittedly modulo assuming certain standard results); and
- to introduce, along the way, some further concepts and methods which are of fundamental importance throughout the proof of the Classification Theorem.

Some general preliminaries. In the next few lemmas, we let the superscript $*$ denote images under the natural surjection of G onto $G^* := G/O_{p'}(G)$.

We had mentioned after 0.3.6 that the subgroups $O_{p'}(H)$ of p-local subgroups H of G can be an obstruction to establishing various properties of G. To help overcome this problem, we would like to show that generically the image $O_{p'}(H)^*$ is small (ideally even trivial).

Such analysis uses results on “coprime action”, as in Section B.1 of Appendix B; here we mention two useful elementary facts (cf. Lemma B.1.8):

LEMMA 0.3.11 (Coprime Action). *Let P be a p-subgroup of G. Then*
(1) $N_{G^*}(P^*) = N_G(P)^*$.
(2) If X is a p'-subgroup of G normalized by elementary abelian P, then:
$$X = \langle C_X(Q) : Q \leq P \text{ with } P/Q \text{ cyclic}\rangle.$$

PROOF. See e.g. 18.7.4 in [**Asc00**], and 11.13 in [**GLS96**]. □

There is a global-implies-local principle (see also Lemma B.1.4) for the condition in (0.3.4):

LEMMA 0.3.12. *If $F^*(G) = O_p(G)$, then G is of characteristic p type.*

PROOF. See e.g. 31.16 in [**Asc00**]—which is an easy consequence of Bender's self-centralizing property 0.3.1 of $F^*(G)$, and the Thompson $A \times B$-Lemma B.1.1. □

This in turn leads to a similar global-implies-local principle for the triviality of the layer of H; and indeed to a sufficient condition for the triviality of the image $O_{p'}(H)^*$ in G^* of the local obstruction $O_{p'}(H)$:

LEMMA 0.3.13. *Assume that $L_{p'}(G) = 1$. Then for each p-local subgroup H of G, $L_{p'}(H) = 1$ and $O_{p'}(H) \le O_{p'}(G)$.*

PROOF. Let $H = N_G(P)$ for some nontrivial p-subgroup P of G, so by Coprime Action 0.3.11.1, $N_{G^*}(P^*) = N_G(P)^* = H^*$. The hypothesis $L_{p'}(G) = 1$ implies $F^*(G^*) = O_p(G^*)$ (cf. B.1.9). So also $F^*\big(N_{G^*}(P^*)\big) = O_p\big(N_{G^*}(P^*)\big)$ by Lemma 0.3.12—that is, $F^*(H^*) = O_p(H^*)$. So $F^*\big(O_{p'}(H^*)\big) \le O_{p'}\big(F^*(H^*)\big) = 1$, and it follows that $O_{p'}(H^*) = 1$ using 0.3.1. Since $O_{p'}(H)^* \le O_{p'}(H^*)$, we get $O_{p'}(H) \le O_{p'}(G)$, as required. Indeed then $O_{p'}(H) = O_{p'}(G) \cap H$, so that we in fact have $H^* \cong \overline{H} := H/O_{p'}(H)$; hence $F^*(\overline{H}) = O_p(\overline{H})$, and we get $L_{p'}(H) = 1$ by B.1.9, completing the proof. □

Some basics of signalizer functor theory. One focus of our proof is the function $a \mapsto O_{p'}\big(C_G(a)\big)$, defined on the set of elements a of G of order p. In fact this function is just one example among many in p-local analysis which map p-elements or p-subgroups of G to suitable p'-subgroups.[6] We will explore this important topic in more detail later, notably in Section B.3 of the Appendix. We begin here with the notion of a signalizer functor:

DEFINITION 0.3.14 (signalizer functor). Let A be an elementary abelian p-subgroup of G of rank at least 3.[7] A *signalizer functor on A* is a function θ defined on the nontrivial elements of A (we denote this subset by $A^\#$) such that

(S1) For each $a \in A^\#$, $\theta(a)$ is a p'-subgroup of $C_G(a)$, normalized by A.

(S2) For all $a, b \in A^\#$, $\theta(a) \cap C_G(b) \le \theta(b)$. ◇

The definition is due to Gorenstein, with simplifications due to Goldschmidt; see e.g. [**GLS96**, Sec 20]. The condition (S2) is usually called *balance*; it reflects a certain naturality (to which we will return just before Definition 0.3.16) of θ with respect to the inclusions of $\langle a \rangle$ and $\langle b \rangle$ in $\langle a, b \rangle$, and indeed in A. (We consider variations on the notion of balance in Section B.3.)

In proving the Dichotomy Theorem, we will see at (0.3.21) that once we assume that G is not of component type, the absence of 2-components in involution centralizers will cause the abovementioned function $a \mapsto O_{p'}\big(C_G(a)\big)$ to satisfy the balance condition (S2), and so to define a signalizer functor. This will then be a key tool in proving G is of characteristic 2 type.

And to make that tool effective, we will need to *assume* the Signalizer Functor Theorem—a nontrivial but fundamental result due to Gorenstein, Goldschmidt,

[6] Such p'-subgroup images were called *signalizers* by Thompson.

[7] It turns out that the concept is not very interesting for A of rank at most 2.

Glauberman, and McBride (see e.g. [**GLS96**, 21.3, 21.4]). One conclusion of the result, usually called *completeness*, shows that the values of θ generate a p'-subgroup $\theta(A)$; and conversely this *single* group determines $\theta(a)$ via $\theta(a) = \theta(A) \cap C_G(a)$—so that the balance condition (S2) in Definition 0.3.14 is satisfied in an obvious way. We will give a more detailed statement in the Appendix as B.3.13; here is a rough statement for our present introductory exposition:

THEOREM 0.3.15 (Signalizer Functor Theorem). *Let G be a finite group, and assume that θ is a signalizer functor on some elementary abelian p-subgroup A of G with $m_p(A) \geq 3$. For $B \leq A$, set $\theta(B) := \langle \theta(b) : b \in B^\# \rangle$. Then*
(1) $\theta(A)$ is a p'-group.
(2) For each $a \in A^\#$, $\theta(a) = C_{\theta(A)}(a)$.
(3) For each noncyclic subgroup B of A, $\theta(B) = \theta(A)$. □

In fact (3) is a corollary of (1) and (2): We are done if $B = A$, so consider any $a \in A \setminus B$. Then B normalizes $\theta(a)$ by (S1) of Definition 0.3.14; so by Coprime Action 0.3.11.2, $\theta(a)$ is generated by the subgroups $C_{\theta(a)}(C)$, for the maximal subgroups C of the noncyclic group B. And for $b \in C^\#$, using (2) we get $C_{\theta(a)}(C) \leq C_{\theta(A)}(b) = \theta(b) \leq \theta(B)$. Thus $\theta(a) \leq \theta(B)$, and varying a we conclude that $\theta(A) = \theta(B)$.

Next let Ω be the category where subgroups of G are objects, and for $P, Q \leq G$, we take $\mathrm{Mor}(P, Q) := \{g \in G : P^g \leq Q\}$; and let Ω_p be the full subcategory of Ω on $\Lambda_2^p(G)^\circ$. We extend the notion of a signalizer functor[8] to obtain contravariant functors $\overline{\theta}$ and $\overline{C} : B \mapsto C_G(B)$ from Ω_p to Ω (where $\overline{\theta}(b) = \theta(b) \cap C_G(B)$ for $b \in B^\#$ and $\overline{\theta}(g) = \overline{C}(g) = g^{-1}$) such that the inclusion map is a natural transformation from $\overline{\theta}$ to $\overline{C}$.

DEFINITION 0.3.16. Let $\mathcal{I}(G)$ denote the set of elements $a \in G$ of order p which also satisfy $m_p\big(C_G(a)\big) \geq 3$. Define a *$G$-equivariant signalizer functor on $\mathcal{I}(G)$* to be a map θ from $\mathcal{I}(G)$ into the set of p'-subgroups of G such that
(I1) For each $a \in \mathcal{I}(G)$ and $g \in G$, $\theta(a) \leq C_G(a)$ and $\theta(a^g) = \theta(a)^g$.
(I2) For each pair of commuting elements $a, b \in \mathcal{I}(G)$, $\theta(a) \cap C_G(b) \leq \theta(b)$. ◊

Observe that if A is an elementary abelian p-subgroup of G with $m_p(A) \geq 3$, then $A^\# \subseteq \mathcal{I}(G)$. Also (I1) for g in A gives the A-invariance for (S1) of 0.3.14, while (I2) restricted to $A^\#$ is just (S2). Thus $\theta|_A$ is a signalizer functor on A, so the signalizer functor theorem allows us to define $\theta(B)$ for any nontrivial subgroup B of A.

In the generic situation, the commuting graph $\Lambda_2^p(G)^\circ$ (from Notation 0.2.1), on suitable subgroups of rank at least 2, is connected. In this case condition 0.3.15.3 says that $\theta(B) = \theta(B')$ for each B, B' in a connected component of that graph:

LEMMA 0.3.17. *Assume that $m_p(G) \geq 3$, with $\Lambda_2^p(G)^\circ$ connected, and that θ is a G-equivariant signalizer functor on $\mathcal{I}(G)$. Then $\langle \theta(a) : a \in \mathcal{I}(G) \rangle$ is a normal p'-subgroup of G. In particular for each $a \in \mathcal{I}(G)$, $\theta(a) \leq O_{p'}(G)$.*

PROOF. For $E \in \Lambda_2^p(G)^\circ$, set $\theta(E) := \langle \theta(e) : e \in E^\# \rangle$.

For A an elementary abelian p-subgroup of G with $m_p(A) \geq 3$, we saw after Definition 0.3.16 that $\theta|_A$ is a signalizer functor on A. So by the Signalizer Functor

[8] The argument from here on is a simplified version of the material leading up to Theorem B.3.25 in the Appendix.

Theorem 0.3.15, $\theta(A)$ is a p'-subgroup of G, and $\theta(A) = \theta(B)$ for each noncyclic subgroup B of A.

In particular if A' is another elementary abelian subgroup with $m_p(A') \geq 3$ which contains B, then $\theta(A) = \theta(B) = \theta(A')$. So as $\Lambda_2^p(G)^\circ$ is connected by hypothesis, it follows that for each pair $B, D \in \Lambda_2^p(G)^\circ$, $\theta(B) = \theta(D)$; and so we see that $\theta(B) = \langle \theta(a) : a \in \mathcal{I}(G) \rangle$. Then by the G-equivariance condition (I1) of 0.3.16, for $g \in G$, $\theta(B)^g = \theta(B^g) = \theta(B)$, completing the proof. □

Using $O_{p'}\big(C_G(-)\big)$ as a signalizer functor. We can now complete the proof of the Dichotomy Theorem 0.3.10—by using signalizer functor theory to show that groups which are not of Gorenstein-Walter type are of characteristic 2 type. First for any prime p we have:

LEMMA 0.3.18. *Assume that $O_{p'}(G) = 1$, $m_p(G) \geq 3$, and $\Lambda_2^p(G)^\circ$ is connected. Then either*

(1) $L_{p'}\big(C_G(a)\big) \neq 1$ for some $a \in \mathcal{I}(G)$, or

(2) for each $a \in \mathcal{I}(G)$, $F^*\big(C_G(a)\big) = O_p\big(C_G(a)\big)$.

PROOF. We may assume (1) fails, so

$$L_{p'}\big(C_G(a)\big) = 1 \text{ for all } a \in \mathcal{I}(G). \tag{0.3.19}$$

Then in particular we obtain

$$F^*\Big(C_G(a)/O_{p'}\big(C_G(a)\big)\Big) = O_p\Big(C_G(a)/O_{p'}\big(C_G(a)\big)\Big) \tag{0.3.20}$$

using Appendix Lemma B.1.9. We also get:

(0.3.21) (0.3.19) *implies $O_{p'}\big(C_G(-)\big)$ is a G-equivariant signalizer functor.*

For take commuting $a, b \in \mathcal{I}(G)$. It is elementary that

$$O_{p'}\big(C_G(a)\big) \cap C_G(b) \leq O_{p'}\big(C_G(a) \cap C_G(b)\big);$$

and now in view of (0.3.19), we may apply Lemma 0.3.13 to the right-hand side, with $C_G(a) \cap C_G(b)$ and $C_G(b)$ in the roles of "H" and "G", to get

$$O_{p'}\big(C_G(a) \cap C_G(b)\big) \leq O_{p'}\big(C_G(b)\big).$$

Thus we obtain the balance condition (I2) of Definition 0.3.16; while condition (I1) there is immediate.

Using Lemma 0.3.17 and the hypothesis that $\Lambda_2^p(G)^\circ$ is connected, we conclude that $O_{p'}\big(C_G(a)\big) \leq O_{p'}(G)$. Then as $O_{p'}(G) = 1$ by hypothesis, (0.3.20) becomes $F^*\big(C_G(a)\big) = O_p\big(C_G(a)\big)$, giving conclusion (2). □

For $p = 2$, Lemma 0.3.18 will lead to our desired dichotomy—although now in the generic case we require only that the commuting graph on involutions should be connected:

LEMMA 0.3.22. *Assume that $G = O^2(G)$, $O_{2'}(G) = 1$, $m_2(G) \geq 3$, and $\Lambda_1^2(G)$ is connected. Then G is either of component type, or of characteristic 2 type.*

PROOF. We may assume G is not of component type. Thus $L_{2'}\big(C_G(a)\big) = 1$ for all involutions a in G; so as in (0.3.21), we also get that $O_{2'}\big(C_G(-)\big)$ is a signalizer functor. However, so far we have only a weaker connectivity hypothesis than in Lemma 0.3.18.

Furthermore we may assume that $Z^*(G) = 1$:[9] For otherwise, our hypothesis that $O_{2'}(G) = 1$ implies that $Z(G)$ contains an involution x, so that $G = C_G(x)$; and then by the noncomponent type hypothesis, we have $F^*(G) = O_2(G)$, which means that G is of characteristic 2 type by 0.4.1, and so we would be done.

Now, since $p = 2$, $G = O^2(G)$, $Z^*(G) = 1$, and $m_2(G) \geq 3$, our hypothesis of connectivity for $\Lambda_1^2(G)$ implies the connectivity of $\Lambda_2^2(G)$. The proof is elementary but somewhat lengthy, so we have placed it in the Appendix, as the final statement in B.4.11.[10]

Now note that if B and C are (distinct) commuting 4-groups[11] defining an edge of $\Lambda_2^2(G)$, then BC is elementary abelian of rank at least 3; thus $B, C \in \Lambda_2^2(G)^\circ$, and $BC^\# \subseteq \mathcal{I}(G)$. Hence the connectivity of $\Lambda_2^2(G)$ implies that every 4-subgroup of G lies in $\Lambda_2^2(G)^\circ$, so that $\Lambda_2^2(G)^\circ = \Lambda_2^2(G)$, and thus $\Lambda_2^2(G)^\circ$ is connected. Also since every involution is in some 4-group as $m_2(G) > 1$, it follows that every involution of G is in $\mathcal{I}(G)$. Thus as G is not of component type, 0.3.18 now gives $F^*\big(C_G(a)\big) = O_2\big(C_G(a)\big)$ for all a in $\mathcal{I}(G)$, i.e. for all involutions of G.

But (see B.1.6) this condition implies $F^*(H) = O_2(H)$ for all 2-local subgroups H; that is, G is of characteristic 2 type. □

To include the atypical case (where $\Lambda_1^2(G)$ is disconnected) in our proof, we also *assume* the Bender-Suzuki Theorem 0.2.3, which we can use to show:

THEOREM 0.3.23. *Assume G is a finite group of even order, with $G = O^2(G)$ and $O_{2'}(G) = 1$. Then one of the following holds:*
(1) $m_2(G) \leq 2$.
(2) G is of component type.
(3) G is of characteristic 2 type.
In particular this holds for any finite simple group of even order.

PROOF. We may assume that (1) fails, so that $m_2(G) \geq 3$.

Assume first that $\Lambda_1^2(G)$ is disconnected. Then as $G = O^2(G)$ and $O_{2'}(G) = 1$, the general version of the Bender-Suzuki Theorem (which we have stated for simple groups as 0.2.3) says that G is of Lie type in characteristic 2 and of BN-rank 1. So, as we mentioned before Definition 0.3.5 (and state later in 0.4.1.3), G is of characteristic 2 type, so that (3) holds.

Thus we may now assume that $\Lambda_1^2(G)$ is connected. Then 0.3.22 shows that (2) or (3) holds. □

From Definition 0.3.9, G is of Gorenstein-Walter type if it satisfies (1) or (2) of Theorem 0.3.23, so that theorem is a restatement of the Dichotomy Theorem 0.3.10—and hence completes our proof of 0.3.10.

We will also discuss in more detail the role of the Dichotomy Theorem in the proof of the CFSG, later in remarks leading up to 1.3.10.

---

[9]We recall that $Z^*(G)$ is the preimage of $Z\big(G/O_{2'}(G)\big)$.

[10]We mention that the argument at B.4.11 also makes use (via the reference to B.4.10) of the $\mathcal{K}$-group Hypothesis 0.5.4, that all proper simple sections of G are known. By contrast, in Section B.5 we will provide a proof of the Dichotomy Theorem (due essentially to Thompson and Gorenstein) which does not appeal to B.4.11, and hence requires no $\mathcal{K}$-group assumptions. It also replaces the hypothesis that $Z^*(G) = 1$ by the hypothesis that G is not of characteristic 2 type—see especially the argument from just before B.5.3 through B.5.4. However, we felt that the shorter proof given here is more suitable for expository purposes.

[11]A (Klein) 4-*group* (or *fours-group*) is an elementary abelian group of order 4, i.e. $\mathbb{Z}_2 \times \mathbb{Z}_2$.

0.4. The finite simple groups

Denote by $\mathcal{K}$ (for "known") the set of (isomorphism classes of) simple groups appearing in the Classification Theorem 0.1.1. In this section, we give a brief description of the groups in $\mathcal{K}$. Further details on the groups can be found in Appendix A.

The groups of prime order in case (1) of the Classification Theorem are the *abelian* simple groups. For each prime p, there exists a unique such group, the cyclic group $\mathbb{Z}_p$ of integers modulo p under the group operation of addition of congruence classes.

Sporadic groups. Each of the groups in cases (1)–(3) of the Classification Theorem is a member of a natural infinite family of simple groups. But the 26 sporadic groups in case (4) live in no such family that we know of; hence the designation of "sporadic group".

The sporadic groups are usually denoted just via the names of their discoverers. Here we will just list the corresponding names of the groups, leaving a discussion of their properties to Appendix A and the references there. The 26 sporadic groups are:

the five *Mathieu groups* M_{11}, M_{12}, M_{22}, M_{23}, and M_{24};
the four *Janko groups* J_1, J_2, J_3, and J_4;
the three *Conway groups* Co_1, Co_2, and Co_3;
the three *Fischer groups* Fi_{22}, Fi_{23}, and Fi'_{24};
the *Higman-Sims group* HS;
the *McLaughlin group* McL;
the *Lyons group* Ly;
the sporadic *Suzuki group* Suz;
the *Held group* He;
the *Rudvalis group* Ru;
the *O'Nan group* $O'N$;
the *Harada-Norton group* HN (or F_5);
the *Thompson group* Th (or F_3);
the *Baby Monster* BM (or F_2); and
the *Monster* M (or F_1).

Natural structures for action of simple groups. The non-sporadic finite simple groups G can be described as (essentially) the group of automorphisms of a suitable highly symmetric natural object X. The representation of G on X allows us to study the group and establish many of its properties.

Alternating groups. For example, consider the alternating group A_n of degree n, which is the normal subgroup of all even permutations in the symmetric group S_n on a set X of order n. The group S_n is the full automorphism group of X in the category of sets, and A_n is of index 2 in S_n. For $n \geq 5$, A_n is simple, and indeed we have $A_n = F^*(S_n)$.

Lie type groups. The groups of Lie type are responsible for most of the complexity and difficulty in the proof of the Classification Theorem. Thus we devote most of the rest of this section to a discussion of those groups (again with more details postponed to Appendix A). In particular, we begin by indicating several versions of a natural structure "X", that can be used to describe the finite simple groups G of Lie type.

In one approach (via the "Chevalley construction") such a G is (essentially) the group of automorphisms of some suitable polynomial function f on[12] a vector space V over some finite field F. Thus the object X defining G is the pair (V, f). More precisely, if $A := Aut(V, f)$ is the group of automorphisms, and $P : GL(V) \to PGL(V)$ is the projection map of $GL(V)$ into the group of automorphisms of the full projective geometry $PG(V)$ on V, then G arises as the generalized Fitting subgroup $F^*(PA)$ of the full automorphism group PA of the subgeometry X of $PG(V)$.

For another approach via algebraic groups: Let $\overline{F}$ denote the algebraic closure of F; and then write $\overline{V} := \overline{F} \otimes_F V$, and $\overline{f} := \overline{F} \otimes_F f$—a polynomial function on $\overline{V}$. Then $\overline{A} := Aut(\overline{V}, \overline{f})$ is an algebraic group, as is $\overline{G} := F^*(P\overline{A})$. And we can instead regard $(\overline{V}, \overline{f})$ as the fundamental structure "X"—since as we vary the particular finite subfield F of $\overline{F}$, the finite groups A and G arise as forms determined by F of the algebraic groups $\overline{A}$ and $\overline{G}$.

Finally in a third approach, the polynomial f induces a subgeometry $\mathcal{B}$ of the projective geometry $PG(V)$, called the *(Tits) building* of G; and we can also take $\mathcal{B}$ for our structure "X"—indeed with the exception of the case of G of Lie rank 1, we have $G = F^*\big(Aut(\mathcal{B})\big)$.

We give a few specific examples of f and X below.

Some facts about groups of Lie type. We next provide a preview of some material from Remark A.3.1 in Appendix A.

The groups of Lie type live in infinite families, indexed by finite fields, and by certain kinds of polynomials.

The *(ordinary) Chevalley groups* (also called *untwisted groups*) are the groups corresponding to the simple Lie algebras over $\mathbb{C}$ or their Dynkin diagrams. These are the groups

$$A_n(q) \cong L_{n+1}(q), B_n(q) \cong P\Omega_{2n+1}(q), C_n(q) \cong PSp_{2n}(q), D_n(q) \cong P\Omega^+_{2n}(q),$$
$$E_6(q), E_7(q), E_8(q), F_4(q), G_2(q).$$

There are also the *twisted groups*

$$^2A_n(q) \cong U_{n+1}(q), {}^2D_n(q) \cong P\Omega^-_{2n}(q), {}^3D_4(q), {}^2E_6(q),$$
$$^2B_2(2^{2m+1}) = Sz(2^{2m+1}), {}^2G_2(3^{2m+1}) = Ree(3^{2m+1}), {}^2F_4(2^{2m+1})'.$$

Some polynomials f and their natural structures X. For example, the *symplectic groups $PSp_m(q)$* and the *orthogonal groups $P\Omega^\epsilon_m(q)$* are the generalized Fitting subgroup of the (projective) group of isometries of a symplectic or orthogonal form f (respectively), on an m-dimensional vector space V over $F := \mathbb{F}_q$. Thus in this case the polynomial f is of degree 2. These groups, together with the linear groups $L_n(q)$ and the unitary groups $U_n(q)$, are the *classical matrix groups*.

As in Example 0.3.2, we will often take as our standard example of a group of Lie type, the general linear group $G := GL(V) = GL_n(q)$ of all invertible linear maps on an n-dimensional vector space over the field $F := \mathbb{F}_q$. Of course G is usually "nearly-simple" rather than actually simple; but we use it for expository purposes, since minor adjustments to our statements about $GL_n(q)$ will hold for the simple section $L_n(q) = PSL_n(q)$. In the case of the linear group, f can be

[12]Sometimes we will of course mean "on $V \times V$", e.g. in the case of an inner product.

taken to be the zero polynomial; the corresponding algebraic group $\overline{G}$ is $GL(\overline{V})$, and the building $\mathcal{B}$ is the full projective geometry $PG(V)$.

Some notions involving linear groups. We regard our group G of Lie type as a linear group on a vector space V over a finite field F of characteristic p. Thus each $g \in G$ is a linear map on V. We call g *semisimple* if g is diagonalizable over the algebraic closure $\overline{F}$, and we call g *unipotent* if all its eigenvalues are 1. As F is finite of characteristic p, the semisimple elements are the p'-elements, and the unipotent elements are the p-elements of G.

A subgroup H of G is *unipotent* if there is a chain (or *flag*):

$$0 = V_0 \leq V_1 \leq \cdots \leq V_m = V,$$

of subspaces such that H induces the identity on each quotient space V_{i+1}/V_i. The unipotent subgroups are in fact precisely the p-subgroups of G.

The *parabolic subgroups* of G are the stabilizers of simplices in the building $\mathcal{B}$ for G. For example in the case of the classical groups, the simplices are flags[13] of subspaces of V which are totally singular with respect to the relevant form f. Given a parabolic P, the *unipotent radical* of P is the largest normal unipotent subgroup $R(P)$ of P. From the previous paragraph, $R(P) = O_p(P)$.

The group G is transitive on the maximal simplices of $\mathcal{B}$, each of which contains exactly one vertex from each G-orbit on vertices. The number of vertices in a maximal simplex is the *Lie rank* or *BN-rank* of G.

Recall that the *Bender groups*, arising as conclusions in the Bender-Suzuki Theorem 0.2.3, are the groups of Lie type over fields of characteristic 2 which have Lie rank 1; these are the groups $L_2(q)$, $U_3(q)$, and $Sz(q) = {}^2B_2(q)$, q a power of 2.

In our standard example $G = GL(V)$ of a group of Lie type, the simplices of $\mathcal{B}$ are just all flags $\mathcal{F} = (V_0 < \cdots < V_m)$, and the unipotent radical $R(P)$ of the parabolic P stabilizing $\mathcal{F}$ is the unipotent subgroup centralizing each quotient space in the flag $\mathcal{F}$. The Lie rank of G is the length $\dim(V) - 1$ of a maximal flag.

Next for r a prime, we wish to examine the r-local structure of G.

p-local structure. Suppose first that $r = p$, the "natural" characteristic of G (i.e., that of the underlying field F). We mentioned above that the p-subgroups of G are the unipotent subgroups, and the p-elements are the unipotent elements.

The following result is fundamental; its core is the Borel-Tits Theorem (see e.g. [**GLS98**, Sec 3.1]). Conclusion (3) had been mentioned earlier.

LEMMA 0.4.1. *Let H be a nontrivial p-subgroup of G of Lie type in characteristic p. Then:*
(1) There exists a proper parabolic subgroup P of G such that $H \leq R(P)$ and $N_G(H) \leq P$.
(2) $F^*(P) = R(P) = O_p(P)$.
(3) G is of characteristic p type.

PROOF. For part (1), see [**GLS98**, 3.1.3.a]; for (2) see [**GLS98**, 2.6.5.e].

By (2) we may apply Lemma 0.3.12 to the parabolic P containing the 2-local H in (1), to conclude that $F^*(H) = O_p(H)$. Thus (3) holds. □

[13]The notion of "flag" is slightly adjusted in the case of orthogonal spaces of plus type and even dimension.

Note that when $G = GL(V)$, (1) and (2) are easy to prove using some elementary linear algebra, along with the facts that H is unipotent and $\mathcal{B} = PG(V)$ is the full projective geometry.

Next let $\mathcal{M}$ be the set of proper parabolics of G containing a Sylow p-subgroup S of G. Then we can choose the collection of Principle I to be the set of unipotent radicals $\mathcal{Q} := \{R(P) : P \in \mathcal{M}\}$. Indeed the set $\mathcal{N}$ of normalizers of members of $\mathcal{Q}$ in Section 0.2 is actually equal to $\mathcal{M}$, and the simplicial complex $\mathcal{O}(\mathcal{M})$ determined by all cosets of the members of $\mathcal{M}$ gives one version of the building $\mathcal{B}$.

Furthermore a fundamental theorem of Tits (cf. [**Tit81**, p. 541]) states that $\mathcal{B}$ is simply connected if (and indeed only if) the Lie rank of G is at least 3. So in that case, G is determined up to isomorphism by the amalgam $\mathcal{A}(\mathcal{M})$, as discussed in Section 0.2 (cf. also 3.3.14).

r-local structure for $r \neq p$. Next we consider the case when the prime r is different from the natural characteristic p of G. Recall that the r-elements of G are semisimple, that is, diagonalizable over the algebraic closure $\overline{F}$ of F.

For expository purposes, consider our standard example $G = GL(V)$, where we have $g \in G$ of order r, and diagonalizable even over F. Then the r-th roots of 1 $\{\lambda_i : 1 \leq i \leq r\}$ are the (potential) eigenvalues of g, and $V = V_1 \oplus \cdots \oplus V_r$, where V_i is the g-eigenspace corresponding to λ_i. Then $C_G(g) = G_1 \times \cdots \times G_r$, where G_i acts faithfully as $GL(V_i)$ on V_i, and centralizes V_j for $j \neq i$. Subject to our convention that any $GL(U)$ is a nearly-simple group, each factor G_i is nearly-simple, and hence $(G_i : 1 \leq i \leq r)$ are essentially components in $C_G(g)$.

This generalizes our earlier Example 0.3.2, when p is odd and $r = 2$. Indeed, with few exceptions, if G is a group of Lie type over a field of odd characteristic p, then G is of component type. Moreover in this case, we have

$$L_{2'}\big(C_G(g)\big) = E\big(C_G(g)\big); \tag{0.4.2}$$

that is, the 2-components in the definition of component type turn out to be actual components.

In fact, this statement holds for involution centralizers in all finite simple groups; it is essentially the content of Thompson's *B-Conjecture* (cf. 1.6.1). The B-Conjecture makes precise the sense in which, for a 2-local H in a finite group X, the image of $O_{2'}(H)$ in $X/O_{2'}(X)$ is small. We will see in the next section that we need a result of this sort to avoid obstructions in analyzing the 2-local structure of a simple group G arising from the subgroups $O_{2'}(H)$. The proof of the B-Conjecture was one of the important steps in the Classification; we will say a little more about it in the next section (with a fuller discussion later in Sections 1.6–1.8).

Of course our summary of the Classification Theorem 0.1.1 can be expanded into a detailed statement listing the individual groups in $\mathcal{K}$, by adding to conclusions (3) and (4) the specific families of groups Lie type and the specific sporadic groups indicated in this section.

0.5. The Classification grid

The Dichotomy Theorem 0.3.10 partitions the finite simple groups into the groups of Gorenstein-Walter type and the groups of characteristic 2 type. As we saw in Section 0.4, the groups of Lie type over fields of characteristic 2 are of characteristic 2 type, while most groups of Lie type over fields of odd characteristic are of Gorenstein-Walter type. (Indeed, the only exceptions are $PSp_4(3)$, $U_4(3)$, and $G_2(3)$.)

A partition by size. Our first objective in this final section of the Introduction is to give a partition of the groups of each type into large and small groups. This provides the Classification grid 0.5.3 below, which partitions the finite simple groups into four classes.

Size for groups of GW type. When G is of Gorenstein-Walter type, we decree that G is *small* if $m_2(G) \leq 2$, and *large* if $m_2(G) \geq 3$. So from Definition 0.3.9, the large groups are of component type.

REMARK 0.5.1 (2-rank and signalizer functors). The primary reason for this choice is technical: As we saw in Section 0.3, for Principle II we must control the subgroups $O_{2'}(H)$, for H a 2-local subgroup of G; and our principal tool for doing so is the Signalizer Functor Theorem 0.3.15, which requires $m_2(G) \geq 3$. ♢

But there is another important reason for this choice, having to do instead with the recognition of G as in Principle I. Namely the large groups should be our *generic* examples (that is, large enough to exhibit "typical" behavior), and to classify the generic groups using the approach described in Section 0.2, we want a suitable complex of 2-local subgroups to be simply connected. Such complexes will tend to have the homotopy type of the *clique complex* $K_1^2(G)$, whose simplices are the cliques of the graph $\Lambda_1^2(G)$. And a necessary condition for $K_1^2(G)$ to be simply connected is that $m_2(G) \geq 3$ (cf. Theorem 2 in [**Asc93**]).

Size for groups of characteristic 2 *type.* Next suppose G is of characteristic 2 type. In this case the structure of a 2-local H is dominated by the 2-subgroup given by $F^*(H) = O_2(H)$, so it is hard to get a handle on H in terms of inductive knowledge of *simple* groups. For this reason, we would like to transfer attention to a suitable odd prime p, and seek components in centralizers of elements of order p.

However we mentioned in Section 0.2 that no analogue of the Bender-Suzuki Theorem 0.2.3 is available for odd primes, so we must also continue to keep the prime 2 in the picture. In practice this means that the p-local structure of some 2-local H should be rich; and from the discussion above, this means in turn that we want $m_p(H) \geq 3$. This leads to the following definitions, due to Thompson in his work [**Tho68**] on N-groups (roughly, minimal simple groups).

DEFINITION 0.5.2 (2-local p-rank and $e(G)$). Given a finite group G and an odd prime p, define the 2-*local p-rank* of G to be

$$m_{2,p}(G) := \max\{m_p(H) : H \text{ is a 2-local subgroup of } G\}.$$

Then define

$$e(G) := \max_{p \text{ odd}} m_{2,p}(G).$$

♢

As in Remark 0.5.1, Principle II and the Signalizer Functor Theorem 0.3.15 (for odd p) suggest that it is natural to define a group G of characteristic 2 type to be *small* if $e(G) \leq 2$, and to define G to be *large* if $e(G) > 2$. (The small groups G with $e(G) \leq 2$ are usually called *quasithin*.)

Moreover, again Principle I also suggests that this is the correct place to make our subdivision, as we now indicate. Recall that by 0.4.1.3, a natural example of a simple group of characteristic 2 type is a group of Lie type over a field of characteristic 2. In such a group G, the parameter $e(G)$ is usually maximized on a diagonal subgroup in a Borel subgroup—and so is a good approximation of the Lie rank l of G. Moreover as we saw in Section 0.4, given a fixed Sylow 2-subgroup S of G, each 2-local of the form $N_G(Q)$, $1 \neq Q \leq S$, is contained in one of the proper parabolics over S. Now if $\mathcal{N}$ is the set of such parabolics, then we saw in Section 0.2 that we want the complex $\mathcal{O}(\mathcal{N})$ to be simply connected. Further we indicated in Section 0.4 that $\mathcal{O}(\mathcal{N})$ is a version of the building $\mathcal{B}$, which by the result of Tits is simply connected iff $l \geq 3$.

REMARK 0.5.3 (The "grid" for the original proof of the CFSG). In summary, we have a four-part subdivision of the Classification:

	GW type	characteristic 2 type
small	2-rank ≤ 2	$e(G) \leq 2$ (quasithin)
large	component type	$e(G) \geq 3$

$\diamondsuit$

The standard form approach to components in centralizers. We close this chapter by describing more precisely the sorts of restrictions on 2-locals we aim for in Principle II, when G is a large simple group of Gorenstein-Walter type (that is, of component type). Analogous restrictions will apply to large groups of characteristic 2 type.

Some general background. The proof of the Classification Theorem is inductive, so one considers a counterexample G of minimal order to the Theorem. For any proper subgroup H of G, the simple composition factors of H are then known by induction; so G satisfies:

HYPOTHESIS 0.5.4 ($\mathcal{K}$-Group Hypothesis). *Define a finite group G to be a $\mathcal{K}$-*group *if the composition factors of each proper subgroup of G are in the set $\mathcal{K}$ of known simple groups—that is, in the list of the Classification Theorem 0.1.1. (The term $\mathcal{K}$-*proper *is also used .)* $\diamondsuit$

Thus many of the subtheorems used in the proof of the Classification are theorems about $\mathcal{K}$-groups. One such result is the following theorem of Gorenstein and Walter, which provides an important tool in local group theory:

THEOREM 0.5.5 (Gorenstein-Walter L-Balance Theorem [**GW75**]). *Let G be a finite $\mathcal{K}$-group. Then for each p-subgroup P of G, $L_{p'}\big(N_G(P)\big) \leq L_{p'}(G)$. (Indeed if $O_{p'}(G) = 1$, then $L_{p'}\big(N_G(P)\big)$ centralizes $O_{p'}\big(N_G(P)\big)$.)*

PROOF. See e.g. 31.17.2 in [**Asc00**]. $\square$

(We discuss various general notions of "balance" in the Appendix starting at (B.3.20).)

Standard form for groups of component type. In our discussion in (0.4.2) we mentioned that one of the important steps in the proof of the Classification was the proof of the B-Conjecture 1.6.1; our discussion below will assume that the B-conjecture has already been established—and we state the result in the form:

If G is a finite group with $O_{2'}(G) = 1$, then for each 2-local subgroup H of G, we have $L_{2'}(H) = E(H)$.

Now suppose that G is our minimal counterexample to the Classification. Let t and s be commuting involutions in G, and assume that we have a component L of $C_G(\langle t, s\rangle)$. (So in particular we are assuming that G is of component type.)

Then L is a component of $C_{C_G(s)}(t)$. As the B-conjecture holds in G, the L-Balance Theorem 0.5.5 says that $L \leq L_{2'}\big(C_G(s)\big) = E\big(C_G(s)\big)$. Similarly we get $L \leq E\big(C_G(t)\big)$. This embedding of L in $E\big(C_G(s)\big)$ and $E\big(C_G(t)\big)$ leads to a partial ordering defined on the components of involution centralizers; and makes it possible to prove that (generically) if t is an involution in G, and L is a component of $H := C_G(t)$ which is maximal in that ordering, then H is a so-called *standard subgroup* of G.

A standard subgroup satisfies various strong properties (see 1.6.3 for details). In particular, if L is standard then $C_G(L)$ is essentially cyclic, and we will assume $C_G(L)$ *is* cyclic for expository purposes. Thus $H = N_G(L)$, and the kernel of the conjugation map $c : H \to Aut(L)$ is $C_G(L)$ which is cyclic. So the structure of H is controlled by L, and is highly restricted, as desired for Principle II.

Let's consider a variation on Example 0.3.2, now taking G to be the nearly-simple group $SL_n(q)$, with both n and q odd. If we simply order components of involution centralizers by cardinality, then the maximal components correspond to $k = n - 1$ in that Example, and are isomorphic to $L = SL_{n-1}(q)$; they arise when

$$t := \left(\begin{array}{c|c} -I_{n-1} & 0 \\ \hline 0 & +I_1 \end{array}\right).$$

Moreover $H = N_G(L) = C_G(t) \cong GL_{n-1}(q)$, and in fact (when we pass to the quotient $PSL_n(q)$) we see that $C_G(L) = Z(H)$ is the cyclic group of diagonal matrices which are scalar on L.

In this example, assuming $L \cong SL_{n-1}(q)$ is standard should also lead to the identification of G as the simple group $L_n(q)$. More generally, a standard subgroup L from a given family of groups of Lie type of odd characteristic in Section 0.4 should lead to a larger group in the same family. A similar remark will hold for L (and hence G) of alternating type. For large groups of GW type, this is the approach to recognizing the group via its 2-locals, as in Principle I.

An analogue for odd centralizers in characteristic 2 *type.* The analysis in the above example also remains valid if we take G to be $SL_n(2^m)$ for $m > 1$—but now take t instead to be a diagonal element of odd prime order p, which has one eigenspace of dimension $n - 1$: This again yields a component of the form $L \cong SL_{n-1}(2^m)$ in $H = C_G(t)$, with $C_G(L) = Z(H)$ cyclic.

Furthermore we can again hope to produce a standard subgroup in a p-local, as we did above in the centralizer of an involution, since the following generalization of the B-Conjecture in fact holds:

B_p-Theorem: Let G be a finite simple group, p a prime, and g an element of G of prime order p. Then $L_{p'}\big(C_G(g)\big) \leq E\big(C_G(g)\big)$.

In fact the only known proof of the B_p-Theorem involves quoting the full Classification Theorem. However, when G is a simple group of characteristic 2 type, and p

is odd with $m_{2,p}(G) \geq 3$, it is possible to prove a weak version of this generalized B-Conjecture. That result,[14] in conjunction with the L-Balance Theorem 0.5.5, still yields an inclusion of a component L of $C_G(\langle t, s\rangle)$ in $E\big(C_G(s)\big)$ and $E\big(C_G(t)\big)$, and, as before, this leads in turn to a standard subgroup in some p-local for our odd prime p; see *standard type* at 5.1.4.

For G of characteristic 2 type, usually the standard subgroup L will be of Lie type in characteristic 2; and again Principle I will aim to show that G is a larger group in the same Lie family as L.[15]

Summary. Thus, whenever G is "large", it is possible to find an element t of prime order p (with $p = 2$ when G is of GW type, and p is odd when G is of characteristic 2 type) for which we will obtain $H := C_G(t)$ as a standard subgroup of G; whence the structure of H is controlled by the structure of its component L and is highly restricted, as desired for Principle II. We then use these restrictions on the p-local subgroup H to characterize G—as in Principle I.

This completes our quick overview of the general standard-subgroup strategy underlying the proof of the Classification Theorem. In 1.1.1 in the following chapter, we will supply a more detailed outline of an idealized approach to pinning down the centralizer of some element of prime order.

The remainder of the next chapter will then provide a somewhat more detailed historical discussion of the portion of the proof which treats the simple groups of Gorenstein-Walter type. (A far more detailed presentation of that case is given by Gorenstein in [**Gor83**].) The other chapters of the present volume will provide a rather detailed discussion of the other half of the proof of the Classification, namely the part treating the simple groups of characteristic 2 type.

---

[14] See (14-5) in [**GL83**]. The related discussion in Section I.18 there uses the language of "$L_{p'}$-balance" rather than L-balance; cf. our exposition after 5.2.4, and after 5.3.4.

[15] A more refined ordering of components is necessary to make this statement accurate.