1 Introduction

Let $G$ be a group and $\{g_1, \ldots, g_r\}$ a finite subset of $G$. If $G$ is nilpotent, then

(A) $G = G' \langle g_1, \ldots, g_r \rangle$ implies $G = \langle g_1, \ldots, g_r \rangle$;

(B) $G = \langle g_1, \ldots, g_r \rangle$ implies $G' = [G, g_1] \cdots [G, g_r]$, where $G'$ denotes the derived group of $G$ and $\langle g_1, \ldots, g_r \rangle$ the subgroup generated by $\{g_1, \ldots, g_r\}$; for $g \in G$ we write

$$[G, g] = \{[x, g] \mid x \in G\}$$

where $[x, g] = x^{-1}g^{-1}xg$ is the usual commutator.

(A) is an easy folklore result; (B) is also well known, and first appeared in the unpublished 1966 PhD thesis of Peter Stroud; it is a key element in Serre’s proof that subgroups of finite index are open in a finitely generated pro-$p$ group. Neither (A) nor (B) is true in general for groups that are not nilpotent. Rather surprisingly, however, similar results hold without assuming nilpotency, as long the group $G$ is assumed to be finite. These are very much harder, relying in their most general form on the classification of finite simple groups. The main technical results of the paper [NS], which enabled us to generalize Serre’s theorem to all finitely generated profinite groups, imply the following for a finite $d$-generator group $G$:

(C) every element of $G'$ is equal to a product of $f_1(d)$ commutators;

(D) if $G = G_* \langle g_1, \ldots, g_r \rangle$ then $G = \langle g_{ij} \mid i = 1, \ldots, r, j = 1, \ldots, f_2(d, \alpha) \rangle$ with $g_{ij}$ conjugate to $g_i$ for all $i$ and $j$;

here $G_*$ is a certain characteristic subgroup of $G$ with the property that $G/G_*$ is semisimple-by-soluble, and $\alpha = \alpha(G)$ is a certain measure of the complexity of $G$ (the largest $n$ such that $G$ has $\text{Alt}(n)$ as a section). In [NS] we left open the question of whether $f_2$ can be made independent of $\alpha(G)$; it appears as Problem 4.7.1 in the book [S2], where further background may be found.
The primary purpose of this paper is to answer that question, and more
general versions of it, positively. Although at first glance this may seem a mere
technical improvement, we shall see that it has diverse applications. These are
described in more detail below; among them are the new theorems:

- If $G$ is any compact Hausdorff topological group, then every finitely gen-
erated (abstract) quotient of $G$ is finite.
- Let $G$ be a compact Hausdorff group such that $G/G^0$ is (topologically)
finitely generated. Then $G$ has a countably infinite (abstract) quotient if
and only if $G$ has an infinite virtually-abelian (continuous) quotient.

(Here, $G^0$ denotes the connected component of 1 in $G$).

Indeed, what motivated the present work was the need to develop machinery
powerful enough to establish results of this kind for profinite groups, for which
the methods of [NS] are insufficient; the extension to more general compact
groups was then a relatively natural step.

Our second purpose is to provide a new and more streamlined route to the
results of [NS] and [NS2] – including the solution of Serre’s problem on finite-
index subgroups in finitely generated profinite groups – and of [NSP], where it
is proved that in those groups the power subgroups are open. In setting out to
prove stronger results, we have found an approach that is both more unified and
in some respects simpler than the original proofs. Thus in a sense the present
paper is a ‘mark 2’ version of [NS] + [NS2] + [NSP].

In the course of the proofs, we shall quote a few self-contained propositions
from [NS]. Apart from these, this work is independent of [NS]. In particular, we
shall not be needing the difficult structural results about finite simple groups
that form the substance of [NS2]; these are replaced by the material of Subsec-
tion 4.1. Some discussion of the new ideas that we use instead appears at the
end of this introduction.

1.1 Main results on finite groups

In this subsection all groups are assumed to be finite. The minimal size of a
generating set for $G$ is denoted $d(G)$. To a finite group $G$ we associate the
characteristic subgroup $G_0 = \bigcap \{T \triangleleft G \mid G/T \text{ is almost-simple}\}$

$$= \bigcap_{M \in \mathcal{S}} C_G(M)$$

where $\mathcal{S}$ is the set of all non-abelian simple chief factors of $G$ (a group $H$ is
almost-simple if $S \triangleleft H \leq \text{Aut}(S)$ for some non-abelian simple group $S$). We
remark that $G/G_0$ is an extension of a semisimple group by a soluble group
of derived length at most 3, because the outer automorphism group of any
simple group is soluble of derived length at most 3 (strong form of the Schreier
conjecture, see Subsection 1.3.2). (Note that $G_0 = G$ if $\mathcal{S}$ is empty, by the usual
convention.)
1.1.1 Generators

In Section 2 we prove

**Theorem 1.1** Let $G$ be a group and $K \leq G_0$ a normal subgroup of $G$. Suppose that $G = K \langle y_1, \ldots, y_r \rangle = G' \langle y_1, \ldots, y_r \rangle$. Then there exist elements $x_{ij} \in K$ such that

$$G = \langle y_i^{x_{ij}} \mid i = 1, \ldots, r, j = 1, \ldots, f_0 \rangle$$

where $f_0 = f_0(r, d(G)) = O(rd(G)^2)$.

It is clear that the $y_i$ must be assumed to generate $G$ modulo $G'$; the definition of $G_0$ serves to exclude obvious counterexamples of the form $G = K \times \langle y_1, \ldots, y_r \rangle$ where $K$ is simple or $G = \text{Sym}(n)$ with $y_1$ a transposition.

Recall that $G_0 = G$ if every non-abelian chief factor of $G$ has composition length at least 2, in particular if $G$ is soluble; the result in the soluble case was established in [S1].

1.1.2 Commutators

For a subset $X$ of a group $G$, we write

$$X_{[f]} = \{x_1x_2 \ldots x_f \mid x_1, x_2, \ldots, x_f \in X\}.$$ 

The subset $X$ is symmetric if $x \in X$ implies $x^{-1} \in X$.

For subgroups $H, K$ of $G$,

$$[H, K] = \langle [x, y] \mid x \in H, y \in K \rangle.$$ 

**Theorem 1.2** Let $G$ be a group and $\{y_1, \ldots, y_r\}$ a symmetric generating set for $G$. If $H$ is a normal subgroup of $G$ then

$$[H, G] = \left( \prod_{i=1}^{r} [H, y_i] \right)^{f_1}$$

where $f_1 = f_1(r, d(G)) = O(r^2 d(G)) = O(r^3)$.

This is proved in Section 3 together with the following ‘relative’ version, our main result on finite groups:

**Theorem 1.3** Let $G$ be a group, $H \leq G_0$ a normal subgroup of $G$, and $\{y_1, \ldots, y_r\}$ a symmetric subset of $G$. If $H \langle y_1, \ldots, y_r \rangle = G' \langle y_1, \ldots, y_r \rangle = G$ then

$$[H, G] = \left( \prod_{i=1}^{r} [H, y_i] \right)^{f_2}$$

where $f_2 = f_2(r, d(G)) = O(r^6 d(G)^6)$. 

3
This is in effect ‘Key Theorem C’ of [NS], with the fundamental improvement that \( f_2 \) no longer depends on \( \alpha(G) \). In fact Theorem 1.3 simultaneously generalizes all three versions of the said ‘Key Theorem’ (and strengthens them, with our new definition of \( G_0 \)).

A variant of Theorem 1.2 also holds, where \( \{y_1, \ldots, y_r\} \) is merely assumed to generate \( G \) modulo \( C_G(H) \) and \( f_1 = O(r^3) \) is independent of \( d(G) \); the proof is a little more involved and will appear elsewhere.

Sharper estimates for the functions \( f_0, f_1, f_2 \) will appear in the course of the proofs.

1.1.3 Verbal subgroups

A group word \( w \) has width \( m \) in a group \( G \) if every product of \( w \)-values in \( G \) is equal to such a product of length \( m \); here, by \( w \)-value we mean an element of the form \( w(g)^{\pm 1} \) with \( g \in G \), where \( w \) is a word on \( k \) variables. In Subsection 5.3 we show how the following theorem, originally established in [NSP], easily follows from the above results:

**Theorem 1.4** Let \( w \) be a non-commutator word and \( G \) a finite \( d \)-generator group. Then \( w \) has width \( f(w, d) \) in \( G \), where \( f(w, d) \) depends only on \( w \) and \( d \).

1.2 Algebraic properties of compact groups

A compact group (which we take to mean a compact Hausdorff topological group) is an extension \( G \) of a compact connected group \( G^0 \), its identity component, by a profinite group \( G/G^0 \). The Levi-Mal’cev Theorem shows that the connected component is essentially a product of compact Lie groups; this makes it relatively tractable, and most of our attention will be focused on the profinite case.

1.2.1 Finitely generated profinite groups

The significance of uniform bounds relating to all \( d \)-generator finite groups is that they reflect qualitative properties of \( d \)-generator profinite groups. Thus (C) implies that the derived group is closed in every finitely generated profinite group; and the main ‘finite’ results of [NS] were used to show that every subgroup of finite index in a finitely generated profinite group \( G \) is open. A more roundabout argument, using results from [NS] related to (D), was used in [NSP] to show that the ‘power subgroups’ \( G^q \) are open in \( G \). The sharper results now at our disposal yield further dividends when applied in the profinite context.

Routine compactness arguments (recalled in Subsection 5.2) transform Theorems 1.1, 1.2 and 1.3 into the following.

**Theorem 1.5** Let \( G \) be a finitely generated profinite group and \( K \leq G^0 \) a closed normal subgroup of \( G \). Suppose that \( G = K\langle y_1, \ldots, y_r \rangle = G^q\langle y_1, \ldots, y_r \rangle \). Then there exist elements \( x_{ij} \in K \) such that

\[
G = \langle y_i^{x_{ij}} \mid i = 1, \ldots, r, \ j = 1, \ldots, f_0 \rangle
\]
where \( f_0 = f_0(r, d(G)) \).

Here, \( G_0 \) is defined by (11) with \( T \) ranging over open normal subgroups; and \( \overline{X} \) denotes the closure of a subset \( X \) in \( G \). As in the finite case, \( G/G_0 \) is an extension of a semisimple group by a soluble group of derived length at most 3 (a semisimple profinite group is a Cartesian product of finite simple groups).

**Theorem 1.6** Let \( G \) be a profinite group and \( \{y_1, \ldots, y_r\} \) a symmetric (topological) generating set for \( G \). If \( H \) is a closed normal subgroup of \( G \) then

\[
[H, G] = \left( \prod_{i=1}^{r} [H, y_i] \right)^{f_1}_r
\]

where \( f_1 = f_1(r, d(G)) \).

This implies that \([H, G]\) is closed in \( G \), a result already established in [NS].

**Theorem 1.7** Let \( G \) be a finitely generated profinite group, \( H \leq G_0 \) a closed normal subgroup of \( G \), and \( \{y_1, \ldots, y_r\} \) a symmetric subset of \( G \). If \( H \langle y_1, \ldots, y_r \rangle = G' \langle y_1, \ldots, y_r \rangle = G \) then

\[
[H, G] = \left( \prod_{i=1}^{r} [H, y_i] \right)^{f_2}_r
\]

where \( f_2 = f_2(r, d(G)) \).

Why is this important? Suppose that \( N \) is a proper normal subgroup in a group \( G \). If \( G \) is finite, then \( N \) is contained in some maximal normal subgroup \( M \) of \( G \). If \( G/M \) is abelian, then \( NG' \leq M < G \); if not, then \( G/M \) is a simple chief factor of \( G \), so \( M \geq G_0 \) and \( NG_0 \leq M < G \). So far, so trivial. Now suppose that \( G \) is a profinite group: unless we assume that \( N \) is closed in \( G \), we have no grounds to assert that \( N \) is contained in a maximal open normal subgroup – indeed \( N \) could be dense in \( G \). If \( G \) is a finitely generated profinite group, however, we claim that at least one of \( NG' \), \( NG_0 \) is necessarily properly contained in \( G \). For suppose that \( NG' = NG_0 = G \). If \( G \) is topologically generated by \( d \) elements, we can find \( 2d \) elements \( y_1, \ldots, y_{2d} \in N \) such that \( G_0 \langle y_1, \ldots, y_{2d} \rangle = G' \langle y_1, \ldots, y_{2d} \rangle = G \), and Theorem 1.7 (with \( H = G_0 \)) then implies that

\[
[G_0, G] \leq \langle [G_0, y_i], [G_0, y_i^{-1}] \mid 1 \leq i \leq 2d \rangle \leq N.
\]

But then

\[
G = NG' = N[NG_0, G] = N.
\]

Thus we may state

**Corollary 1.8** Let \( G \) be a finitely generated profinite group and \( N \) a normal subgroup of (the underlying abstract group) \( G \). If \( NG' = NG_0 = G \) then \( N = G \).
This is the key to understanding ‘abstract’ normal subgroups. For example, it quickly reduces Serre’s problem on finite-index subgroups ((E) stated below) to the special cases of abelian groups and semisimple groups, where the answer has long been known: see Subsection 5.1. More generally, it shows that if $G$ has a dense proper normal subgroup, then at least one of $G/G'$ or $G/G_0$ has a dense proper normal subgroup; the point is that each of these quotients has relatively transparent structure. This is exploited to good effect in Subsections 5.6 and 5.7.

In Subsection 5.3 we discuss the profinite version of Theorem 1.4.

**Theorem 1.9** [NSP] Let $G$ be a finitely generated profinite group and $w$ a non-trivial non-commutator word. Then the verbal subgroup $w(G)$ is open in $G$.

Such results also imply certain rigidity properties for profinite groups, that is, conditions under which abstract group homomorphisms are forced to be continuous. Let $G$ be a profinite group, $Q \neq 1$ an abstract group, and $f : G \to Q$ a surjective homomorphism, with kernel $N$.

We can restate the main result of [NS] (re-proved in Subsection 5.1) as:

**(E)** If $G$ is finitely generated (topologically) and $Q$ is finite, then $N$ is open.

This is also true if $G$ is a connected compact group instead of profinite: indeed, such a group is divisible, hence has no nontrivial finite quotients at all ([HM], Theorem 9.35).

An immediate consequence of (E) is

**(F)** If $G$ is finitely generated and $Q$ is residually finite, then $N$ is closed, so $Q$ is profinite (with topology inherited from $G/N$ via $f$); hence $Q$ cannot be countably infinite.

Rather surprisingly, it is easy to find countably infinite non-(residually finite) images (if using the axiom of choice counts as ‘finding’): if $\phi : Q_p \to Q$ is any $Q$-vector space epimorphism then $Z_p\phi$ is a countably infinite image of $Z_p$ (in fact it is an exercise, given (F), to show that $Z_p\phi = Q$). This suggests the question: can $Q$ be finitely generated and infinite? This is answered below.

### 1.2.2 Compact groups

Many of the above results hold more generally for compact groups $G$, assuming usually that the profinite quotient $G/G^0$ is finitely generated ($G^0$ denotes the connected component of the identity in $G$). The structure of a connected compact group is relatively straightforward: it is semisimple modulo its centre (where by a connected compact semisimple group we mean a Cartesian product of compact connected simple Lie groups). In Subsection 5.6 we prove:

**Theorem 1.10** Let $G$ be a semisimple compact group that is either finitely generated profinite or connected. If $Q$ is an infinite quotient of $G$ then $|Q| \geq 2^{2^{\aleph_0}}$. 

6
In the profinite case, we also give a complete classification of the maximal normal subgroups of $G$. Both results depend on associating to each normal subgroup an ultrafilter on the underlying index set of the Cartesian product.

The main results on quotients of compact groups are established in Section 5.6 using Corollary 1.8 and Theorem 1.10.

**Theorem 1.11** Let $G$ be a compact group such that $G/G^0$ is (topologically) finitely generated. Let $N$ be a normal subgroup of (the underlying abstract group) $G$. If $G/N$ is countably infinite then $G/N$ has an infinite virtually-abelian quotient.

**Corollary 1.12** Let $G$ be a compact group such that $G/G^0$ is (topologically) finitely generated. Then $G$ has a countably infinite (abstract) quotient if and only if $G$ has an infinite virtually-abelian (continuous) quotient.

Using (F) in conjunction with Theorem 1.11 it is easy to deduce

**Theorem 1.13** Let $G$ be a compact group and $N$ a normal subgroup of (the underlying abstract group) $G$ such that $G/N$ is finitely generated. Then $G/N$ is finite.

If $G/N$ is a countable quotient of $G$ then the closure of $N$ must be open in $G$; in this case we say that $N$ is virtually dense in $G$. More generally, one might ask: under what conditions is it possible for a normal subgroup of infinite index to be virtually dense? The answer is ‘always’ in abelian groups – for example, $\mathbb{Z}$ is dense in $\mathbb{Z}_p$; and the results of Subsection 5.5 show that a semisimple group can have uncountably many dense normal subgroups. When $G$ is finitely generated profinite, Corollary 1.8 shows that these extreme cases essentially account for all possibilities; when $G$ is connected, the proof of Theorem 1.10 enables us to draw a similar conclusion. Let us say that a semisimple compact group is strictly infinite if it is the product of an infinite set of simple connected Lie groups or finite simple groups. In Subsection 5.7 we prove

**Theorem 1.14** Let $G$ be a compact group such that $G/G^0$ is (topologically) finitely generated. Then $G$ has a virtually dense normal subgroup of infinite index if and only if some open normal subgroup of $G$ has an infinite abelian quotient or a strictly infinite semisimple quotient.

An easy consequence is

**Corollary 1.15** Let $G$ be a finitely generated just-infinite profinite group that is not virtually abelian. Then every normal subgroup of $G$ is closed.

($G$ is just-infinite if $G$ is infinite and every closed non-identity normal subgroup is open. The corollary generalizes a result of A. Jaikin [JZ], who proved it for pro-$p$ groups.)

If $G$ is connected, a virtually dense subgroup is the same thing as a dense subgroup; if $G$ is profinite, however, the conditions for the existence of a proper
dense normal subgroup are more stringent. Their precise characterization (which
depends only on $G/G'$ and $G/G_0$) is stated in our final theorem, whose proof
will appear elsewhere.

1.3 Overview of the paper, conventions, remarks

The basic idea is very simple. Suppose that $G = \langle g_1, \ldots, g_r \rangle$ is a finite group.
If $M$ is a non-central chief factor of $G$ then at least one of the generators $g_i$
must centralize a relatively small proportion of the points of $M$, so the set of
commutators $[M, g_i]$ must be relatively large. Although we can’t predict which
value of $i$ is the relevant one, we can in any case infer that the set

$$\prod_{i=1}^{r}[M, g_i]$$

is relatively large: thus ‘many’ of the elements of $M$ can be expressed as products,
of bounded length, of commutators with the original generators $g_i$.

For this to be of any use, we need to replace ‘many’ with ‘all’. The most
difficult parts of [NS] and [NS2] were devoted to that end; we can now replace
some of those arguments with the help of a new ‘portmanteau’ result, which we
call ‘the Gowers trick’. This is explained below.

For many applications, one needs to have an analogous result for a subset
$\{g_1, \ldots, g_r\}$ which may not generate the whole group. This was achieved in [NS]
(‘Key Theorem C’) only under severe restrictions on the structure of the group
$G$. Somewhat to our surprise, these restrictions turn out to be unnecessary: in
Section 2 we show that the $g_i$ have the necessary ‘fixed-point’ property on chief
factors provided only that $\{g_1, \ldots, g_r\}$ satisfies the hypotheses of Theorem 1.1.
The proof is in principle elementary, relying on the O’Nan-Scott Theorem to
analyse the action of $G$ on its chief factors.

In Section 3 the main results on products of commutators are reduced to
Theorem 4.28: this technical result, the hard core of the paper, concerns a
(quasi-)semisimple group $N$ with operators $y_{ij}$, and shows that every element
of $N$ is equal to a certain product of ‘twisted commutators’ with the $y_{ij}$. The
whole of Section 4 is devoted to the proof of this theorem. While the combinatorial
reduction arguments are still quite complicated, the proofs in Subsection 4.1 of the necessary results about finite simple groups are relatively short and
transparent.

The final Section 5 can be read independently of the rest. Here we derive all
the above-stated applications to topological groups, using only the statements
of Theorems 1.5 – 1.7 and Corollary 1.8 with some additional material relating
to connected compact groups.

The main theorems stated above are not all stated in their sharpest form:
sharper, but less succinct, versions are formulated and proved in the body of the paper.
We take as given the classification of finite simple groups. Some of the main results depend on general consequences of CFSG, such as the facts that finite simple groups can be generated by a bounded number of elements, have bounded commutator width, and have soluble outer automorphism groups (the Schreier conjecture). Others depend on specific properties of groups of Lie type, such as the proportion of regular semisimple elements in these groups, and the detailed structure of their automorphisms. Recent results such as the proof of the Ore Conjecture \cite{LOST}, which says that simple groups have commutator width equal to one, lead to sharper estimates for the implied constants in our main theorems, but are not necessary if one is satisfied with qualitative statements as given above.

1.3.1 The ‘Gowers trick’

A key tool in some of the proofs is a remarkable combinatorial result discovered by Tim Gowers. The basic idea is this: to show that a finite group is equal to the product of some of its subsets, it is enough to know merely that the subsets have sufficiently big cardinalities. We will need the following generalization of Gowers’s result.

For a finite group $G$ let $l(G)$ denote the minimal dimension of any non-trivial $\mathbb{R}$-linear representation of $G$.

**Theorem 1.16** \cite{BNP} Corollary 2.6) Let $X_1, \ldots, X_t$ be subsets of $G$, where $t \geq 3$. Then

$$
\prod_{i=1}^{t} |X_i| \geq |G|^t \cdot l(G)^{2-t} \text{ implies } X_1 \cdot \ldots \cdot X_t = G.
$$

This holds in particular if $|X_i| \geq |G| \cdot l(G)^{-\mu}$ for each $i$, where $t\mu \leq t - 2$.

1.3.2 Facts about simple groups

Here we list some frequently quoted results, for ready reference. Here $S^*$ will denote a quasisimple group (see below) and $S = S^*/Z(S^*)$ a finite (non-abelian) simple group.

**Proposition 1.17** \cite{AG} $S^*$ can be generated by 2 elements.

(This is usually stated for simple groups, but of course any generating set for $S$ lifts to a generating set of $S^*$.)

**Proposition 1.18** \cite{GLS}, Sections 7.1, 2.5) The outer automorphism group $Out(S)$ is soluble of derived length at most 3.

**Proposition 1.19** (i) \cite{W}, Proposition 2.4) There exists $\delta \in \mathbb{N}$ such that every element of $S$ is a product of $\delta$ commutators.

(ii) There exists $\delta^* \in \mathbb{N}$ such that every element of $S^*$ is a product of $\delta^*$ commutators.
(ii) follows from (i) by a theorem of Blau [3], which asserts that every element of $Z(S^*)$ is a commutator unless $S^*$ is one of finitely many exceptions.)

**Corollary 1.20** $S^*$ can be generated by $2\delta$ commutators.

For the record, we recall the validity of the Ore Conjecture (not strictly necessary for our results but yielding better values for the constants):

**Proposition 1.21** ([LOST1, LOST2]) $\delta = 1$, $\delta^* = 2$.

**Proposition 1.22** ([LaS], [KLL] Table 5.3A.) Let $S^*$ be a quasisimple group of Lie type, of untwisted Lie rank $r$ over $F_q$ where $q^r > 27$. Then $l(S^*) \geq (q^r - 1)/2$.

**Proposition 1.23** ([LiSh]) There is an absolute constant $c'$ such that: if $Y$ is a normal subset of $S$ then

$$|Y|^n \geq |S| \implies Y^c n = S.$$ 

It is convenient to define the rank of a simple group as follows: if $S$ is of Lie type, rank($S$) is the (untwisted) Lie rank of $S$; if $S \cong \text{Alt}(n)$, rank($S$) = $n$; if $S$ is sporadic, rank($S$) = 0. The next result is essentially a special case of the main theorem of [BCP]:

**Proposition 1.24** If $C$ is a proper subgroup of $S$ then $|S:C| \geq |S|^{\varepsilon(r)}$ where $\varepsilon(r) > 0$ depends only on $r = \text{rank}(S)$.

### 1.3.3 Notation

For a group $G$, the centre is $Z(G)$ and the derived group is $G'$. For $n > 1$, $G^{(n)} = (G^{(n-1)})'$ where $G^{(1)} = G'$.

For a subset $X$ and an element $y$ of $G$, $[X, y]$ denotes the set $\{[x, y] \mid x \in X\}$. When $X$ and $Y$ are both subgroups of $G$, $[X, Y]$ denotes the subgroup $\langle [x, y] \mid x \in X, y \in Y \rangle$. In particular, the terms of the lower central series are defined by $\gamma_1(G) = G$, $\gamma_2(G) = G'$, and for $n > 1$

$$\gamma_n(G) = [G, \gamma_{n-1}(G)].$$

$\gamma(G) = \bigcap_{n=1}^\infty \gamma_n(G)$ is the nilpotent residual of $G$. If $G$ is finite, then for some $n$ we have $\gamma_\omega(G) = \gamma_n(G) = [\gamma(G), G]$.

The notation $G^{(m)}$ is also used for the Cartesian power $G \times \cdots \times G$ with $m$ factors; which meaning is intended should be clear from the context. For $a, b \in G^{(m)}$ and $\alpha \in \text{Aut}(G)^{(m)}$,

$$a \cdot b = (a_1 b_1, \ldots, a_m b_m)$$

$$[a, \alpha] = ([a_1, \alpha_1], \ldots, [a_m, \alpha_m])$$

$$c(a, \alpha) = \prod_{j=1}^m [a_j, \alpha_j]$$
where as usual \([a, β] = a^{-1}a^β\).

In sections 2 - 4, 'group' means 'finite group', and 'simple group' means 'non-abelian simple group'.

A direct (or Cartesian) product of simple groups is called semisimple. A group \(G\) is quasisimple if \(G\) is perfect (i.e. \(G = G'\)) and \(G/Z(G)\) is simple. A central product of quasisimple groups is called quasi-semisimple.

For a topological group \(G\), the connected component of the identity is denoted \(G^0\) (not to be confused with \(G_0\) defined above (1)).

For \(m \in \mathbb{N}\) we write \([m] = \{1, \ldots, m\}\).

When, occasionally, a lemma is stated without proof, it can be verified by a short direct calculation.

2 Generators

2.1 Fixed-point properties

We begin by defining a key technical concept, in three flavours: the fixed-point property (fpp), the fixed-point space property (fsp), and the fixed-group property (fgp):

**Definition** Let \(Ω\) be a finite \(G\)-set, \(V\) a finite-dimensional \(kG\)-module (\(k\) some field), and \(M\) a \(G\)-group (group acted on by \(G\)). Let \(ε \in (0, 1]\). An element \(y \in G\) has the

- \(ε\)-fpp on \(Ω\) if \(y\) moves at least \(ε |Ω|\) points of \(Ω\),
- \(ε\)-fsp on \(V\) if \(\dim V(y - 1) ≥ ε \dim V\).

If \(Y\) is a subset of \(G\), we say that \(Y\) has the \(ε\)-fpp etc. if there exists \(y \in Y\) having the given property.

- \(Y\) has the \(ε\)-fgp on \(M\) if (i) \(M = S_1 × \cdots × S_n\) with \(n ≥ 2\) and the action of \(G\) permutes the factors \(S_i\) transitively, and (ii) for each such decomposition of \(M\), the set \(Y\) has the \(ε\)-fpp on the set \(\{S_1, \ldots, S_n\}\).

**Remarks.** (i) Each \(ε\) property implies the corresponding \(ε'\) property for any \(ε' ≤ ε\). If \(Y\) acts non-trivially on \(Ω\), respectively \(V\), then \(y\) has the \(2/|Ω|\)-fpp on \(Ω\) and the \(1/\dim(V)\)-fsp on \(V\).

(ii) Suppose that \(G\) is imprimitive on \(Ω\) and acts transitively on a set \(Ω\) of blocks. If \(y\) has the \(ε\)-fpp on \(Ω\) then \(y\) has the \(ε\)-fpp on \(Ω\).

(iii) If \(M\) is a \(G\)-group and \(y\) has the \(ε\)-fgp on \(M\), then \(|C_M(y)| ≤ |M|^{1−ε/2}\).

(iv) Suppose that \(G\) acts as an imprimitive linear group on \(V\), permuting a system of imprimitivity \(Ω\) transitively. If \(y\) has the \(ε\)-fpp on \(Ω\) then \(y\) has the \(ε/2\)-fsp on \(V\).

(v) Say \(|Y| = r\). If \(C_V(⟨Y⟩) = 0\) then \(Y\) has the \(1/r\)-fsp on \(V\); if \(⟨Y⟩\) has no fixed points in \(Ω\) then \(Y\) has the \(1/r\)-fpp on \(Ω\).
These are all easy to see; for (iii) and (iv), suppose that $M = S_1 \times \cdots \times S_n$ is a $G$-group and $y \in G$ permutes the factors $S_i$, according to a permutation with $r$ cycles, including exactly $k$ cycles of length 1. Choose representatives $i(1), \ldots, i(r)$ for these cycles. Then any fixed point of $y$ in $M$ is determined by its projections to $S_{i(1)}, \ldots, S_{i(r)}$, so $|C_M(y)| \leq |S|^{r} = |M|^{r/n}$ if $S_1 \cong \cdots \cong S_n \cong S$.

On the other hand, we have $r \leq (n + k)/2 \leq n(1 - \varepsilon/2)$ if $y$ has the $\varepsilon$-fpp on $\{S_1, \ldots, S_n\}$. This gives (iii), and (iv) is similar, using $\dim V$ in place of $|M|$.

We recall the

**Definition.**

$$G_0 = \bigcap_{M \in \mathcal{S}} C_G(M)$$

where $\mathcal{S}$ denotes the set of all (non-abelian) simple chief factors of $G$.

Recall also that $\delta$ is a number such that every quasisimple group can be generated by $2\delta$ commutators. Note that we can take $\delta = 1$ (see Subsection 1.3.2).

**Theorem 2.1** Suppose $G = G' \langle Y \rangle = G_0 \langle Y \rangle$. Then $Y$ has the $\varepsilon/2$-fsp on every non-central abelian chief factor of $G$ and the $\varepsilon$-fgp on every non-abelian chief factor of $G$ inside $G_0$, where

$$\varepsilon = \min \left\{ \frac{1}{1 + 6\delta}, \frac{1}{|Y|} \right\}.$$

**Reductions.** Let $M = S_1 \times \cdots \times S_n$ be a non-abelian chief factor of $G$, where $n > 1$ and $G$ permutes $\Omega = \{S_1, \ldots, S_n\}$ transitively. Let $\Omega$ be a primitive quotient of the $G$-set $\Omega$. Suppose that $|\Omega| = 2$. Then $G'$ acts trivially on $\Omega$, so $\langle Y \rangle$ acts transitively on $\Omega$, and it follows by Remarks (i) and (ii) that $Y$ has the 1-fpp on $\Omega$. Thus $y$ has the $\varepsilon$-fgp on $M$.

Let $V$ be a non-central abelian chief factor of $G$, so $G$ acts as an irreducible $\mathbb{F}_p$-linear group on $V$.

(i) Suppose that this action is not primitive, so it induces a primitive permutation action of $G$ on a system of imprimitivity $\Omega$. If $|\Omega| = 2$ then as above we may deduce that $Y$ has the 1-fpp on $\Omega$, hence the 1/2-fsp on $V$, by Remark (iv).

(ii) Suppose that $V$ is not inside $G_0$. Then $G_0$ centralizes $V$, so $V$ is a non-trivial simple $\mathbb{F}_p \langle Y \rangle$-module, and then $Y$ has the $\varepsilon$-fgp on $V$ by Remark (v).

(iii) Suppose that $\dim_{\mathbb{F}_p} V = 1$. Then $G'$ centralizes $V$, whence $V(y - 1) = V$ for some $y \in Y$; thus $Y$ has the $\varepsilon$-fgp on $V$.

Arguing by induction on the number of non-central factors in a chief series of $G$ inside $G_0$, it will therefore suffice to prove the following proposition.
The hypotheses imply that every abelian normal subgroup of $G$ acts freely on $V$.

Theorem 5.4.9) that $F$, with each $S_i$ simple and $n \geq 2$, then $\langle Y \rangle$ does not normalize every $S_i$. Put $\varepsilon = \min\{1/(1 + 6\delta), 1/r\}$. Then $Y$ has the $\varepsilon$-fsp on every primitive irreducible $F\rho G$-module of dimension at least two, and the $\varepsilon$-fpp on every primitive $G$-set of size at least 3.

2.1.1 Primitive modules

Let $G$ be a group and $Y$ a subset of $G$ of size $r \geq 1$ such that $G = G' \langle Y \rangle = G_0 \langle Y \rangle$. Suppose that $\langle Y \rangle$ does not centralize any non-central abelian chief factor of $G$, and that if $M = S_1 \times \cdots \times S_n$ is a non-abelian chief factor of $G$, with each $S_i$ simple and $n \geq 2$, then $\langle Y \rangle$ does not normalize every $S_i$.

Proposition 2.2 Let $G$ be a group and $Y$ a subset of $G$ of size $r \geq 1$ such that $G = G' \langle Y \rangle = G_0 \langle Y \rangle$. Suppose that $\langle Y \rangle$ does not centralize any non-central abelian chief factor of $G$, and that if $M = S_1 \times \cdots \times S_n$ is a non-abelian chief factor of $G$, with each $S_i$ simple and $n \geq 2$, then $\langle Y \rangle$ does not normalize every $S_i$. Put $\varepsilon = \min\{1/(1 + 6\delta), 1/r\}$. Then $Y$ has the $\varepsilon$-fsp on every primitive irreducible $F\rho G$-module of dimension at least two, and the $\varepsilon$-fpp on every primitive $G$-set of size at least 3.

2.1.1 Primitive modules

Let $G$ be a group and $Y$ a subset of $G$ of size $r$ satisfying the hypotheses of Proposition 2.2. Let $V$ be a primitive irreducible $F\rho G$-module of dimension at least two; we may assume that $G$ acts faithfully on $V$. Put $F = \text{Fit}(G)$, the Fitting subgroup of $G$.

Lemma 2.3 Let $S$ be a quasisimple subgroup of $G$ and $y$ an element of $G$ such that $[S, S^y] = 1$. Then there exist $a_j, b_j \in S$ such that $S \leq \langle y, y^{a_j}, y^{b_j}, y^{a_jb_j} | 1 \leq j \leq 2\delta \rangle$.

Proof. For $u, v \in S$ we have

$[u^{-1}, v] = [[u, y], v] = y^{-1}yuy^{-uv}y^v$.

The lemma follows since $S$ is generated by $2\delta$ commutators.

Lemma 2.4 If $y \in G$ satisfies $[F, y] \neq \{1\}$ then $y$ has the $\frac{1}{2}$-fsp on $V$.

Proof. (cf. [GSS], proof of Theorem 5.3) For $x \in G$ put $c(x) = \dim C_V(x)$. As $C_V([x_1, x_2]) \geq C_V(x_1) \cap C_V(x_2) \cdot x_2$ we have

$c([x_1, x_2]) \geq 2c(x_1) - \dim(V). \quad (2)$

The hypotheses imply that every abelian normal subgroup of $G$ is cyclic and acts freely on $V$. It follows by a theorem of P. Hall (see [AS], 23.9 or [Go], Theorem 5.4.9) that $F$ is metabelian. Thus if $1 \neq t \in F' \cup Z(F)$ then $c(t) = 0$.

Now there exists $x \in F$ such that $[x, y] \neq 1$. If $[x, y] \in Z(F)$ we may infer using (2) that $c(y) \leq \frac{1}{2} \dim(V)$. If $[x, y] \notin Z(F)$ then for some $h \in F$ we have $1 \neq [[x, y], h] \in F'$. Then using (2) twice gives

$c(y) \leq \frac{1}{2}(c([x, y] + \dim(V)) \leq \frac{1}{2}\left(\frac{1}{2} \dim(V) + \dim(V)\right) = \frac{3}{4}\dim(V)$.

The result follows.

In view of the preceding lemma, we may suppose for the rest of this subsection that $[F, Y] = \{1\}$. Since $Y$ does not centralize any non-central abelian chief factor of $G$, this implies that $F$ is contained in the hypercentre of $G$, and hence that $[F, \gamma_\omega(G)] = 1$. But $G = G' \langle Y \rangle$ implies $G = \gamma_\omega(G) \langle Y \rangle$; therefore $[F, G] = 1$, and so $F = Z(G)$. 

13
Now let $F^* = F^*(G_0)$ denote the generalized Fitting subgroup of $G_0$ (see [AS], Section 31). Then
\[ C_{G_0}(F^*) = Z(F^*) = F \cap G_0. \]

**Case 1.** Suppose that $F^* \leq F$. Then $F^*$ is central in $G_0$ and it follows that $G_0 = F \cap G_0 \leq Z(G)$. Hence $C_V(\langle Y \rangle)$ is a $G$-submodule of $V$; as $V$ is faithful and irreducible for $G$ and $\langle Y \rangle \neq 1$ it follows that $C_V(\langle Y \rangle) = 0$. Hence $Y$ has the $1/r$-fsp on $V$ by Remark (v).

**Case 2.** Suppose that $F^* \not\leq F$. Then $F^* = E \cdot F_0$ where $E$ is a non-empty central product of quasisimple groups, $F_0 = F \cap G_0$, and $E$ is characteristic in $G_0$ with centre $Z_0 = E \cap F$. Let $N = N/Z_0$ be a minimal normal subgroup of $G/Z_0$ contained in $E/Z_0$. Then $N = S_1 \times \cdots \times S_n$, where each $S_i = S_i/Z_0$ is a simple group. By hypothesis, there exists $y \in Y$ such that $y$ moves at least one of these factors; say $y$ moves $S_1$. Then $[S_1, S_1^y] = 1$, and Lemma 2.3 now shows that $S_1 \leq \langle y_1, \ldots, y_t \rangle$ where $t = 1 + 6\delta$ and each $y_j$ is a conjugate of $y$.

We claim that $C_V(S_1) = 0$. Accepting the claim for now, it follows by Remark (v) that some $y_j$ has the $1/t$-fsp on $V$; as $y_j$ is conjugate to $y$ we may conclude that $y$ has the $1/t$-fsp on $V$.

Since $V$ is a primitive irreducible $\mathbb{F}_pG$-module it is a direct sum of copies of some simple $\mathbb{F}_pN$-module $W$. If $C_V(S_1) \neq 0$ then $W$ is a composition factor of the $\mathbb{F}_pN$-module $C_V(S_1)$, so $W(S_1 - 1) = 0$. But then $V(S_1 - 1) = 0$, a contradiction since $V$ is faithful for $G$. Thus $C_V(S_1) = 0$ as claimed.

The first claim of Proposition 2.2 clearly follows.

### 2.1.2 Primitive $G$-sets

Let $G$ be a group and $Y$ a subset of $G$ of size $r$ satisfying the hypotheses of Proposition 2.2. Let $\Omega$ be a primitive $G$-set of size $n \geq 3$, on which $G$ acts faithfully.

If $\langle Y \rangle$ has no fixed points in $\Omega$ then $Y$ has the $1/r$-fpp on $\Omega$, by Remark (v). We assume henceforth that $\langle Y \rangle$ has at least one fixed point in $\Omega$; since $G = G_0(\langle Y \rangle)$ is transitive this implies also that $G_0 \neq 1$.

According to [DM] Theorem 4.3B (part of the O'Nan-Scott Theorem), one of the following holds:

(a) $G$ has a unique minimal normal subgroup $N = C_G(N)$ and $N$ acts regularly on $\Omega$;

(b) $G$ has exactly two minimal normal subgroups $N$ and $C_G(N)$, and each of them acts regularly on $\Omega$;

(c) $G$ has a unique minimal normal subgroup $N$ and $C_G(N) = 1$. 

14
Since $G_0 > 1$, in cases (a) and (c) we have $N \leq G_0$; in case (b) at least one of $N$ and $C_G(N)$ must lie in $G_0$, and we choose to call that one $N$.

**Case 1.** Suppose that the minimal normal subgroup $N$ of $G$ contained in $G_0$ acts regularly on $\Omega$. Then $|N| = n$ and $N$ is a non-central chief factor of $G$ ([DM], Theorem 4.3B). Let $\alpha \in \Omega$ be a fixed point for $\langle Y \rangle$. Then for $x \in N$ and $y \in Y$ we have $(\alpha x)y = ax^y$, so $y$ has exactly $|C_N(y)|$ fixed points on $\alpha N = \Omega$.

By hypothesis, there exists $y \in Y$ such that $C_N(y) \neq N$. The number of fixed points of $y$ in $\Omega$ is then at most $|C_N(y)| \leq \frac{1}{2} |N| = \frac{1}{2}n$, so $y$ has the $\frac{1}{2}$-fpp on $\Omega$.

**Case 2.** The unique minimal normal subgroup $N$ of $G$ is not regular on $\Omega$. Then $N$ is not abelian, so $N = S_1 \times \cdots \times S_m$ where each $S_i$ is simple and $m \geq 2$ since $N \leq G_0$. According to [DM] Theorem 4.6A there are now two possibilities.

**Subcase 2.1.** $G$ acts as a group of diagonal type on $\Omega$. Fixing an identification of each $S_i$ with a group $T$, we identify $\Omega$ with the right coset space $T^* \setminus T^m$ where $T^*$ denotes the diagonal subgroup. The action of $N$ is induced by the right regular action, so

$$T^*(t_1, \ldots, t_m) \cdot s_1 \ldots s_m = T^*(t_1 s_1, \ldots, t_m s_m)$$

for $(t_1, \ldots, t_m) \in T^m$ and $s_i \in S_i$. Write $k = |T|$, so that $n = k^{m-1}$.

Let $\alpha = T^*(t_1, \ldots, t_m)$ be a fixed point for $y \in G$. The stabilizer of $\alpha$ in $N$ is

$$N_\alpha = \{ (u^{t_1}, \ldots, u^{t_m}) \mid u \in T \},$$

so for $x \in N$ we have

$$(ax) \cdot y = ax \iff ax^y = ax \iff x^y x^{-1} = (u^{t_1}, \ldots, u^{t_m}),$$

some $u \in T$. (3)

Suppose that the conjugation action of $y$ permutes $S_1, S_2, \ldots, S_e$ cyclically, and that (3) holds with $x = s_1 s_2 \cdots s_m$ ($s_i \in S_i$). Then $s_2, \ldots, s_e$ are uniquely determined by $u$ and $s_1$. Thus if $y$ has $q = q(y)$ cycles in its action on $\{ S_1, S_2, \ldots, S_m \}$, then the number of $x \in N$ satisfying (3) is at most $k \cdot k^q$. The mapping $x \mapsto \alpha x$ from $N$ to $\Omega$ is surjective and each fibre has size $k$. It follows that $y$ has at most $k^q$ fixed points in $\Omega$.

Suppose that some $y \in Y$ moves at least 3 of the $S_i$. Then $q(y) \leq m - 2$, and so the number of fixed points of $y$ in $\Omega$ is at most

$$k^{q(y)} \leq k^{m-2} = nk^{-1}.$$
If this holds for no element \( y \in Y \), then \( Y \) must contain an element \( y_1 \) that acts as a transposition (12), say, on \( \{ S_1, S_2, \ldots, S_m \} \).

Assume first that \( m \geq 3 \). There exists \( g \in G \) such that \( S_3^g = S_3 \); then \( y = [y_1, g] \) moves at least 3 of the \( S_i \), and hence fixes at most \( nk^{-1} \) points in \( \Omega \). It follows that \( y_1 \) has at most \((n + nk^{-1})/2 \) fixed points.

Suppose now that \( m = 2 \). Set \( y = [y_1, a] \) where \( 1 \neq a \in S_1 \). Suppose that \( y \) fixes \( \alpha \in \Omega \). Each element of \( \Omega \) can be put uniquely in the form \( \alpha x \) with \( x = (s, 1) \), \( s \in S_1 \), and then (3) gives \((\alpha x) \cdot y = \alpha \) \iff \( (s^{\alpha_1}, u_1) = (s^t, u_2) \) \iff \( s \in C_{S_1}(a) \).

Thus \( y \) has at most \(|C_{S_1}(a)| \leq \frac{1}{5} |S_1| = \frac{1}{5} n \) fixed points in \( \Omega \). It follows that \( y_1 \) has at most \( \frac{3}{5} n \) fixed points.

Thus in any case, we may conclude (since \( k \geq 60 \)) that \( y \) contains an element with the \( \varepsilon \)-fpp as long as \( \varepsilon \leq \frac{2}{5} \).

Subcase 2.2. \( G \) is contained in a wreath product \( W = H \wr \pi(G) \) where \( H \leq \text{Sym}(\Gamma) \), \( \pi : G \to \text{Sym}(d) \) where \( d > 1 \), and \( W \) acts on \( \Omega = \Gamma(d) \) by the product action. In this case \( N = N_1 \times \cdots \times N_d \leq H^{(d)} \), and \( G \) permutes the factors \( N_i \) via \( \pi \). Put \( k = |\Gamma| \), so \( n = k^d \). Note that \( k \geq 5 \) since \( N \) is not soluble.

Suppose that \( y = b \cdot \pi(y) \) fixes \((\gamma_1, \ldots, \gamma_d) \in \Gamma^{(d)} \), where \( b \in H^{(d)} \). If \( \pi(y) \) has a cycle \((1, 2, \ldots, e) \) then \( \gamma_{i+1} = \gamma_i^{b_i} \) for \( i = 1, \ldots, e - 1 \). Thus if \( \pi(y) \) has \( q = q(y) \) cycles then the number of fixed points of \( y \) in \( \Omega \) is at most \( k^q \).

By hypothesis, there exists \( y \in Y \) such that \( \pi(y) \neq 1 \). Then \( q(y) \leq d - 1 \) and so \( y \) has at most \( k^{d-1} \leq n/5 \) fixed points in \( \Omega \). Thus \( y \) has the \( \frac{3}{5} \)-fpp on \( \Omega \).

The proof of Proposition 2.2 is now complete.

### 2.2 Small chief factors

We quote a mild generalization of a well-known result due to Gaschütz [Gch] ; the proof given (for example) in [FJ], Lemma 15.30 adapts easily to yield this version:

**Lemma 2.5** Let \( Y_1 \subseteq G \) and \( D \lhd G \). Suppose that

\[
G = D \langle y_1, \ldots, y_d, Y_1 \rangle
\]

where \( d \geq d(G) \). Then there exist \( h_1, \ldots, h_d \in D \) such that \( G = \langle h_1y_1, \ldots, h_4y_d, Y_1 \rangle \).

We have defined \( \delta \) to be a number such that each element of every simple group is a product of \( \delta \) commutators, and observed that in fact one can take \( \delta = 1 \) (Subsection 1.3.2).
Lemma 2.6 Suppose that $M = S_1 \times S_2$ and $\alpha \in \text{Aut}(M)$ satisfies $S_1^\alpha = S_2$, $S_2^\alpha = S_1$. Let $C = \{[x,y] \mid x, y \in S_1\}$. Then

$$C \subseteq [M, \alpha]^{*4}.$$  

If $S_1$ is simple then

$$M = [M, \alpha]^{*8\delta}.$$  

Proof. Let $x, y \in S_1$. Then

$$[y, x^{-1}] = [y, [x, \alpha]] = [y, \alpha][y, \alpha^{-x}][[y, \alpha^{-x}], \alpha] = [y, \alpha] \cdot [y^{x^{-1}}, \alpha^{-1}] \cdot [[[y, \alpha^{-x}], \alpha]$$

and the middle factor lies in $[M, \alpha]^{*2}$ because for any $z \in S_1$ we have

$$[z, \alpha^{-1}]^{x} = [zx, \alpha^{-1}][x^{\alpha^{-1}}, \alpha] = ([zx]^{\alpha^{-1}}, \alpha)[x^{\alpha^{-1}}, \alpha]$$

(for the final equality note that $(zx)^\alpha$ commutes with $zx$). This establishes the first claim.

If $S_1$ is simple, then $S_1 = C^{*\delta}$, so $M = C^{*\delta} \cdot (C^{*\delta})^\alpha \subseteq [M, \alpha]^{*8\delta}$ since $[M, \alpha] = [M, \alpha]^\alpha$. ■

For technical reasons, we need to introduce a slightly smaller analogue of the subgroup $G_0$:

**Definition** For a group $G$, let

$$G_1 = \bigcap_{M \in \mathcal{C}(G)} C_G(M)$$  

(4)

where $\mathcal{C}(G)$ denote the set of all non-abelian chief factors of $G$ that have composition length at most two. We shall call such chief factors ‘bad’.

**Remarks.** (vi) A non-abelian chief factor belongs to $\mathcal{C}(G)$ if and only if it is either simple or a product of two simple groups. Hence such a factor that occurs inside $G_1$ is a product of at least 3 simple groups.

(vii) $(G^2)^{(3)}G_1/G_1$ is semisimple: for if $M \in \mathcal{C}(G)$ then $G/C_G(M)$ is an extension of $M$ by $\text{Out}(M)$, $\text{Out}(M)$ is isomorphic to $\text{Out}(S)$ or $\text{Out}(S) \wr C_2$ where $S$ is simple, and $\text{Out}(S)^{(3)} = 1$ (Proposition 1.18).

(viii) If $G > 1$ then $G_1 < G$ or $G' < G$.

**Proposition 2.7** Let $G$ be a group and $W = \{w_1, \ldots, w_s\}$ a subset such that $G = D \langle W \rangle$ where $D \leq G_0 \cap G^{(4)}G_1$. Then there exist elements $b_{ij} \in D$ such that

$$G = \langle w_i^{b_{ij}} \mid i = 1, \ldots, s, \ j = 1, \ldots, m \rangle (D \cap G_1)$$

where $m = 1 + 8\delta d(G)$.  

17
Proof. Note that $G_1 = \bigcap_{M \in C \setminus S} C_{G_0}(M)$, where $S$ denotes the set of all simple chief factors of $G$. The section $G^{(4)}G_1/G_1$ is semisimple, and is a product of minimal normal subgroups of $G/G_1$ belonging to $C$. We may suppose that $D \cap G_1 = 1$. In that case, $D$ is a product of minimal normal subgroups of $G$ belonging to $C \setminus S$.

Let $M = S_1 \times S_2$ be one of these. Then $D$ normalizes $S_1$ and $S_2$, so there exists $y \in W$ such that $S_1^y = S_2$ and $S_2^y = S_1$. Now Lemma 2.6 shows that

$$M = [M, y]^{8\delta}.$$ 

As $D$ is the direct product of such normal subgroups $M$ of $G$, it follows that

$$D = [D, y]^{8\delta}.$$ 

If $r < d = d(G)$ put $w_{r+1} = \ldots = w_d = w_r$. Now applying Lemma 2.5 we find elements $h_j \in D$ such that $G = \langle h_1 w_1, \ldots, h_d w_d, w_{d+1}, \ldots, w_r \rangle$. Each $h_j$ lies in the subgroup generated by $W$ and $8\delta D$-conjugates of the $w_i$. The result follows.

2.3 Lifting generators

Recall that a chief factor of $G$ is bad if it is either simple or the product of two simple groups.

Proposition 2.8 Let $G = N \langle y_1, \ldots, y_m \rangle$ be a $d$-generator group where $N$ is a non-central minimal normal subgroup of $G$. If $N$ is non-abelian, assume that $N$ is not bad. Let

$$X = \left\{ a \in N^{(m)} \mid \langle y_1^{a_1}, \ldots, y_m^{a_m} \rangle = G \right\}.$$ 

(i) Suppose that $N$ is abelian and that $y_j$ has the $\varepsilon$-fsp on $N$ for at least $k$ values of $j$. Then

$$|X| \geq |N|^m \left( 1 - |N|^{d-\varepsilon k} \right).$$

(ii) Suppose that $N$ is non-abelian and that $y_j$ has the $\varepsilon$-fsp on $N$ for at least $k$ values of $j$, where $k \varepsilon \geq \max\{2d+4, C\}$ for a certain absolute constant $C$. Then

$$|X| \geq |N|^m \left( 1 - 2^{2-k\varepsilon} \right).$$

Proof. Part (i) is [NS], Proposition 5.1(i). In the situation of (ii), the proof of [NS], Proposition 5.1(ii) shows that $|X| \geq |N|^m (1 - z)$ where $z \leq \zeta(\varepsilon) - 1$ (Riemann zeta function). A crude estimate gives $\zeta(t) - 1 \leq 2^{2-t}$ for $t > 2$.

The main result is now

Theorem 2.9 Let $G$ be a group and $K \leq G_0$ a normal subgroup of $G$. Let $Y = \{y_1, \ldots, y_r\}$ be a subset of $G$ such that $G = G' \langle Y \rangle = K \langle Y \rangle$. Then there exist elements $x_{ij} \in K$ such that

$$G = \langle y_i^{x_{ij}} \mid i = 1, \ldots, r, \ j = 1, \ldots, k \rangle.$$
where

\[ k = \max\{(1 + 2d(G)\hat{r})(1 + 8\delta d(G)), \hat{r}C\} \]

\[ = f_0(r, d(G)) \leq C_0rd(G)^2, \]

\[ \hat{r} = \max\{r, 1 + 6\delta\}, \text{ and } C \text{ and } C_0 \text{ are absolute constants.} \]

**Corollary 2.10** If \( G \) has no simple chief factors and \( G = G' \langle y_1, \ldots, y_r \rangle \) then \( G = \langle y_i^{a_{ij}} \mid i = 1, \ldots, r, j = 1, \ldots, k \rangle \).

**Proof of Theorem 2.9.** Write \( d = d(G) \) and set \( \varepsilon = \hat{r} - 1 \). Let \( N \) be a non-central chief factor of \( G \). We will say that a subset \( W \) of \( G \) has the \((k, \varepsilon)\)-property w.r.t. \( N \) if \( N \) is abelian and at least \( k \) elements of \( W \) have the \( \varepsilon/2 \)-fsp on \( N \), or if \( N \) is a product of at least 3 simple groups and at least \( k \) elements of \( W \) have the \( \varepsilon \)-fpp on the set of simple factors of \( N \). According to Theorem 2.1, the set \( \{y_1, \ldots, y_r\} \) has the \((1, \varepsilon)\)-property w.r.t. \( N \).

Put \( D = K \cap G^{(4)}G_1 \). We begin by proving

\((*)\) there exists elements \( a_{ij} \in K \) such that

\[ G = D \langle y_i^{a_{ij}} \mid i = 1, \ldots, r, j = 1, \ldots, k_1 \rangle \]

where \( k_1 = 1 + 2d\hat{r} \).

Replacing \( G \) by \( G/D \) for the moment, we may assume that \( K \) is soluble. If \( K = 1 \) we can take all \( a_{ij} = 1 \) and there is nothing to prove.

Suppose that \( K > 1 \) and let \( N \) be a minimal normal subgroup of \( G \) contained in \( K \); then \( N \) is abelian. Arguing by induction on \(|K|\), we may suppose that \( G = N \langle W \rangle \) where

\[ W = \{y_i^{a_{ij}} \mid i = 1, \ldots, r, j = 1, \ldots, k_1\} \]

and each \( a_{ij} \in K \). If \( N \leq Z(G) \) then \( G' \leq \langle W \rangle \), so \( \langle W \rangle \geq G' \langle Y \rangle = G \) and we are done.

If \( N \) is non-central, the set \( W \) has the \((k_1, \varepsilon)\)-property w.r.t. \( N \). As \( d - k_1\varepsilon/2 < 0 \), Proposition 2.8(i) shows that there exist elements \( b_{ij} \in N \) such that

\[ G = \langle y_i^{a_{ij}b_{ij}} \mid i = 1, \ldots, r, j = 1, \ldots, k_1 \rangle, \]

and \((*)\) follows on replacing \( a_{ij} \) by \( a_{ij}b_{ij} \). This completes the proof of \((*)\).

Now we apply Proposition 2.7 to find elements \( c_{ijl} \in D \) such that

\[ G = G_1 \langle y_i^{a_{ij}c_{ijl}} \mid i = 1, \ldots, r, j = 1, \ldots, k_1, l = 1, \ldots, m \rangle \]

where \( m = 1 + 8\delta d \).

If \( G_1 = 1 \) we are done. Otherwise, let \( N \) be a minimal normal subgroup of \( G \) contained in \( G_1 \), and suppose inductively that \( G = N \langle W \rangle \) where \( W = \{y_i^{x_{ij}} \mid i = 1, \ldots, r, j = 1, \ldots, k_1m\} \). If \( N \) is abelian we deduce as above that
$G$ is generated by a set of the form \( \{ y_{ij}^{x, b_{ij}} \mid i = 1, \ldots, r, \ j = 1, \ldots, k \} \), and

the result follows since \( k = \max \{ k_1 m, \hat{r} C \} \geq k_1 m \).

Suppose that \( N \) is non-abelian; then \( N \) is not bad. If \( k_1 m < \hat{r} C \), enlarge the family \( W \) by repeating some of its elements to obtain a family containing \( k \) conjugates of each \( y_i \) \( (i = 1, \ldots, r) \). Then in any case, \( W \) has the \( (k, \varepsilon) \)-property w.r.t. \( N \); Proposition 2.8(ii) now shows that \( G \) is generated by a set of the form \( \{ y_{ij}^{x, c_{ij}} \mid i = 1, \ldots, r, \ j = 1, \ldots, k \} \) with \( c_{ij} \in N \), as required. (Note that \( k \varepsilon \geq \max \{ 2d + 4, C \} \) since \( k_1 m > 18d \).)

**Remarks.** (ix) Recall that \( \delta = 1 \) if we accept the validity of the Ore Conjecture (Subsection 1.3.2).

(x) If we assume that \( K \leq G_1 \) we can take \( k = \hat{r} \cdot \max \{ 2d + 4, C \} = O(rd) \).

In particular, if \( G \) has no bad chief factors then the Corollary holds with this smaller value of \( k \).

## 3 Commutators

In this section we begin the proof of the two main ‘commutator’ results.

**Theorem 3.1** Let \( G = \langle g_1, \ldots, g_r \rangle \) be a group and \( H \) a normal subgroup of \( G \). Then

\[
[H, G] = \left( \prod_{i=1}^{r} [H, g_i]^H \right)^{f_3} = \left( \prod_{i=1}^{r} [H, g_i][H, g_i^{-1}] \right)^{f_3}
\]

where \( f_3 = O(rd) = O(r^2) \) depends only on \( r \) and \( d = d(G) \).

**Theorem 3.2** Let \( G = G' \langle g_1, \ldots, g_r \rangle \) be a group and \( H \) a normal subgroup of \( G \) such that \( H \langle g_1, \ldots, g_r \rangle = G \).

(i) If \( H \leq G_0 \) then

\[
[H, G] = \left( \prod_{i=1}^{r} [H, g_i]^H \right)^{f_4} = \left( \prod_{i=1}^{r} [H, g_i][H, g_i^{-1}] \right)^{f_4}
\]

(ii) if \( H \leq G_1 \) then

\[
[H, G] = \left( \prod_{i=1}^{r} [H, g_i]^H \right)^{f_5},
\]

where \( f_4 = O(r^5 d^f) \) and \( f_5 = O(rd) \) depend only on \( r \) and \( d = d(G) \).

These are not quite the same as Theorems 1.2 and 1.3, which refer to a symmetric set \( Y = \{ y_1, \ldots, y_r \} \), and omit the factors \([H, g_i^{-1}]\). To deduce the stated results, note that if \( Y \) is symmetric then

\[
\left( \prod_{j=1}^{r} [H, y_j] \right)^{2r} \supseteq \prod_{j=1}^{r} [H, y_j][H, y_j^{-1}],
\]
and we may take \( f_1 = 2rf_3, f_2 = 2rf_4 \); of course if we are allowed to order \( Y \) so that \( y_{2i} = y_{2i-1}^{-1} \) for \( i = 1, \ldots, r/2 \) then we can take \( f_1 = f_3 \) and \( f_2 = f_4 \).

### 3.1 Acceptable normal subgroups

Suppose that \( A < B \) are normal subgroups of a group \( G \). Recall that \( B/A \) is a **bad chief factor** of \( G \) if \( B/A \) is a minimal normal subgroup of \( G/A \) and \( B/A \) is either simple or the direct product of two simple groups. Thus \( G_1 \) (defined in Subsection 2.1) is precisely the intersection of the centralizers of all bad chief factors of \( G \).

A normal subgroup \( H \) of \( G \) is said to be **acceptable** in \( G \) if

(a) \( H = [H, G] \) and

(b) if \( A < B \leq H \) are normal subgroups of \( G \) then \( B/A \) is not a bad chief factor of \( G \).

Here we show how the main results may be reduced to the consideration of acceptable normal subgroups.

**Lemma 3.3** \( H \triangleleft G \) is acceptable if and only if \( H = [H, G] \leq G' \cap G_1 \).

**Proof.** If \( H \geq B > A \) and \( B/A \) is a bad chief factor then \( H \) does not centralize \( B/A \), so \( H \nleq G_1 \). Conversely, if \( H \nleq G_1 \) then \( H \) does not centralize some bad chief factor \( B/A \); then \( (B \cap H)A = A \) so \( (B \cap H)/(A \cap H) \cong B/A \) and \( A \cap H < B \cap H \leq H \) contradicts (b), showing that \( H \) is not acceptable. ■

The next result is elementary; it is the general form of facts (A) and (B) mentioned in the introduction:

**Lemma 3.4** Let \( H \triangleleft G = G' \langle g_1, \ldots, g_r \rangle \) and let \( n \geq 1 \). Then

\[
[H, G] = [H, nG] \prod_{i=1}^{r} [H, g_i].
\]

If in addition we have \( G = H \langle g_1, \ldots, g_r \rangle \) then

\[
G = [H, nG] \langle g_1, \ldots, g_r \rangle.
\]

**Proof.** The first claim is [NS], Lemma 2.4 or [S2], Prop. 1.2.5. For the second, we argue by induction on \( n \) and reduce to the case where \( [H, G] = 1 \). Then \( G' \leq \langle g_1, \ldots, g_r \rangle \) and the claim is evident. ■

**Lemma 3.5** Let \( G \) be a group and \( \alpha, \beta \in \text{Aut}(G) \). Then

\[
[G, \alpha] \subseteq [G, \alpha][G, \alpha^{-1}],
\]

\[
[G, \alpha \beta] \subseteq [G, \beta][G, \beta^{-1}][G, \alpha][G, \alpha^{-1}][G, \beta][G, \beta^{-1}][G, \alpha][G, \alpha^{-1}]
\]

\[
[G, \alpha^{-1} \beta \alpha] \subseteq [G, \beta][G, \beta^{-1}][G, \alpha][G, \alpha^{-1}]
\]
Proof.

\[[x, \alpha]^k = [xy, \alpha][y^\alpha, \alpha]^{-1},\]
\[[x, \alpha \beta] = [x, \beta][x, \alpha][x, \alpha][\beta]\]
\[[x, \alpha^{-1} \beta \alpha] = [x^{\alpha^{-1}}, \beta][x^{\alpha^{-1}}, \beta, \alpha].\]

\hfill ■

Lemma 3.6 Let \(G\) be a quasisimple group and \(\alpha \in \text{Aut}(G)\). Put \(\overline{G} = G/Z(G)\).

If \(|[\overline{G}, \alpha]|^s \geq |\overline{G}|\) then

\[G = ([G, \alpha]^G)^{cs}\]

where \(c \in \mathbb{N}\) is an absolute constant.

Proof. Proposition 1.23 shows that if \(Y\) is a normal subset of \(\overline{G}\) with \(|Y|^s \geq |\overline{G}|\) then \(\overline{G} = Y^{s c' s}\), where \(c'\) is an absolute constant. Applying this with \(Y = XZ(G)/Z(G)\) where \(X = [G, \alpha]^G\) we get

\[G = X^{s c' s}Z(G)\]

Now for \(g, h, k \in G\) we have

\[[[g, \alpha]^k, h] = [g, \alpha]^{-k}[g, \alpha]^{kgh} = [g^{-1}, \alpha]^{gk}[g, \alpha]^{kh} \in X^{s 2},\]

so if \(w \in X^{s c' s}\) then

\[[w, h] \in X^{* 2 c' s}.\]

According to Proposition 1.19 there exists an absolute constant \(\delta^*\) such that every element of \(G\) is a product of \(\delta^*\) commutators (In fact \(\delta^* = 2\)). It follows that

\[G = X^{* 2 c' s \delta^*}.\]

\hfill ■

Lemma 3.7 Let \(G = \langle g_1, \ldots, g_r \rangle\) and suppose that \(T \trianglelefteq G\) is quasisemisimple with one or two simple composition factors. Then

\[T = \left( \prod_{i=1}^r [T, g_i]^T \right)^{* k_0 r},\]

where \(k_0\) is an absolute constant.

Proof. Suppose that \(T\) is quasisimple, with centre \(Z\). Put \(\overline{T} = T/Z\). Then \(C_\overline{T}(G) = 1\) so \(|\overline{T}| \leq \prod_{i=1}^r [\overline{T}, g_i]| \) and so \(|[\overline{T}, g_i]| \geq |\overline{T}|^{1/c}\) for some \(i\). Now Lemma 3.6 implies that \(T = ([T, g_i]^T)^{* c r}\).

If \(T\) is not quasisimple, then \(T = S_1S_2\) with each \(S_i\) quasisimple and \([S_1, S_2] = 1\). If \(G\) normalizes the factors \(S_i\), we apply the preceding paragraph to each factor and obtain the same result as before. Otherwise, \(G\) permutes them
transitively by conjugation. The action of $G$ lifts to an action on the universal cover $\tilde{T} = \tilde{S}_1 \times \tilde{S}_2$, and for some $i$ we have $\tilde{S}_1^{g_1} = \tilde{S}_2$, $\tilde{S}_2^{g_2} = \tilde{S}_1$. Let $C_j$ denote the set of commutators in $\tilde{S}_j$; then Lemma 2.6 shows that

$$C_j \subseteq [\tilde{T}, g_i]^{*4}$$

for $j = 1, 2$. Since $\tilde{S}_j = C_j^{*\delta^*}$ (Proposition 1.19), it follows that

$$\tilde{T} = [\tilde{T}, g_i]^{*4\delta^*},$$

which implies $T = [T, g_i]^{*4\delta^*}$.

The result follows on setting $k_0 = \max\{c, 4\delta^*\}$.

Let us say that $N \triangleleft G$ is narrow if

$$\bigcap_{T \in \mathcal{M}} T \leq Z(N)$$

where $\mathcal{M}$ is the set of normal subgroups $T$ of $G$ contained in $N$ such that $N/T$ is semisimple with composition length at most two. This is equivalent to saying that $N/Z(N)$ is a direct product of bad chief factors of $G$ (occurring as minimal normal subgroups of $G/Z(N)$).

**Lemma 3.8** Let $G = \langle g_1, \ldots, g_r \rangle$ and let $N$ be a perfect narrow normal subgroup of $G$. Then

$$N = \left( \prod_{i=1}^r [N, g_i]^N \right)^{k_0 r}$$

where $k_0$ is given in Lemma 3.7.

**Proof.** The hypotheses imply that $N$ is a central product $N = T_1 \cdots T_n$ where each $T_i$ is a quasisemisimple normal subgroup of $G$ having one or two simple composition factors. As the $T_i$ commute elementwise the claim follows from Lemma 3.7.

**Proposition 3.9** Let $G = \langle g_1, \ldots, g_r \rangle$ and let $H \triangleleft G$. Then $G$ has normal subgroups $H_3 \leq H_2 \leq H_1 \leq [H, G]$ such that

$$[H, G] = \prod_{i=1}^r [H, g_i] \cdot H_1, \quad (7)$$

$$H_2 = \left( \prod_{i=1}^r [H_2, g_i]^H \right)^{k_0 r} \cdot H_3, \quad (8)$$

$H_1/H_2$ is acceptable in $G/H_2$ and $H_3$ is acceptable in $G$. 

23
Proof. Let $G_s$ be the soluble residual of $G$ and set $H_1 = [H, G_s]$, $H_2 = [H, G_s]$. Let $D$ be the intersection of all $M < G$ such that $M < H_2$ and $H_2/M$ is either simple or a product of two simple groups, and put $H_3 = [D, G_s]$.

Then (8) follows from Lemma 3.4. Also $H_1 = [H, G_s]$, $H_2 = [H, G_s]$, and $H_1/H_2$ is soluble, so $H_1/H_2$ is acceptable in $G/H_2$.

Now

$$[H_3, G] ≥ [H_3, G_s] = [D, G_s, G] = [D, G_s] = H_3$$

since $G_s$ is perfect. To complete the proof that $H_3$ is acceptable, suppose that $K ≤ H_3$ is a minimal normal subgroup of $G$ and that $K$ is either simple or a product of two simple groups. Then $G/KCG(K)$ is soluble by the Schreier conjecture (Proposition 1.18), so $G_s ≤ KCG(K)$ and as $K ≤ H_2 ≤ G_s$ it follows that $H_2 = K × CH_2(K)$. This implies that $CH_2(K) ≥ D ≥ K$, a contradiction. Applying this argument to an arbitrary quotient of $G$ we infer that $H_3$ is acceptable in $G$.

Finally, $H_2/H_3$ is narrow in $G/H_3$ so Lemma 3.8 gives (8).

3.2 The ‘Key Theorem’

The ‘Key Theorem’ of [NS] described certain product decompositions of an acceptable normal subgroup in a $d$-generator group. As one of us wrote in [S2], 'each part has an undesirable feature in either its hypothesis or its conclusion'. These are now swept away in our core technical result. To state this we need some notation:

Definition For $g, v ∈ G(m)$ and $1 ≤ j ≤ m$,

$$τ_j(g, v) = v_j[g_{j-1}, v_{j-1}]⋯[g_1, v_1].$$

Theorem 3.10 There exists a function $k : \mathbb{N}^{(2)} → \mathbb{N}$ with the following property. Let $G$ be a $d$-generator group and $H$ an acceptable normal subgroup of $G$. Suppose that $G = H \langle g_1, \ldots, g_r \rangle$. Put $m = r · k(d, r)$, and for $1 ≤ j < k(d, r)$ and $1 ≤ i ≤ r$ set

$$g_{i+jr} = g_i.$$

Then for each $h ∈ H$ there exist $v(i) ∈ H(m)$ ($i = 1, \ldots, 10$) such that

$$h = \prod_{i=1}^{10} \prod_{j=1}^{m} [v(i)_j, g_j]$$

and

$$\langle g_1^{τ_1(g, v(i))}, \ldots, g_m^{τ_m(g, v(i))} \rangle = G$$

for $i = 1, \ldots, 10$. (10)

In fact we can take

$$k(d, r) = 1 + \max\{r, 1 + 6δ\} · \max\{4d + 4, \tilde{C}\} ≤ C_1dr,$$
where \( \hat{C} \) and \( C_1 \) are absolute constants.

The proof will occupy the next three subsections. Accepting the theorem for now, we deduce the main results stated above.

**Proof of Theorem 3.1.** We are given \( H \trianglelefteq G = \langle g_1, \ldots, g_r \rangle \). Let \( H_3 \leq H \leq H_1 \leq [H, G] \) be the normal subgroups given by Proposition 3.9. Thus \( H_1/H_2 \) is acceptable in \( G/H_2 \) and \( H_3 \) is acceptable in \( G \). Theorem 3.10 shows that

\[
H_1 = \left( \prod_{i=1}^{r} [H_1, g_i] \right)^{\ast 10k(d, r)} \cdot H_2
\]

and that

\[
H_3 = \left( \prod_{i=1}^{r} [H_3, g_i] \right)^{\ast 10k(d, r)}
\]

where \( d = d(G) \). Combining these with (7) and (8) from Proposition 3.9 we deduce that

\[
[H, G] = \left( \prod_{i=1}^{r} [H, g_i]^H \right)^{\ast f_3}
\]

where \( f_3 = 1 + k_0r + 20k(d, r) \); here \( k_0 \) is the absolute constant introduced in Lemma 3.7. Finally, Lemma 3.5 shows that \([H, g_j]^H\) can be replaced by \([H, g_j][H, g_j^{-1}]\) for each \( j \).

We observe that \( f_3 = O(r + k(d, r)) = O(dr) = O(r^2) \).

**Proof of Theorem 3.2 (i).** Now \( H \trianglelefteq G \) satisfies \( H \leq G_0 \), and \( G = G/\langle g_1, \ldots, g_r \rangle = H \langle g_1, \ldots, g_r \rangle \). According to Theorem 2.9 there exist element \( x_{ij} \in H \) such that

\[
G = \langle g_i^{x_{ij}} \mid i = 1, \ldots, r, \; j = 1, \ldots, k \rangle
\]

where \( k = f_0(r, d(G)) \). Using this generating set in Theorem 3.1 gives

\[
[H, G] = \left( \prod_{i=1}^{k} \prod_{j=1}^{r} [H, g_i^{x_{ij}}]^H \right)^{\ast f_3(kr)}
\]

\[
= \left( \prod_{i=1}^{k} \prod_{j=1}^{r} [H, g_i]^H \right)^{\ast f_3(kr)} = \left( \prod_{i=1}^{r} [H, g_i]^H \right)^{\ast f_4}
\]

where \( f_4 = kf_3(kr) \). Again, we may replace \([H, g_j]^H\) by \([H, g_j][H, g_j^{-1}]\), by Lemma 3.5.

Since \( k = f_0(r, d) \leq C_0rd^2 \) where \( d = d(G) \), we have \( f_4 = O(k^3r^2) = O(r^5d^5) \).

25
We remark that this bound for $f_4(r,d)$ is very crude; a much better bound emerges if, instead of quoting Theorem 2.9, one uses the method of proof of that theorem to reduce Theorem 3.2 (i) to Theorem 3.2 (ii).

**Proof of Theorem 3.2 (ii).** Now we assume that $H$, as above, satisfies $H \leq G_1$. Put $H_1 = [H, \omega G]$. Then $H_1$ is acceptable in $G$, by Lemma 3.3, and $G = H_1 \langle g_1, \ldots, g_r \rangle$ by Lemma 3.4 Thus Theorem 3.10 and Lemma 3.4 together yield

$$[H,G] = \prod_{j=1}^{r} [H,g_j] \cdot H_1$$

$$= \prod_{j=1}^{r} [H,g_j] \cdot \left( \prod_{j=1}^{r} [H_1,g_j] \right)^{*10k(d,r)} = \left( \prod_{j=1}^{r} [H,g_i] \right)^{*f_5}$$

where $f_5 = 1 + 10k(d,r) = O(rd)$.

### 3.3 Proof of the Key Theorem: reductions

We follow the strategy of [NS], Section 4.

**Notation** For $u, g \in G^{(m)}$,

$$u \cdot g = (u_1 g_1, \ldots, u_m g_m), \quad c(u, g) = \prod_{j=1}^{m} [u_j, g_j].$$

**Lemma 3.11**

$$\left( \prod_{i=1}^{s} c(a(i) \cdot u(i), g) \right) \left( \prod_{i=1}^{s} c(u(i), g) \right)^{-1} = \prod_{i=1}^{s} \left( \prod_{j=1}^{m} [a(i)_j, g_j]^{\tau_j(g,u(i))} \right)^{w(i)}$$

where $w(i) = c(u(i-1), g)^{-1} \cdots c(u(1), g)^{-1}$.

This is a direct calculation. The next lemma is easily verified by induction on $m$ (see [NS], Lemma 4.5):

**Lemma 3.12**

$$\left\langle g_j^{\tau_j(g,u)} \mid j = 1, \ldots, m \right\rangle = \left\langle g_j^{u_h} \mid j = 1, \ldots, m \right\rangle$$

where $h_j = g_{j-1}^{-1} \cdots g_1^{-1}$.
Now let $H < G = H \langle g_1, \ldots, g_r \rangle$ be as in Theorem 3.10. If $H = 1$ there is nothing to prove, so we suppose that $H > 1$ and argue by induction on $|H|$. Since $H$ is acceptable, we have $H = [H, G]$. Choose $N < G$ with $N \leq H$ minimal subject to $1 < N = [N, G]$ (in $N$ such an $N$ was called a quasi-minimal normal subgroup of $G$). Let $Z$ be a normal subgroup of $G$ maximal subject to $Z < N = [N, G]$. Then $[Z, G] = 1$ for some $n$, which implies (i) that $Z = N \cap \zeta_n(G)$ is uniquely determined, and (ii) that $[Z, N] \leq [Z, H] \leq [Z, G] = 1$. By definition, $\overline{N} = N/Z$ is a chief factor of $G$; it is not bad because $H$ is acceptable.

Applying Lemma 3.4 to $Z$ we note that $Z$ is contained in the Frattini subgroup $\Phi(G)$ of $G$.

We fix a natural number $k$, the candidate for $k(d, r)$, and define $g_j$ for $j = 1, \ldots, kr$ as in Theorem 3.10. Depending on the nature of $N$, we shall choose a certain normal subgroup $K$ of $G$ with $1 \neq K \leq N$.

Suppose now that $h \in H$. We have to find elements $v(i) \in H^{(m)}$ (i = 1, \ldots, 10) such that (9) and (10) hold. By inductive hypothesis, we can do this ‘modulo $K$’: thus there exist $u(i) \in H^{(m)}$ and $\kappa \in K$ such that

$$h = \kappa \prod_{i=1}^{10} c(u(i), g)$$

and

$$G = K \left\langle g_j^{\tau_j(g, u(i))} \mid j = 1, \ldots, m \right\rangle = K \left\langle g_j^{u(i), h_j} \mid j = 1, \ldots, m \right\rangle \quad \text{for } i = 1, \ldots, 10,$$

the second equality thanks to Lemma 3.12.

The idea now is to find elements $a(i) \in N^{(m)}$ such that (9) and (10) are satisfied on setting $v(i) = a(i) \cdot u(i)$.

Lemma 3.11 shows that (9) is then equivalent to

$$\prod_{i=1}^{10} \left( \prod_{j=1}^{m} [a(i)_j, g_j]^{\tau_j(g, u(i))} \right)^{w(i)} = \kappa.$$  \hspace{1cm} (12)

This can be further simplified by setting

$$y(i)_j = g_j^{\tau_j(g, u(i)) w(i)} \quad \text{and} \quad t(i)_j = g_j^{u(i)_j h_j},$$

$$b(i)_j = a(i)^{\tau_j(g, u(i))} w(i), \quad c(i)_j = a(i)^{u(i)_j h_j}.$$  \hspace{1cm} (13)

Define $\phi(i) : N^{(m)} \to N$ by

$$b \phi(i) = c(b, y(i)).$$

27
Then (12) becomes
\[ \prod_{i=1}^{10} b(i)\phi(i) = \kappa, \]
and (11) is equivalent to
\[ G = K \langle y(i)_1, \ldots, y(i)_m \rangle \]
\[ = K \langle t(i)_1, \ldots, t(i)_m \rangle \quad \text{for } i = 1, \ldots, 10. \quad (15) \]

Similarly, (10) holds if and only if for \( i = 1, \ldots, 10 \) we have
\[ G = \langle t(i)_j^{c(i)_j} \mid j = 1, \ldots, m \rangle Z \quad (16) \]
(where \( Z \) is added harmlessly since \( Z \leq \Phi(G) \)). Let \( X(i) \) denote the set of all \( c(i)_j \in N(m) \) such that (16) holds, and write \( W(i) \) for the image of \( X(i) \) under the bijection \( N(m) \to N(m) \) defined in (13) sending \( c(i) \mapsto b(i) \).

To sum up: to establish the existence of \( a(1), \ldots, a(10) \in N(m) \) such that the \( v(i) = a(i) \cdot u(i) \) satisfy (9) and (10), it suffices to find \((b(1), \ldots, b(10)) \in W(1) \times \cdots \times W(10) \) such that (14) holds.

Set \( \varepsilon = \min\{1, 1+\frac{6}{1+6\delta} \} \), and write \( - : G \to G/Z \) for the quotient map. Now we separate four cases.

3.3.1 The easy case
If \([Z, G] > 1 \) we define \( K = [Z, G] \). Since \([Z, H] = 1 \) and \( G = H \langle g_1, \ldots, g_r \rangle \), we have \( K = \prod_{j=1}^{r} [Z, g_j] \). Thus \( \kappa = \prod_{j=1}^{r} [z_j, g_j] \) with \( z_1, \ldots, z_r \in Z \). In this case, (14) is satisfied if we set
\[ b(1)_j = z_j \quad (1 \leq j \leq r) \]
\[ b(1)_j = 1 \quad (r < j \leq m) \]
\[ b(i)_j = 1 \quad (2 \leq i \leq 10, \ 1 \leq j \leq m), \]
because \( y(i)_j \) is conjugate to \( g_j \) under the action of \( H \) and \([Z, H] = 1 \).

For each \( i \) we have \( W(i) \supseteq Z(m) \), since in this case (15) implies (16) if \( c(i)_j \in Z \) for all \( j \). So \( b(i)_j \in W(i) \) for each \( i \), as required.

3.3.2 The abelian case
If \([Z, G] = 1 \) and \( N \) is abelian we set \( K = N \). We use additive notation for \( N \) and consider it as a \( G \)-module. Then (15) implies that
\[ \phi(1) : b \mapsto \sum_{j=1}^{m} b_j(y(1)_j - 1) \]
is a surjective \((Z\text{-module})\) homomorphism \( N(m) \to N \). It follows that
\[ |\phi(1)^{-1}(c)| = |\ker \phi(1)| = |N|^{m-1} \]
for each $c \in N$.

Now fix $i \in \{1, \ldots, 10\}$. According to Theorem 2.1, at least one of the elements $g_j$ has the $\varepsilon/2$-fsp on $N$; therefore at least $k$ of the elements $f(i)_j$ have this property. Now we apply Proposition 3.13: this shows that (10) holds for at least $|N|^m (1 - |N|^{-d/k\varepsilon/2})$ values of $c(i)$ in $|N|^m$. It follows that

$$|W(i)| = |X(i)| \geq |Z|^m \cdot |N|^m (1 - |N|^{-d/k\varepsilon/2}) = |N|^m (1 - |N|^{-d/k\varepsilon/2}). \quad (17)$$

We need to compare $|N|$ with $|N|$. Let $\{x_1, \ldots, x_d\}$ be a generating set for $G$. Then $b \mapsto \sum_{j=1}^d b_j(x_j - 1)$ induces an epimorphism from $N^{(d)}$ onto $N$; consequently $|N| \leq |N|^d$. Thus provided $k\varepsilon/2d > 1$ we have

$$|W(i)| \geq |N|^m (1 - |N|^{1-k\varepsilon/2d}).$$

Assume now that $k\varepsilon > 4d$. Then $W(i)$ is non-empty for each $i$. For $i = 2, \ldots, 10$ choose $b(i) \in W(i)$ and put

$$c = \kappa \left( \prod_{i=2}^{10} b(i)\phi(i) \right)^{-1}.$$ 

Then

$$|\phi(1)^{-1}(c)| + |W(1)| \geq |N|^m \cdot |N|^{-1} + 1 - |N|^{1-k\varepsilon/2d}) \geq |N|^m.$$ 

It follows that $\phi(1)^{-1}(c) \cap W(1)$ is non-empty. Thus we may choose $b(1) \in \phi(1)^{-1}(c) \cap W(1)$ and ensure that (14) is satisfied.

### 3.3.3 The soluble case

Suppose next that $[Z, G] = 1$ and $N > N' > 1$. In this case we take $K = N'$. Since $N' \leq Z$, the argument above again gives (17).

The maps $\phi(i)$ are no longer homomorphisms, however, and it is quite a major undertaking to obtain a good estimate for the fibres. The outcome is Proposition 7.1 of [NS]; translated into the present notation it is

**Proposition 3.13** Assume that $G = Z < (y(i)_{1, \ldots, y(i)_{m}})$ for $i = 1, 2, 3$. Then for each $c \in N'$ there exist $c_1, c_2, c_3 \in N$ such that $c = c_1c_2c_3$ and

$$|\phi(i)^{-1}(c_i)| \geq |N|^m \cdot |N|^{-d-1} \quad (i = 1, 2, 3). \quad (18)$$

The initial hypothesis follows from (15) since now $K \leq Z$.

Assume now that $k\varepsilon > 4d + 2$. Then (17) and (18) together imply that $\phi(i)^{-1}(c_i) \cap W(i)$ is non-empty for $i = 1, 2, 3$, while (17) implies that $W(i)$ is non-empty for every $i$. 

29
Choose \( b(i) \in W(i) \) for \( i = 4, \ldots, 10 \). Put
\[
c = \kappa \left( \prod_{i=4}^{10} b(i) \phi(i) \right)^{-1},
\]
and choose \( c_1, c_2, c_3 \) as in Proposition 3.13. Then for \( i = 1, 2, 3 \) we can find \( b(i) \in \phi(i)^{-1}(c_i) \cap W(i) \), and so ensure that (14) is satisfied.

3.3.4 The semisimple case

If \( [Z, G] = 1 \) and \( N = N' \), define \( K = N \). Now \( N' \) is semisimple with at least 3 simple factors, and \( N \) is quasi-semisimple. In this case, Theorem 2.11 shows that at least one the elements \( g_j \) has the \( \varepsilon \)-fgp on \( N \); therefore for each \( i \), at least \( k \) of the elements \( t(i) \) and at least \( k \) of the elements \( y(i) \) have this property.

Proposition 2.8(ii) now shows that
\[
|X(i)| \geq |N|^m \left( 1 - 2^{2-k\varepsilon} \right),
\]
provided we assume that \( k\varepsilon \geq \max \{ 2d + 4, C \} \) for a certain absolute constant \( C \). This implies
\[
|W(i)| = |X(i)| \geq |N|^m \left( 1 - 2^{2-k\varepsilon} \right).
\]

Now Theorem 4.28 proved below in Subsection 4.2 gives the following: there are absolute constants \( D, \varepsilon_0 \) such that if for each \( i = 1, \ldots, 10 \)

\( a \) the group \( \langle y(i), \ldots, y(i)_m \rangle \) permutes the quasisimple factors of \( N \) transitively,

\( b \) at least \( k \) of the \( y(i)_j \) have the \( \varepsilon \)-fgp on \( N \), where \( k\varepsilon \geq 4 + 2D \),

\( c \) the subset \( W(i) \subseteq N^{(m)} \) satisfies \( |W(i)| \geq (1 - \varepsilon_0/6) |N|^m \),

then
\[
\prod_{i=1}^{10} W(i) \phi(i) = N.
\]

Condition (a) follows from (15). Thus we can find \( b(i) \in W(i) \) \( (i = 1, \ldots, 10) \) such that (14) is satisfied provided we assume that
\[
k\varepsilon > \max \{ 2d + 4, C, 4 + 2D, 2 + \log_2(6/\varepsilon_0) \}
\]
\[
= \max \{ 2d + 4, C^* \}
\]
where \( C^* \) is an absolute constant.

3.3.5 Conclusion of the proof

Recall that we defined \( \varepsilon = \min \{ \frac{1}{1+6\delta}, \frac{1}{r} \} \). So if we now define
\[
k(d, r) = 1 + \max \{ r, 1 + 6\delta \} \cdot \max \{ 4d + 4, [C^*] \},
\]
then \( k = k(d, r) \) fulfils the requirements of all the preceding steps. This concludes the proof of Theorem 3.10 modulo Proposition 2.8, Theorem 4.28 and [NS], Proposition 6.2.

30
4 Semisimple groups

This section is devoted to the proof of Theorem 4.28. This will be stated in Subsection 4.2. Like Proposition 9.2 of [NS], which it in effect generalizes, its proof has two components: (1) a result about products of commutators in quasisimple groups, and (2) a complicated combinatorial reduction argument. These will occupy the next two subsections.

As remarked in the Introduction, the proof of (1) given here is significantly simpler (and shorter) than [NS2], which played the analogous role in our earlier work. The reduction argument (2) is essentially the same as in [NS], though we are now using it to prove something different (specifically, we have to control the image of a certain mapping rather than its fibres). We have re-cast the argument from scratch, in an attempt to make it more transparent (the reader will judge whether we have succeeded!) However, we shall quote one combinatorial result from Section 8 of [NS].

4.1 Twisted commutators in quasisimple groups

For automorphisms $\alpha, \beta$ of a group $S$ and $x, y \in S$ we write

$$T_{\alpha, \beta}(x, y) = x^{-1} y^{-1} x^{\alpha} y^{\beta}.$$ 

For $\alpha = (\alpha_1, \ldots, \alpha_D)$ and $\beta = (\beta_1, \ldots, \beta_D)$ in $\text{Aut}(S)^{(D)}$ the mapping $T_{\alpha, \beta} : S^{(D)} \times S^{(D)} \to S$ is defined by

$$T_{\alpha, \beta}(x, y) = \prod_{i=1}^{D} T_{\alpha_i, \beta_i}(x_i, y_i).$$

**Theorem 4.1** There exist $\varepsilon > 0$ and $D \in \mathbb{N}$ such that if $S$ is a finite quasisimple group, $\alpha, \beta \in \text{Aut}(S)^{(D)}$, and $X \subseteq S^{(2D)}$ has size at least $(1 - \varepsilon) \left| S^{(2D)} \right|$, then $|T_{\alpha, \beta}(X)| \geq \lambda |S|$, where

$$\lambda = \begin{cases} 
  l(S)^{-3/5} & \text{if } l(S) \geq 3 \\
  1 & \text{if } l(S) = 2 
\end{cases}.$$  \hspace{1cm} (19)

The following corollary is Theorem 1.1 of [NS2].

**Corollary 4.2** There exists $D_1 \in \mathbb{N}$ such that if $S$ is a finite quasisimple group and $\alpha, \beta \in \text{Aut}(S)^{(D_1)}$ then

$$\prod_{i=1}^{D_1} T_{\alpha_i, \beta_i}(S, S) = S.$$ 

**Proof.** Set $D_1 = 5D$, and divide $\alpha$ and $\beta$ into 5 $D$-tuples $\alpha(j)$, $\beta(j)$. Taking $X = S^{(2D)}$ in the theorem gives $|T_{\alpha(j), \beta(j)}(X)| \geq l(S)^{-3/5} |S|$ for $j = 1, \ldots, 5$. The result now follows by the ‘Gowers trick’, since $5 \times \frac{1}{5} = 5 - 2$. \hfill \blacksquare
4.1.1 Reductions for Theorem 4.1

In this subsection, we fix a finite group $S$ and $\lambda \in (0,1]$. For any $\alpha, \beta \in \text{Aut}(S)^{(D)}$ we consider the statement

$$\mathcal{P}(\alpha, \beta; D, \varepsilon): \text{For } X \subseteq S^{(2D)},$$

$$|X| \geq (1-\varepsilon)|S^{(2D)}| \implies |T_{\alpha,\beta}(X)| \geq \lambda|S|.$$  

If $\Gamma$ is a subgroup of $\text{Aut}(S)$, we write

$$\mathcal{P}(\Gamma; D, \varepsilon) \Leftrightarrow \mathcal{P}(\alpha, \beta; D, \varepsilon) \forall \alpha, \beta \in \Gamma^{(D)}.$$  

Thus Theorem 4.1 asserts the existence of $D$ and $\varepsilon$ such that $\mathcal{P}(\text{Aut}(S); D, \varepsilon)$ holds with $\lambda$ defined by (19) for every quasisimple group $S$.

Our aim in the rest of this subsection is to establish the reduction steps Propositions 4.3, 4.4 and 4.12.

**Proposition 4.3** If $D_1 \leq D$ and $\varepsilon_1 \geq \varepsilon$ then $\mathcal{P}(\Gamma; D_1, \varepsilon_1)$ implies $\mathcal{P}(\Gamma; D, \varepsilon)$.

**Proof.** If $D_1 = D$ the claim is obvious. Suppose that $D > D_1$. We write

$$T_{\alpha,\beta}(x, y) = T_{\alpha',\beta'}(x', y')T_{\alpha'',\beta''}(x'', y''),$$

where $x' = (x_1, \ldots, x_{D_1})$, $x'' = (x_{D_1+1}, \ldots, x_D)$ etc. Now if $X \subseteq S^{(2D)}$ satisfies $|X| \geq (1-\varepsilon)|S^{(2D)}|$ then there exist $(x', y') \in S^{2(D-D_1)}$ and $X_1 \subseteq S^{(2D_1)}$ such that $X_1 \times \{(x'', y'')\} \subseteq X$ and $|X_1| \geq (1-\varepsilon)|S^{(2D_1)}|$. Then

$$T_{\alpha,\beta}(X) \supseteq T_{\alpha',\beta'}(X_1) \cdot T_{\alpha'',\beta''}(x'', y''),$$

a set of size at least $\lambda |S|$ since $1-\varepsilon \geq 1 - \varepsilon_1$. 

**Proposition 4.4** If $\Delta \triangleleft \Gamma$ and $|\Gamma : \Delta| \leq n$ then $\mathcal{P}(\Delta; D, \varepsilon)$ implies $\mathcal{P}(\Gamma; n^2D, \varepsilon)$.

This is a little more complicated. It will follow from

**Proposition 4.5** Let $\Delta$ be a normal subgroup of index $n$ in $\Gamma$, and let $\alpha, \beta \in \Gamma^{(n^2D)}$. Then there exist $\alpha', \beta' \in \Delta^{(D)}$ and a bijection $\pi : S^{(n^2D)} \to S^{((2n^2-2)D)} \times S^{(2D)}$ such that, for each $x \in S^{(n^2D)}$,

$$\prod_{i=1}^{n^2D} T_{\alpha,\beta}(x_{2i-1}, x_{2i}) = \prod_{i=1}^{D} T_{\alpha',\beta'}(\hat{x}_{2i-1}, \hat{x}_{2i}) \cdot R(\hat{x}),$$

where $(\hat{x}, \hat{\hat{x}}) = x\pi$ and $R(\hat{x})$ depends only on $\hat{x}$.

Accepting this for now we deduce Proposition 4.4. Let $\varepsilon \in (0,1)$ and suppose that $W \subseteq S^{(2n^2D)}$ satisfies $|W| \geq (1-\varepsilon)|S^{(2n^2D)}|$. Then $|W\pi| = |W|$; so for at least one element $u \in S^{((2n^2-2)D)}$ the set

$$Y_u := \left\{ y \in S^{(2D)} \mid (u, y) \in W\pi \right\}$$

32
satisfies $|Y_u| \geq (1 - \varepsilon) |S^{(2D)}|$. Then
\[
T_{\alpha, \beta}(W) \supseteq T_{\alpha, \beta}(Y_u) \cdot R(u).
\]
If $\mathcal{P}(\Delta; D, \varepsilon)$ holds then $|T_{\alpha, \beta}(Y_u)| \geq \lambda |S|$, and so $|T_{\alpha, \beta}(W)| \geq \lambda |S|$. Thus $\mathcal{P}(\Gamma; n^2D, \varepsilon)$ holds as claimed.

Now we embark on the proof of Proposition 4.5.

**Lemma 4.6**
\[
T_{a_1, b_1}(x_1, x_2) T_{a_2, b_2}(x_3, x_4) = A(x \mu) \cdot T_{a_1, b_1}(x \mu) \cdot B(x \nu) = C(x \tau) \cdot T_{a_1, b_1}(x \sigma)
\]
where $x \mapsto (x \mu, x \nu)$ and $x \mapsto (x \sigma, x \tau)$ are bijections from $S^4$ to $S^2 \times S^2$.

**Proof.** Take $z = x_3(x_1^{\alpha_1}, x_2^{\beta_1})^{-1}$, $w = z^{-\alpha_2}x_4z x_2$ and $t = x_4^{\beta_1\alpha_1^{-1}}z^{\alpha_2\beta_2\alpha_1^{-1}}$ and set
\[
x \mu = (t, w), \quad x \nu = (x_2, z),
\]
\[
A(y, z) = (y^{\beta_1}z^{\beta_2\alpha_1^{-1}})^{\alpha_1^{-1}}, \quad B(y, z) = (yz)^{-\beta_2}.
\]
Take $u = x_2^{\beta_1} x_3^{\alpha_1} x_4^{\alpha_1^{-1}} = x_1 u^{\alpha_1} x_3^{\alpha_2\alpha_1^{-1}} u^{-\beta_2\alpha_1^{-1}}$ and set
\[
x \sigma = (v, x_2), \quad x \tau = (u, x_3),
\]
\[
C(u, y) = (y^{\alpha_2}u^{-\beta_2})^{\alpha_1^{-1}}.
\]

**Lemma 4.7**
\[
z T_{a, b}(x, y) = T_{a, b}(x', y') z^{-\gamma}
\]
where $x' = z^{\alpha\beta^{-1}a^{-1}} x z^{-1}, y' = z^{\alpha\beta^{-1}a^{-1}} y z^{-\alpha\beta^{-1}a^{-1}}$ and $\gamma = [\alpha^{-1}, \beta]$.

**Lemma 4.8** Suppose that $|\Gamma : \Delta| = 2$. Given $\alpha_i, \beta_i \in \Gamma (i = 1, \ldots, 4)$, there exist $\gamma, \delta \in \Delta$, a bijection $x \mapsto (x^*, \bar{x})$ from $S^4$ to $S^2 \times S^6$ and maps $P, Q : S^6 \to S$ such that
\[
\prod_{i=1}^{4} T_{a_i, b_i}(x_{2i-1}, x_{2i}) = P(\bar{x}) T_{a, b}(x^*) Q(\bar{x}).
\]

**Proof.** Define
\[
(x_1, x_2, x_3, x_4; \alpha_1, \beta_1) \quad \text{if} \quad \alpha_1 \in \Delta
\]
\[
(x_3, x_4, x_1, x_2; \alpha_2, \beta_2) \quad \text{if} \quad \alpha_2 \in \Delta
\]
\[
(x_\mu, x_\nu; \alpha_1, \alpha_2, \beta_2) \quad \text{if} \quad \alpha_1 \alpha_2 \in \Delta
\]
(assuming in the 2nd and 3d lines that $\alpha_1 \notin \Delta$). Then (using Lemma 4.6 in the 3d case) we see that

$$T_{\alpha_1, \beta_1}(x_1, x_2)T_{\alpha_2, \beta_2}(x_3, x_4) = P_1(x_1, x_2)T_{\gamma_1, \eta_1}(x_1, x_2)Q_1(x_1, x_2)$$

for suitable maps $P_1, Q_1$. Note that $\gamma_1 \in \Delta$ and $(x_1, x_2, x_3, x_4) \mapsto (\bar{x}_1, \bar{x}_2, \bar{x}_1, \bar{x}_2)$ is bijective. Similarly

$$T_{\alpha_3, \beta_3}(x_5, x_6)T_{\alpha_4, \beta_4}(x_7, x_8) = P_2(x_3, x_4)T_{\gamma_2, \eta_2}(x_3, x_4)Q_2(x_3, x_4)$$

where $\gamma_2 \in \Delta$ and $(x_3, x_4, x_5, x_6) \mapsto (\tilde{x}_3, \tilde{x}_4, \tilde{x}_3, \tilde{x}_4)$ is bijective.

Put $z = Q_1(x_1, x_2)P_2(x_3, x_4)$ and set

$$\bar{x}_3 = z^{\gamma_2 \eta_2^{-1}}x_3^{\gamma_1 \eta_1^{-1}}z^{-1}, \quad \bar{x}_4 = z^{\gamma_2 \eta_2^{-1}}x_4^{\gamma_1 \eta_1^{-1}}z^{-1}.$$

Lemma 4.7 gives

$$T_{\gamma_1, \eta_1}(\bar{x}_1, \bar{x}_2) \cdot z \cdot T_{\gamma_2, \eta_2}(\tilde{x}_3, \tilde{x}_4) = T_{\gamma_1, \eta_1}(\bar{x}_1, \bar{x}_2)T_{\gamma_2, \eta_2}(\bar{x}_3, \bar{x}_4)$$

where $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4)$. Now we repeat the first procedure, applied to the second pair of automorphisms $\eta_1, \eta_2$. This gives $\delta \in \{\eta_1, \eta_2, \eta_1 \eta_2\} \cap \Delta$, $\gamma \in \{\gamma_1, \gamma_2\}$ and a bijection $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) \mapsto (x_1^*, x_2^*, x_5^*, x_6^*)$ such that

$$T_{\gamma_1, \eta_1}(\bar{x}_1, \bar{x}_2)T_{\gamma_2, \eta_2}(\bar{x}_3, \bar{x}_4) = P_3(\bar{x}_5, \bar{x}_6)T_{\gamma, \delta}(x_1^*, x_2^*)Q_3(x_5^*, x_6^*).$$

Then

$$\prod_{i=1}^{4} T_{\alpha_i, \beta_i}(x_{2i-1}, x_{2i}) = PT_{\gamma, \delta}(x_1^*, x_2^*)Q$$

where $P = P_1(x_1, x_2)P_3(x_5, x_6)$ and $Q = Q_3(x_5, x_6)R(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4)$. The result follows. ■

**Proof of Proposition 4.5** Suppose first that $n = 2$. Write $T_i = T_{\alpha_i, \beta_i}(x_{2i-1}, x_{2i})$. Grouping these four at a time and applying the preceding lemma we see that

$$\prod_{i=1}^{4D} T_i = \prod_{j=1}^{D} P_j(y_j)T_{\gamma_j, \delta_j}(u_j)Q_j(y_j)$$

where $\gamma_j, \delta_j \in D, y_j \in S^{(6)}, u_j \in S^{(2)}$ and $(x_1, \ldots, x_{4D}) \mapsto (y_1, \ldots, y_D; u_1, \ldots, u_D)$ is a bijection. Using Lemma 4.7 we now conjugate the factors $T_{\gamma_j, \delta_j}(u_j)$ by $z_j = (P_1^{\lambda_{ij}}Q_1^{\mu_{ij}} \ldots Q_{j-1}^{\mu_{ij}^{-1}}P_{j}^{\lambda_{ij}^{-1}})^{-1}$, for suitable automorphisms $\lambda_{ij}, \mu_{ij} \in \Delta$, to obtain

$$\prod_{i=1}^{4D} T_i = \prod_{j=1}^{D} T_{\gamma_j, \delta_j}(\tilde{x}_{2j-1}, \tilde{x}_{2j}) \cdot R(\bar{x})$$

where $\bar{x} = (y_1, \ldots, y_D), R = \prod_{j=1}^{D} P_j(y_j)^{\lambda_j}Q_j(y_j)^{\mu_j}$ for certain automorphisms $\lambda_j, \mu_j \in \Delta$, and $\tilde{x}_{2j-1}, \tilde{x}_{2j}$ are obtained from $y_j$ by multiplying on the left and right by expressions depending only on $z_j = z_j(y_1, \ldots, y_D)$. The result follows.
Now we consider the general case where $|\Gamma : \Delta| = n > 2$. This follows the same pattern. Suppose first that $D = 1$. There exist $i, j$ with $1 \leq i \leq j \leq n$ such that $\gamma = \alpha_i \alpha_{i+1} \ldots \alpha_j \in \Delta$. Using Lemmas 4.6 and 4.7 repeatedly we get

$$\prod_{i=1}^{n} T_i = P(\overline{x}) T_{\gamma, \delta}(x^*) Q(\overline{x})$$

where $x \mapsto (x^*, \overline{x})$ is a bijection $S^{(2n)} \to S(2) \times S^{(2n-2)}$. Grouping the factors together $n$ at a time and applying this to each group of $n$ factors we get

$$\prod_{i=1}^{n^2} T_i = \prod_{i=1}^{n} P(\overline{x}_i) T_{\gamma_i, \delta_{j(i)}}(x^*_i) Q(x_i)$$

$$= \prod_{i=1}^{n} T_{\gamma_i, \delta_{j(i)}}(\overline{x}_i) \cdot R_i(x^1),$$

using Lemma 4.7 for the second step; here $x \mapsto (\overline{x}, x^1)$ is a bijection $S^{(2n^2)} \to S(2n) \times S^{(2n^2-2n)}$ and each $\gamma_i \in \Delta$.

There exist $k, l$ with $1 \leq k \leq l \leq n$ such that $\beta_j(k) \ldots \beta_j(l) = \delta \in \Delta$, and repeating the procedure we get

$$\prod_{i=1}^{n} T_{\gamma_i, \delta_{j(i)}}(\overline{x}_i) = T_{\gamma_k, \delta}(x_1) \cdot R_2(x^5)$$

where $\overline{x} \mapsto (x_1, x^5)$ is a bijection $S^{(2n)} \to S(2) \times S^{(2n-2)}$. So putting $\gamma = \gamma_k$ we have

$$\prod_{i=1}^{n^2} T_i = T_{\gamma, \delta}(x_1) \cdot R(x^1, x^5) \quad (20)$$

with $\gamma, \delta \in \Delta$ and $x \mapsto (x^1, x^1, x^5)$ a bijection $S^{(n^2)} \to S^2 \times S^{(2n^2-2n)} \times S^{(2n-2)}$.

In the general case where $D > 1$ we group the $n^2 D$ factors $T_i$ together $n^2$ at a time, apply (20) to each product of $n^2$ factors, and then conjugate the resulting terms $T_{\gamma(l), \delta(i)}(x^1_l)$ by the intervening factors $R$ using Lemma 4.7 to obtain

$$\prod_{i=1}^{n^2 D} T_i = \prod_{i=1}^{D} \prod_{\gamma(l), \delta(i)}(\overline{x}) \cdot R(\overline{x})$$

for a certain bijection $x \mapsto (\overline{x}, \overline{x}) : S^{(n^2 D)} \to S^{((2n^2-2n)D)} \times S^{(2D)}$.

This completes the proof.

The final reduction step needs the next three lemmas.

**Lemma 4.9** Let $\varepsilon \in (0, 1)$. If $Z \subseteq X \times Y$ satisfies $|Z| \geq (1 - \varepsilon^2) |X \times Y|$ then for at least $(1 - \varepsilon) |X|$ elements $u \in X$ we have $|Z \cap (\{u\} \times Y)| \geq (1 - \varepsilon) |Y|$.
Proof. Suppose the number of such elements \( u \) is \( \rho |X| \). Then
\[
(1 - \varepsilon^2) |X \times Y| \leq (1 - \rho) |X| \cdot (1 - \varepsilon) |Y| + \rho |X| \cdot |Y|
\]
whence \( \rho \geq 1 - \varepsilon \). \( \blacksquare \)

Lemma 4.10

\[
T_{\alpha,\beta}(x, y) = [x, \alpha y][y, \beta]
\]
\[
z[x, \gamma] = [x', \gamma]z^\gamma
\]
where \( x' = xz^{-1} \).

Recall that for \( D \)-tuples \( x, \beta \), we use the notation \( c(x, \beta) = \prod_{i=1}^D [x_i, \beta_i] \), and \( x \cdot \beta = (x_1 \beta_1, \ldots, x_D \beta_D) \).

Lemma 4.11 There is a bijection \( y \mapsto \bar{\gamma} : S(D) \to S(D) \), and for each fixed \( y \in S(D) \) a bijection \( x \mapsto x' : S(D) \to S(D) \) (depending on \( y \)), such that
\[
T_{\alpha,\beta}(x, y) = c(x', \bar{\gamma} \cdot \alpha) \cdot h(y),
\]
where \( h(y) \) depends only on \( y \).

Proof. Using Lemma 4.10 we get
\[
T_{\alpha,\beta}(x, y) = \prod_{i=1}^D [x_i, \alpha_i y_i][y_i, \beta_i]
\]
\[
= \prod_{i=1}^D [x_i', \alpha_i y_i] \cdot z_D
\]
where \( x_i' = x_i z_i^{-1} \) and \( z_1 = 1, z_i = (z_{i-1}[y_{i-1}, \beta_{i-1}])^{\alpha_i y_i} \) for \( 1 < i \leq D \). The result follows on setting \( y_i = y_i^\alpha \). \( \blacksquare \)

Proposition 4.12 Let \( \alpha, \beta \in \text{Aut}(S) \). Suppose that for each \( Y \subseteq S^{(D)} \) with \( |Y| \geq (1 - \varepsilon)|S^{(D)}| \) there exists \( y \in Y \) such that
\[
X \subseteq S^{(D)}, |X| \geq (1 - \varepsilon)|S^{(D)}| \implies |c(X, y \cdot \alpha)| \geq \lambda |S|.
\]
Then \( P(\alpha, \beta; D, \varepsilon^2) \) holds.

Here \( c(X, y \cdot \alpha) = \{ c(x, y \cdot \alpha) \mid x \in X \} \).

Proof. Suppose that \( W \subseteq S^{(2D)} \) satisfies \( |W| \geq (1 - \varepsilon^2)|S^{(2D)}| \). Let
\[
Y = \left\{ y \in S^{(D)} \mid |W \cap (S^{(D)} \times \{ y \})| \geq (1 - \varepsilon)|S^{(D)}| \right\}.
\]
Lemma 4.10 shows that \( |Y| \geq (1 - \varepsilon)|S^{(D)}| \), so we can choose \( y \in Y \) so that \( |W| \geq (1 - \varepsilon)|S^{(D)}| \) holds. There exists \( X \subseteq S^{(D)} \) with \( |X| \geq (1 - \varepsilon)|S^{(D)}| \) and \( X \times \{ y \} \subseteq W \). Let \( x \mapsto x' \) be the bijection \( S^{(D)} \to S^{(D)} \) given in Lemma 4.11. Then
\[
T_{\alpha,\beta}(W) \geq T_{\alpha,\beta}(X \times \{ y \}) = c(X', \bar{\gamma} \cdot \alpha) \cdot h(y),
\]
a set of size at least \( \lambda |S| \). \( \blacksquare \)
4.1.2 Small groups

Let $N^*$ be an upper bound for the orders of quasisimple groups $S$ such that $l(S) = 2$; that $N^*$ is finite follows from Proposition 4.22 and well-known facts about the alternating groups. We fix a natural number $N_0 \geq N^*$, to be specified later, and denote by $\mathcal{S}$ the class of all quasisimple groups of order less than $N_0$. Set

$$N_1 = \max_{S \in \mathcal{S}} |\text{Out}(S)|.$$

There is a natural number $\delta_1$ such that for each $S \in \mathcal{S}$, every element of $S$ is a product of $\delta_1$ commutators (obviously $\delta_1 \leq \delta^*$, given in Proposition 4.19; in fact we can take $\delta_1 \leq 2$).

Define

$$\gamma : G \times G \to G$$

$$\gamma(x, y) = [x, y].$$

**Lemma 4.13** Let $\alpha, \beta \in \text{Inn}(S)$. Then there exist a bijection $(x, y) \mapsto (\overline{x}, \overline{y})$ from $S^{(2)}$ to $S^{(2)}$ and an element $t \in S$ such that

$$T_{\alpha, \beta}(x, y) = [\overline{x}, \overline{y}]t$$

for all $x, y \in S$.

**Proof.** For simplicity, let $\alpha$ and $\beta$ denote also elements of $S$ inducing the given inner automorphisms. Now define

$$t = [\alpha^{-1}, \beta],$$

$$(\overline{x}, \overline{y}) = (t\beta^{-1}x, t\beta^{-1}y\beta t^{-1}).$$

$\blacksquare$

**Proposition 4.14** If $S \in \mathcal{S}$ then $\mathcal{P}(\text{Aut}(S); D, \varepsilon)$ holds for $\lambda = 1$, with

$$D = N_2^2\delta_1, \quad \varepsilon = N_0^{-2\delta_1}.$$ 

**Proof.** In view of Proposition 4.14 it will suffice to establish $\mathcal{P}(\text{Inn}(S); \delta_1, \varepsilon)$. Let $X \subseteq S^{(2\delta_1)}$ satisfy $|X| \geq (1 - \varepsilon)|S|^{2\delta_1}$; then $X = S^{(2\delta_1)}$. Let $\alpha, \beta \in \text{Inn}(S)$. Using Lemma 4.13 we obtain

$$T_{\alpha, \beta}(X) = \prod_{i=1}^{\delta_1} T_{\alpha, \beta_i}(S, S)$$

$$= \prod_{i=1}^{\delta_1} \gamma(S \times S) t_i = \gamma(S \times S)^{\delta_1} \cdot t = S,$$

where $t = t_1 \ldots t_{\delta_1}$. The result follows. $\blacksquare$
4.1.3 Inner automorphisms

The key to this case is a result is due to Garion and Shalev. In [GaSh] they define for each finite group $G$ the invariant

$$
\epsilon(G) = (\zeta^G(2) - 1)^{1/4},
$$

where $\zeta^G(2) = \sum \chi^{-2}$ summed over irreducible characters $\chi$ of $G$.

**Proposition 4.15** ([GaSh], Corollary 1.4(ii)) For $W \subseteq G \times G$ and $\eta \in (0, 1)$,

$$
|W| \geq (1 - \eta)|G|^2 \implies |\gamma(W)| \geq (1 - \eta - 3\epsilon(G))|G|.
$$

This is useful in combination with Theorem 1.1 of [LiSh2], which implies that $\zeta^G(2) \to 1$ as $|G| \to \infty$ when $G$ ranges over quasisimple groups. We may therefore choose $N_2 \in \mathbb{N}$ so that $\epsilon(S) < \frac{1}{24}$ for every quasisimple group $S$ with $|S| \geq N_2$.

**Proposition 4.16** Let $S$ be a quasisimple group with $|S| \geq N_2$. Then $P(\text{Inn}(S); 1, \frac{1}{8})$ holds with $\lambda = \ell(S)^{-3/5}$.

**Proof.** Let $\alpha, \beta \in \text{Inn}(S)$ and let $X \subseteq S^{(2)}$ satisfy $|X| \geq \frac{7}{8}|S|^2$. According to Lemma 4.13 there exist $t \in S$ and a subset $Y$ of $S^{(2)}$ with $|Y| = |X|$ such that $T_{\alpha, \beta}(X) = \gamma(Y)t$.

By Proposition 4.15 we have

$$
|\gamma(W)| \geq \left(\frac{7}{8} - 3\epsilon(S)\right)|S| \geq \frac{3}{4}|S|.
$$

Therefore

$$
|T_{\alpha, \beta}(X)| = |\gamma(Y)| \geq |\gamma(W)| \geq \frac{3}{4}|S| > \ell(S)^{-3/5}|S|
$$

since $\ell(S) \geq 2$. 

4.1.4 Diagonal automorphisms

In this subsection and the next, we consider a quasisimple group $S$ of Lie type, of untwisted rank $r$. This means ([GLS], Section 2.2) that $S$ is the group of fixed points of a Steinberg automorphism $\sigma$ of order $k \in \{1, 2, 3\}$ of some untwisted Lie type group $S^\sigma \leq \text{GL}_d(q^k)$ of rank $r$ (where $k = 1$ precisely when $S = S^\sigma$ is untwisted). We denote by $D \leq \text{GL}_d(q^k)$ the group of diagonal matrices that induce diagonal automorphisms on $S$. Thus $S < SD$ and the restriction to $S$ of the inner automorphisms of $SD$ is the group $\text{InnDiag}(S)$ of inner-diagonal automorphisms of $S$. We will use the facts (loc. cit. Section 2.5):

$$
|SD : S| \leq r + 1,
$$

$$
|Z(SD)| = |Z(S)| \leq r + 1.
$$
An abelian subgroup of $S$ consisting of semisimple elements and maximal with this property will be called a maximal torus of $S$ (this is the same as the intersection with $S$ of a maximal torus in the underlying algebraic group). The following estimate is easily derived from [C], Proposition 3.3.5:

**Lemma 4.17** The size of a maximal torus of $S$ is at most $(q + 1)^r$.

**Proposition 4.18** There exists $N_3 \in \mathbb{N}$ such that if $|S| \geq N_3$, $r \geq 9$ and $q > 10$ then $\mathcal{P}(\text{InnDiag}(S); 8, 10^{-3})$ holds with $\lambda = l(S)^{-3/5}$.

This will be deduced from the next two results:

**Proposition 4.19** [GL] If $S$ is a classical group and $h \in D$ then the number of regular semisimple elements in the coset $Sh$ is at least $(1 - \frac{3}{q-1} - \frac{1}{(q-1)^2}) |S|$, which exceeds $\frac{2}{3} |S|$ if $q > 10$.

(This follows from the proof of [GL], though it is not explicitly stated there in this form.)

**Proposition 4.20** Assume that $r \geq 9$ and $q > 10$. Let $h_1, h_2, \ldots, h_8$ be regular semisimple elements of $SD$, and let $X \subseteq S^{(8)}$ satisfy $|X| \geq \frac{1}{4} |S|^8$. Then provided $|S|$ is sufficiently large, the number of elements $g \in S$ such that

$$c(x, h) = \prod_{i=1}^{8} [x_i, h_i] = g$$

has a solution $x = (x_1, \ldots, x_8) \in X$ is at least $\frac{1}{3} |S|$.

Before proving this let us deduce Proposition 4.18. Let $\alpha, \beta \in \text{InnDiag}(S)^{(8)}$ and let $Y \subseteq S^{(8)}$ satisfy $|Y| \geq (1 - (\frac{2}{3})^8) |S|^8$. There exist $c_i \in S$ and $h_i' \in D$ such that $\alpha_i$ is induced by $c_i h_i'$ ($i = 1, \ldots, 8$). Put $Y' = \{ y \cdot c \cdot h' \mid y \in Y \}$. Then $|Y'| = |Y|$, so Proposition 4.19 ensures that $Y'$ contains at least one element $y \cdot c \cdot h' = (h_1, \ldots, h_8)$ with each $h_i$ regular semisimple. Then provided $|S|$ is sufficiently large, Proposition 4.20 gives

$$|c(X, y \cdot \alpha)| = |c(X, h)| \geq \frac{1}{6} |S|$$

whenever $X \subseteq S^{(8)}$ satisfies $|X| \geq \frac{1}{4} |S|^8$. Applying Proposition 4.12 we infer that $\mathcal{P}(\alpha, \beta; 8, (\frac{2}{3})^8)$ holds with $\lambda = \frac{1}{6}$. Now Proposition 4.18 follows, since $(\frac{2}{3})^{32} > 10^{-3}$ and $l(S) \geq \frac{1}{2}(11^9 - 1) > 6^{5/3}$ by Proposition 1.22.

**Proof of Proposition 4.20** Relabelling $(h_1^{-1}, h_2^{-1}, \ldots, h_8^{-1})$ as $(k_1, k_2, \ldots, k_8)$ and $(x_1, x_2^{-1}, \ldots, x_8^{-1})$ as $(y_1, \ldots, y_8)$, it will suffice to prove that the image of the map

$$f : (y_1, \ldots, y_8) \mapsto k_1^{y_1} k_2^{y_2} \cdots k_8^{y_8} \cdot (k_1 \cdots k_8)^{-1} \in S$$

is at least $\frac{1}{3} |S|$. For each $i = 1, \ldots, 8$ let $x_i = k_i^{-1} y_i k_i$. Then $x_i \in S^{(8)}$ and $|x_i| \geq \frac{1}{4} |S|^8$. Thus $|X| \geq \frac{1}{4} |S|^8$ and the proposition follows.
has size at least $\frac{1}{6}|S|$ when $(y_1, \ldots, y_8)$ ranges over a subset of $S^{(8)}$ of proportion $\frac{1}{4}$.

Write $G = S D$. We observe that if $g$ is a semisimple element of $S$ then $C_G(g)$ contains a maximal torus of $G$ and so maps onto $G/S$. This means that the conjugacy class $g^S$ of $g$ in $S$ is the same as the conjugacy class of $g$ in $G$. Now we count solutions in conjugacy classes of $G$:

**Lemma 4.21** Assume that $r \geq 9$, $q > 10$. Let $\delta > 0$ and let $k_1, \ldots, k_8$ be regular semisimple elements of $G$. Put $c_i = |k_i^G|$. There is an integer $N_\delta$ such that if $|S| \geq N_\delta$ then the following holds:

For every $g \in S$ the number of 8-tuples $(a_1, \ldots, a_8) \in k_1^G \times \cdots \times k_8^G$ such that

$$a_1 \cdots a_8 = g k_1 \cdots k_8$$

is

$$\frac{c_1 \cdots c_8}{|S|} (1 + \gamma_g)$$

where $|\gamma_g| < \delta$.

Assuming this for the moment we can finish the proof of Proposition 4.20.

Take $\delta = \frac{1}{2}$ and assume that $|S| \geq N_\delta$. Then Lemma 4.21 implies that for each $g \in S$ we have

$$|f^{-1}(g)| = \prod_{i=1}^8 |C_S(k_i)| \cdot \frac{c_1 \cdots c_8}{|S|} (1 + \gamma_g)$$

$$= |S|^7 (1 + \gamma_g) < \frac{3}{7} |S|^7.$$

Suppose that $Y \subseteq S^{(8)}$ satisfies $|Y| \geq \frac{1}{4} |S|^8$. Then

$$|f(Y)| > \frac{|Y|}{\frac{3}{7} |S|^7} \geq \frac{1}{6} |S|,$$

as required.

**Proof of Lemma 4.21** Let $\chi$ be an irreducible character of $G$. By Clifford theory $\chi \downarrow S$ is a sum of irreducible characters of $S$, say $\psi + \phi + \cdots$. Then $\chi(1) \geq \psi(1)$. Now if $\chi$ is nonlinear then $\psi \in \text{Irr}(S)$ is also nonlinear, and hence $\chi(1) \geq \psi(1) \geq cq^r$ for some absolute constant $c$, by Proposition 1.22.

Put $p = g k_1 \cdots k_8$ and let $s(p)$ denote the number of the number of 8-tuples $(a_1, \ldots, a_8) \in k_1^G \times \cdots \times k_8^G$ such that $a_1 a_2 \cdots a_8 = p$.

A well-known formula (cf. [SGT], 7.2) gives

$$s(p) = \frac{c_1 \cdots c_8}{|G|} \sum_{\chi \in \text{Irr}(G)} \frac{\chi(k_1) \cdots \chi(k_8) \chi(p^{-1})}{\chi(1)^7}.$$

Since $k_1 \cdots k_8 p^{-1} \in S$ and hence lies inside $\ker \chi$ for any linear character $\chi$ of $G$, these contribute precisely $|G/G'| = |G|/|S|$ to the above sum. It therefore suffices to show that
where ${\text{Irr}}_0(G)$ denotes the set of non-linear irreducible characters of $G$.

Since $|\chi(p^{-1})|/\chi(1) \leq 1$ it is enough to show that

$$V := (r + 1) \sum_{\chi \in {\text{Irr}}_0(G)} \frac{\chi(k_1) \cdots \chi(k_8)}{\chi(1)^6} \to 0 \text{ as } |S| \to \infty.$$  

Now since $k_1$ is regular semisimple, $C_G(k_1)$ is a torus of $G = SD$, and so $|C_G(k_1)| \leq (q + 1)^{r+1}$ by Lemma 4.1.17. Hence $|\chi(k_1)| \leq \sqrt{|C_G(k_1)|} \leq (q + 1)^{(r+1)/2}$, and we obtain

$$|\chi(k_1) \cdots \chi(k_8)| \chi(1)^{-6} \leq \frac{(q + 1)^{r+1}^4}{c_6 q^{6r}} = c \cdot 6(q + 1)^{4+4r} q^{-6r}.$$  

By Corollary 1.2 (3) of [FG], $|{\text{Irr}}(G/Z(G))| \leq 100q^r$, whence $|{\text{Irr}}(G)| \leq 100q^r(r + 1)$. Moreover $q + 1 < q^{1.1}$ when $q > 10$. Consequently

$$V \leq c_5(r + 1)2q^{4.4 - 6r}$$

for some absolute constant $c_5 > 0$.

As $r \geq 9$ we have $0.6r > 4.4$; consequently $V \to 0$ as $|S| \to \infty$, as required.

### 4.1.5 Field automorphisms

As in the preceding subsection, $S$ denotes a quasisimple group of Lie type, of un twisted rank $r$. We assume that $S$ is universal, and introduce some more notation (cf. [GLS], Section 2.2). $L$ is a simple simply connected algebraic group defined over $\mathbb{F}_p$, and $S = L_\sigma$ is the group of $\sigma$-fixed points of a Steinberg automorphism $\sigma$ acting on $L$. Here $k \in \{1, 2, 3\}$ and $\sigma^k$ is the smallest power of $\sigma$ which is a power of the Frobenius automorphism $[p]$ of $L$. In fact $\sigma$ is the product of a graph automorphism of $L$ and some power of $[p]$, so $\sigma$ commutes with all field automorphisms of $L$.

We consider $L$ as embedded in some $\text{GL}_d$. Then $\text{GL}_d$ contains a torus $T$ that normalizes $L$ and induces the diagonal automorphisms on $L$. In the same way $D = T_\sigma$ induces the diagonal automorphisms of $S = L_\sigma$.

We consider a field automorphism $\phi$ of $S$. Thus $\phi$ is the restriction to $S$ of $[p]^f$ for some $f$, and we shall denote $[p]^f$ also by $\phi$. Then $\phi^n = \sigma^k$ where $n$ is the order of $\phi$ as an automorphism of $S$.

Let $q_0$ denote the cardinality of the fixed field of $\phi$. Thus $q_0 = p^f$, while $\mathbb{F}_{q^k}$ is the fixed field of $\phi^n$, so $q^k = p^{nf}$ and

$$q_k = q_0^n. \quad (22)$$

We remark that $k \leq 2$ unless $S$ is of type $^3D_4$, with $r = 4$; and $q$ might be the square root of a non-square integer if $S$ is a Suzuki or Ree group.
For an algebraic subgroup $M$ of $\text{GL}_d(\overline{\mathbb{F}_p})$ we denote by $M_\phi$ the fixed-point set of $\phi$ in $M$. Later, we shall need to consider the groups

$$G = L_\phi, \quad H = T_\phi$$

Thus $G$ is an untwisted quasisimple group of Lie type, say $X$, over $\mathbb{F}_{q^0}$ of rank $r$ equal to the rank of $L$. The group $H$ induces the diagonal automorphisms on $G$. Since $\sigma$ commutes with $\phi$ it preserves $G$ and acts on it as an automorphism of order $k$ (since $\sigma^k = \phi^n$).

We shall consider automorphisms $\alpha = c h \phi^{-1}$ where

- $\phi$ is a field automorphism of $S$ having order $n > 50$,
- $h$ is a diagonal automorphism of $S$ (we identify $h$ with an element of $D$),
- $c$ is an inner automorphism of $S$ (we will identify $c$ with an element of $S$).

**Proposition 4.22** With $\alpha$ as above, $\mathcal{P}(\alpha, \beta; 1, \frac{3}{5})$ and $\mathcal{P}(\beta, \alpha^{-1}; 1, \frac{3}{5})$ hold for every $\beta \in \text{Aut}(S)$, with $\lambda = l(S)^{-3/5}$.

This will follow from

**Proposition 4.23** Let

$$W = \left\{ x \in S \mid |C_S(xh\phi^{-1})| < l(S)^{1/2} \right\}. \quad (23)$$

Then $|W| > \frac{4}{5} |S|$.

To deduce Proposition 4.22 suppose $Y \subseteq S$ satisfies $|Y| \geq \frac{1}{5} |S|$. Then $Y \cap W$ is non-empty; choose $y \in Y$ with $yc \in W$. Then $|C_S(ya)| < l(S)^{1/2}$, so for any subset $X$ of $S$ with $|X| \geq \frac{1}{5} |S|$ we have

$$|c(X, y\alpha)| = |\{ [x, y\alpha] \mid x \in X \}|$$

$$\geq |X| l(S)^{-1/2} \geq \frac{1}{5} l(S)^{-1/2} |S|.$$

With Proposition 4.12 this shows that $\mathcal{P}(\alpha, \beta; 1, \frac{2}{5})$ holds with $\lambda = \frac{1}{5} l(S)^{-1/2}$ (as $\frac{2}{5} < \left(\frac{2}{5}\right)^2$). Since

$$T_{\beta^{-1}, \alpha^{-1}}(x, y) = T_{\alpha, \beta}(y_{\alpha^{-1}, x_{\beta^{-1}}}),$$

this implies also that $\mathcal{P}(\beta^{-1}, \alpha^{-1}; 1, \frac{2}{5})$ holds with the same value of $\lambda$.

Suppose that $k \leq 2$. Then (22) implies that $q > 2^{25}$. Proposition 1.22 then implies that $l(S) \geq (q - 1)/2 \geq 2^{24}$. If $k = 3$ then $S = 3D_4(q)$ and

42
Proposition 1.22 gives $l(S) \geq (q^4 - 1)/2 > (2^{4.50/3} - 1)/2 > 2^{65}$. In any case, then, $l(S)^{-1/10} < 2^{-2.4} < \frac{1}{5}$, whence

$$\frac{1}{5} l(S)^{-1/2} > l(S)^{-3/5}.$$

Proposition 4.22 follows.

We proceed to the proof of Proposition 4.23. We are given $h \in D = T_\sigma$. By Lang’s theorem ([GLS], Theorem 2.1.1) we may choose $\kappa \in T$ with $h = \kappa^{-1} \kappa^\phi$. Put $h' = \kappa \kappa^{-1} \sigma$. Note that

$$(\kappa^{-1} \kappa^\phi)^\sigma = h^\sigma = h = \kappa^{-1} \kappa^\phi,$$

$$h'^\phi = (\kappa \kappa^{-1} \sigma)^\phi = \kappa \kappa^{-1} = h'$$

so $h' \in H$. Define

$$\mu, \nu : L \to LT$$

$$\mu(x) = [x \kappa, \phi], \ \nu(x) = [(x \kappa)^{-1}, \sigma].$$

**Lemma 4.24** (i) $\mu^{-1}(Sh) = \nu^{-1}(Gh')$;

(ii) if $g \in Sh$ then $|\mu^{-1}(g)| = |G|$;

(iii) if $z \in Gh'$ then $|\nu^{-1}(z)| = |S|$.

**Proof.** (i). If $x \in Sh$, then

$$\mu(x) \in Sh \iff \mu(x)^\sigma = \mu(x)$$

$$\iff \kappa^{-1} x^{-1} x^\phi \kappa^\phi = \kappa^{-1} x^{-\sigma} x^\phi \kappa^\sigma \kappa^\phi$$

$$\iff x^\sigma \kappa^{-1} x^{-1} \kappa x = x^\sigma \kappa \kappa^{-1} x^{-1} \kappa x$$

$$\iff \nu(x) = \nu(x)^\phi \iff \nu(x) \in Gh'.$$

(ii), (iii). Let $g \in Sh$. Then $g = \kappa^{-1} g' \kappa^\phi$ with $g' \in L$, and by Lang’s theorem again we have $g' = [x, \phi]$ for some $x \in L$. Then $\mu(x) = g$, and we see that $\mu^{-1}(g) = x L_\phi = x G$. Similarly we find that $\nu^{-1}(z) = y \kappa G_\sigma \kappa^{-1} = y S^{-1}$ where $y = \kappa y_1 \kappa^{-1}$ and $z = \kappa \cdot y_1 y_1^{-\sigma} \cdot \kappa^\sigma$. □

Now consider the semi-direct product $G_1 = GH \rtimes \langle \sigma \rangle$. We define a permutation action of $G$ on $G_1$ as follows: for $x \in G$ and $a \in G_1$,

$$a^x = x^{-1} ax^\sigma.$$

We will call this the *twisted action*. For $a \in G_1$ we denote the stabilizer of $a$ in $G$ under this action by $C(a)$, i.e.

$$C(a) = \{ x \in G \mid ax^\sigma = xa \}.$$

Set $Y = \mu^{-1}(Sh) = \nu^{-1}(Gh')$.

43
Lemma 4.25 Let \( y \in Y \) and put \( g = \mu(y) \), \( z = \nu(y) \). Then
\[
|C_S(g\phi^{-1})| = |C(z)|.
\]

Proof. Let \( a \in L \) and put \( b = y\kappa a^{-1}y^{-1} \). The condition \( a^{g\phi^{-1}} = a \) is equivalent to \( b^\phi = b \), i.e. \( b \in L_\phi = G \). The condition \( a \in S = L_\sigma \) is equivalent to \( (b^\sigma)^\sigma = b^\sigma \), i.e. \( z^\sigma = b z \). So
\[
C_S(g\phi^{-1}) = (y\kappa)^{-1}C(z)y\kappa.
\]

If we put
\[
Z = \left\{ z \in Gh' \mid |C(z)| < l(S)^{1/2} \right\},
\]
\[
Y^* = \nu^{-1}(Z),
\]
the two preceding lemmas give
\[
|W| = |G|^{-1} |Y^*| = |S| |G|^{-1} |Z|.
\]

Lemma 4.26 (i) If \( S \neq 3D_4(q) \) then \( l(S)^{1/2} > q_0^{11}\); (ii) If \( S = 3D_4(q) \) then \( l(S)^{1/2} > |G| \).

Proof. Proposition 1.22 says that \( l(S) \) is at least \( (q^r - 1)/2 \). Also \( q = q_0^{n/k} \geq q_0^{51/k} \).

In case (i) we have \( k \leq 2 \). Then \( l(S) > \frac{1}{2}(q_0^{25r} - 1) \), whence \( l(S) \geq q_0^{24r} \) and the result follows.

In case (ii), \( k = 3 \) and \( G = D_4(q_0) \). In this case, we have
\[
l(S) \geq (q^4 - 1)/2 \geq (q_0^{68} - 1)/2,
\]
\[
|G| < q_0^{28} < l(S)^{1/2}.
\]

Since \( C(z) \leq G \) for each \( z \in Gh' \), it follows in case (ii) that \( Z = Gh' \) and hence that \( |W| = |S| \).

Henceforth, we assume that \( S \neq 3D_4(q) \).

Let \( c(G) \) denote the number of conjugacy classes of \( G \).

Lemma 4.27 The coset \( Gh' \subseteq G_1 \) is a union of at most \( |G_1 : G| c(G) \) orbits of \( G \) with the twisted action.

Proof. For \( z \in Gh' \) and \( x \in G \) we have
\[
(z \cdot \sigma^{-1})^x = x^{-1} z x^\sigma \cdot \sigma^{-1} = z^x \cdot \sigma^{-1}
\]
in $G_1$. This shows that the twisted action on $G$ on the coset $Gh'$ is equivalent to the conjugation action of $G$ on $Gh'\sigma^{-1} \subseteq G_1$. The number of orbits of $G$ acting by conjugation on $G_1$ is

$$|G|^{-1} \sum_{g \in G} |C_{G_1}(g)| \leq |G|^{-1} |G_1 : G| \sum_{g \in G} |C_G(g)|$$

$$= |G|^{-1} |G_1 : G| |G| c(G) = |G_1 : G| c(G).$$

The result follows.

Since $G = L_\phi$ is a quasisimple group of untwisted Lie type,

$$|GH : G| \leq |\text{Outdiag}(G)| \leq \min\{r + 1, q_0 - 1\} < q_0.$$  

The automorphism $\sigma$ has order 1 or 2. Thus $|G_1 : G| \leq 2q_0$. Now Theorem 1.1 (1) in [FG] shows that $c(G) \leq 30q_0^r$. Applying Lemma 4.26, we deduce that if $y \in Gh' \setminus Z$ then

$$|\hat{y}| = \frac{|G|}{|\text{C}(y)|} < \frac{|G|}{q_0^{1r}}.$$  

Hence by Lemma 4.27 $Gh' \setminus Z$ is the union of at most $60q_0^{r+1}$ orbits of this size, whence

$$|Gh' \setminus Z| < 60q_0^{-10r+1} |G|.$$  

Therefore $|Z| \geq \eta |G|$ where $\eta = 1 - 60/2^9 > 4/5$.

Now Proposition 4.23 follows from [21].

4.1.6 Proof of Theorem 4.1

As explained in Subsection 4.1.1, we have to find $D \in \mathbb{N}$ and $\varepsilon > 0$ such that $P(\text{Aut}(S); D, \varepsilon)$ holds with $\lambda$ given by (19) for every quasisimple group $S$: i.e. $\lambda = l(S)^{-3/5}$ if $l(S) \geq 3$, $\lambda = 1$ if $l(S) = 2$. Henceforth, when we say that $P(\ldots)$ holds for some group $S$, we will mean that it holds with $\lambda$ given by (19).

Set $N_0 = \max\{N_2, N_3, 1 + |M|\}$ where $N_i$ are the bounds introduced above and $M$ denotes the largest sporadic (quasi)simple group (it happens to be simple).

Now let $S$ be a quasisimple group. We consider several cases.

Case 1. Where $|S| < N_0$. Proposition 4.14 shows that $P(\text{Aut}(S); D_1, \varepsilon_1)$ holds for some $D_1$ and $\varepsilon_1$.

We assume henceforth that $|S| \geq N_0$. Putting $\Gamma_0 = \text{Inn}(S)$, Proposition 4.16 shows that $P(\Gamma_0; 1, \frac{1}{8})$ holds.

Case 2. Where $S/Z(S)$ is an alternating group. Then $|\text{Aut}(S) : \Gamma_0| = 2$, and Proposition 4.1 gives $P(\text{Aut}(S); 4, \frac{1}{8})$.  

45
From now on, $S$ is a group of Lie type, of rank $r$ over $F_q$. We denote by $\Phi$ the group of field automorphisms of $S$. Then $\text{Aut}(S)$ has normal subgroups

$$\text{Aut}(S) \geq \Gamma \geq \Gamma_1 \geq \Gamma_2 \geq \Gamma_0 = \text{Inn}(S)$$

where $\Gamma_2 = \text{InnDiag}(S)$, $\Gamma = \Gamma_2 \Phi$, and $\Gamma_1 = \Gamma_2 \Phi_1$ where $\Phi_1$ is the subgroup of $\Phi$ generated by all elements of order at most 50.

Put $n_0 = \lcm[50]$, and define $n_1 = \min\{q + 1, r + 1\}$ if $S$ has type $A_r$ or $2A_r$, $n_1 = 4$ otherwise. We have

$$|\text{Aut}(S) : \Gamma| \leq 6,$$

$$|\Gamma : \Gamma_2| \leq \log_p(q^3) \leq 3 \log_2(q),$$

$$|\Gamma_1 : \Gamma_2| \leq n_0,$$

$$|\Gamma_2 : \Gamma_0| \leq n_1$$

where $p = \text{char}(F_q)$ (see [GLS], Section 2.5).

**Case 3.** Where $q \leq 10$. In this case, $|\text{Aut}(S) : \Gamma_0| \leq 600$. As in Case 2, we may deduce that $P(\text{Aut}(S); D_2, \frac{1}{8})$ holds where $D_2 = 360,000$.

**Case 4.** Where $q > 10$. If $r < 9$ we have $|\Gamma_2 : \Gamma_0| \leq n_1 \leq 9$; we deduce as before that $P(\Gamma_2; 81, \frac{1}{8})$ holds. If $r \geq 9$, Proposition 4.18 gives $P(\Gamma_2; 8, \frac{1}{8})$.

Taking $D_3 = 81n_0^2$, we infer in any case that $P(\Gamma_1; D_3, \frac{1}{8})$ holds, whatever the rank $r$.

Now let $\alpha_i, \beta_i \in \Gamma^{(D_3)}$. If $\alpha_i$ and $\beta_i$ lie in $\Gamma_1$ for every $i$ then we have $P(\alpha, \beta; D_3, \frac{1}{8})$. If not, let us suppose for convenience that $\alpha_1 \notin \Gamma_1$. Then $\alpha_1 = c\phi$ where $c \in \Gamma_0$, $h$ is diagonal, and $\phi \in \Phi$ has order exceeding 50. Proposition 4.22 now shows that $P(\alpha_1, \beta_1; 1, 3/5)$ holds. As in the proof of Proposition 4.3, this in turn implies $P(\alpha, \beta; D_3, 3/5)$.

Thus $P(\Gamma; D_3, \frac{1}{8})$ holds in either case. Since $|\text{Aut}(S) : \Gamma| \leq 6$, a final application of Proposition 4.4 gives $P(\text{Aut}(S); D_4, \frac{1}{8})$ where $D_4 = 36D_3$.

**Conclusion.** Take $D = \max\{4, D_1, D_2, D_4\}$ and $\varepsilon = \min\{\varepsilon_1, \frac{1}{8}\}$. Then $P(\text{Aut}(S); D, \varepsilon)$ holds in all cases, by Proposition 4.3.

### 4.2 Commutators in semisimple groups

In this subsection, $D$ and $\varepsilon$ are the constants introduced in subsection 4.1. We will say that a multiset $Y$ has the $(k, \eta)$-fpp on a $(Y)$-set $\Omega$ if at least $k$ elements of $Y$ have the $\eta$-fpp on $\Omega$.

**Theorem 4.28** Let $N$ be a finite quasisemisimple group with at least 3 non-abelian composition factors. Let $y_1, \ldots, y_{10}$ be $m$-tuples of automorphisms of

---

46
N. Assume that for each $i$, the group $\langle y_i \rangle$ permutes the set $\Omega$ of quasisimple factors of $N$ transitively and that $y_i$ has the $(k, \eta)$-fpp on $\Omega$, where $k \eta \geq 4 + 2D$.

For each $i$ let $W(i) \subseteq N^{(m)}$ be a subset with $|W(i)| \geq (1 - \epsilon/6)|N|^m$. Then

$$\prod_{i=1}^{10} W(i) \phi(i) = N$$

where $\phi(i) : N^{(m)} \to N$ is given by

$$(x_1, \ldots, x_m) \phi(i) = \prod_{i=1}^{m} [x_i, y_{ij}].$$

The action of $\text{Aut}(N)$ lifts to an action on the universal cover $\tilde{N}$ of $N$, and $\tilde{N} = S_1 \times \cdots \times S_n$ where the $S_i$ are quasisimple groups. Replacing $N$ by $\tilde{N}$ and each $W(i)$ by its inverse image in $N^{(m)}$, we may suppose that in fact $N = S_1 \times \cdots \times S_n$. Since $\langle y_1 \rangle$ permutes $\Omega = \{S_1, \ldots, S_n\}$ transitively, the groups $S_i$ are all isomorphic to a quasisimple group $S$.

Now let $G = \langle g_1, \ldots, g_m \rangle \leq \text{Aut}(N)$ and denote by $e_i$ the number of cycles (including fixed points) of $g_i$ in its action on $\Omega$. Define $\phi : N^{(m)} \to N$ by

$$x \phi = c(x, g) = \prod_{i=1}^{m} [x_i, g_i].$$

We shall prove

**Proposition 4.29** Suppose that that

$$(m - 2)n - \sum_{i=1}^{m} e_i \geq 2D. \quad (25)$$

1. Let $W \subseteq N^{(m)}$ satisfy $|W| \geq (1 - \epsilon/6)|N|^m$. Then

$$|W\phi| \geq l(S)^{-4/5}|N|.$$

2. If $D$ is replaced by $D_1 = 5D$, then $\phi$ is surjective, and each fibre of $\phi$ has size at least $|N|^{-2D_1/n}|N|^{m-1}$.

Part (2) is a sharper version of [NS], Proposition 9.1; we will not be needing it, and include it in a revisionist spirit, to show how the main results of [NS] can be reproduced using these methods.

To deduce Theorem 4.28 from (1), note that for each $i = 1, \ldots, 10$, the total number of cycles for $y_{i1}, \ldots, y_{im}$ on $\Omega$ is at most

$$(m - k)n + k(1 - \eta/2)n \leq (m - 2)n - nD,$$

which implies condition (25) since $n \geq 3$. So taking $g = y_i$ and writing $\phi(i)$ for the corresponding map $\phi$, we may infer that

$$|W(i)\phi(i)| \geq l(S)^{-4/5}|N|.$$
Now \( l(S) = l(N) = l \), say, and we have
\[
\prod_{i=1}^{10} |W(i)\phi(i)| \geq \frac{|N|^{10}}{l^8}.
\]

It follows by the ‘Gowers trick’ that \( \prod_{i=1}^{m} W(i)\phi(i) = N \), and this is the statement of Theorem 4.28 since \( W(i)\phi(i) = m \prod_{j=1}^{e_i} [W(i), y_{ij}] \).

4.2.1 Proof of Proposition 4.29

**Lemma 4.30** Suppose that \( G = \langle g_1, \ldots, g_m \rangle \) acts transitively on a finite set \( J \). Fix \( t \in J \). Then there is a total order on \( J \) with minimal element \( t \) such that for each \( j > t \) there exist \( i(j) \in [m] \) and \( \varepsilon_j \in \{ \pm 1 \} \) such that \( j \cdot g_{i(j)}^{\varepsilon_j} < j \).

**Proof.** Let \( X \) with \( 1 \in X \) be a Schreier transversal to the right cosets of \( \text{stab}_G(t) \): thus \( X \) is a set of words on \( \{ g_1, \ldots, g_m \} \) such that (1) \( x \mapsto t \cdot x \) is a bijection \( X \to J \) and (2) each initial segment of a word in \( X \) is again in \( X \), i.e. if a word \( vg_i^{\varepsilon} \) is in \( X \) then \( v \in X \). Now define the size of \( j = t \cdot x \) to be the length of \( x \), and finally order \( J \) lexicographically by size. 

Keeping \( G \) and \( J \) as above, we label the elements of \( J \) as \( \{ 1, 2, \ldots, n \} \) in the given order, and fix \( i(j), \varepsilon_j \) \( (j = 2, \ldots, n) \) as in the lemma. Say \( g_i \) has cycles \( \Delta_{il}, l = 1, \ldots, e_i \) (including cycles of length 1); we also write
\[
\Delta_{il} = \Delta_{i,(j)}(j) \text{ if } j \in \Delta_{il}.
\]
Let \( \delta_{il} = \delta_{i,(j)}(j) \) denote the least member of \( \Delta_{il} = \Delta_{i,(j)} \), and set
\[
\hat{j} = \delta_{il(j)}(j),
\]
i.e. \( \hat{j} \) is the least element in the \( \langle g_{i(j)} \rangle \)-orbit of \( j \). This implies that \( \hat{j} < j \) if \( j > 1 \).

Put
\[
\Delta'_{il} = \Delta_{il} \setminus \{ \delta_{il} \},
\]
\[
J' = \bigcup_{l=1}^{e_i} \Delta'_{il}.
\]
In writing products labelled by \( \Delta_{il} \), we will assume that \( \Delta_{il} \) is ordered as a \( g_{il} \)-cycle starting with \( \delta_{il} \) (not with the induced order from \( J \)).

Let \( S \) be a finite group, \( N = S^J \), and suppose that \( G \) acts on \( N \), permuting the factors according to the action of \( G \) on \( J \). Write elements of \( N \) as \( x = (x(j))_{j \in J} \).
For any subset $T$ of $[m] \times J$ write $\pi_T : N^{(m)} \to S^T$ for the projection map

$$(x_1, \ldots, x_m)\pi_T = (x_i(j))_{(i,j)\in T}.$$ 

For $x \in S$ an expression $x^\alpha$ will mean $x^\alpha$ where $\alpha$ is some fixed automorphism of $S$, depending on the context but not on $x$, and $x^{-*} = (x^*)^{-1}$.

We write $[x, g] = ([x_1, g_1], \ldots, [x_m, g_m])$.

**Lemma 4.31** Let $x, y \in N$. Then $[x, g_i] = y$ if and only if

$$y(\delta_{il}) = x(\delta_{il})^{-1} x(\delta_{il})^* \prod_{j \in \Delta_{il}} y(j)^{-*} \quad (26)$$

$$x(j) = x(j)^* y(j)^{-1} \quad (j \in \Delta_i') \quad (27)$$

for $1 \leq l \leq e_i$, where $j^- = j \cdot g_i^{-1}$.

**Proof.** Compare the $j$-components of $u = [x, g_i]$ and of $y$ as $j$ runs over a given cycle $\Delta_{il}$. To simplify notation let’s suppose that $\Delta_{il} = (1, 2, \ldots, s)$, with $\delta_{il} = 1$. For $1 \leq j \leq s$ we have

$$u(j) = x(j)^{-1} x(j - 1)^{\alpha_j}$$

(writing $x(0) = x(s)$) where $\alpha_j \in \text{Aut}(S)$ depends on $j$ and $g_i$. Using these to eliminate $x(2), \ldots, x(s)$ in turn we get

$$x(1)^{-1} x(1)^\beta = u(1) u(s)^{\alpha_s} u(s - 1)^{\alpha_{s-1}} \ldots u(2)^{\alpha_2 \ldots \alpha_s},$$

where $\beta = \alpha_1 \ldots \alpha_s$ is the automorphism induced by $g_i^*$ on the first component of $S^\Delta_{il}$. Thus (26) and (27) hold with $u$ in place of $y$. The lemma follows since these equations determine $y$ uniquely, given $x$. \qed

Put

$$C = \{(i, \delta_{il}) \mid 1 \leq i \leq m, \ 1 \leq l \leq e_i\}$$

$$K = \{(i, j) \mid 1 \leq i \leq m, \ j \in J_i'\}$$

$$K' = K \setminus \{(i(j), j) \mid j = 2, \ldots, n\}.$$ 

Define $\Theta : N^{(m)} \to S^C \times S^K = S^{(mn)}$ by

$$x\Theta = (x\pi_C, [x, g]_{\pi_K}).$$

**Lemma 4.31** shows that $\Theta$ is bijective.

Now define $\phi : N^{(m)} \to N = S^{(n)}$ by

$$x\phi = \prod_{i=1}^m [x_i, g_i]$$

$$= (x\phi_1, \ldots, x\phi_n).$$
Define $\Psi : N^{(m)} \rightarrow S^C \times S^{K'} \times S^{(n-1)} = S^{(mn)}$ by

$$x\Psi = (x\pi_{C}, [x, g]\pi_{K'}, (x\phi_{2}, \ldots, x\phi_{n})).$$

**Lemma 4.32** The mapping $\Psi : S^{(mn)} \rightarrow S^{(mn)}$ is bijective.

**Proof.** Let $(u, v, z_{2}, \ldots, z_{n}) \in S^C \times S^{K'} \times S^{(n-1)}$. We have to show that there exists a unique $x \in N^{(m)}$ such that $x\pi_{C} = u, [x, g]\pi_{K'} = v$ and $x\phi_{j} = z_{j}$ for $j = 2, \ldots, n$.

Since $\Theta$ is bijective, for each tuple $\eta = (\eta_{2}, \ldots, \eta_{n}) \in S^{(n-1)}$ there exists a unique $x \in N^{(m)}$ with

$$x\pi_{C} = u, [x, g]\pi_{K'} = v,$$

$$[x_{i(j)}, g_{i(j)}](j) = \eta_{j} \quad (j = 2, \ldots, n).$$

Write $y_{i} = [x_{i}, g_{i}]$. Then

$$x\phi_{j} = y_{i}(j)y_{2}(j) \ldots y_{m}(j).$$

If $(i, j) \in K'$ then $y_{i}(j)$ is the $(i, j)$-component of $[x, g]\pi_{K'} = v$. If $(i, j) \in C$ then $y_{i}(j)$ is determined by equation (26); this involves $x_{i}(j)$, a component of $x\pi_{C} = u$, and further factors $y_{i}(r)$ where $r > j$.

If $(i, j) \notin C \cup K'$ then $i = i(j)$ and $y_{i}(j) = \eta_{j}$. Now we can solve the equations

$$\eta_{j} = y_{i-1}(j)^{-1} \ldots y_{i}(j)^{-1}z_{j}y_{m}(j)^{-1} \ldots y_{i+1}(j)^{-1}$$

successively for $j = n, n - 1, \ldots, 2$, uniquely for $\eta$. The result follows. □

Observe now that $x\phi = (z_{1}, \ldots, z_{n})$ if and only if

$$x\phi_{1} = z_{1} \quad (29)$$

and

$$x\Psi = (u, v, z_{2}, \ldots, z_{n}) \quad (30)$$

for some $(u, v) \in S^C \times S^{K'}$.

Putting $y_{i} = [x_{i}, g_{i}]$ as above we have

$$x\phi_{1} = y_{1}(1)y_{2}(1) \ldots y_{m}(1). \quad (31)$$

Now the following hold:

If $(i, j) \in C$ then $j = \delta_{il}$ for some $l \leq e_{i}$, and

$$y_{i}(j) = x_{i}(j)^{-1}x_{i}(j)^{*} \prod_{k \in \Delta_{il}} y_{k}(k)^{-*}; \quad (S(i, j))$$

note that for each factor $y_{i}(k)$ occurring on the right we have $(i, k) \notin C$ and $k > j$. 50
If \( i = i(j) \) then

\[
y_i(j)^{-1} = y_{i+1}(j) \cdots y_m(j) z_j^{-1} y_i(j) \cdots y_i^{-1}(j); \tag{S(j)}
\]

note that for each factor \( y_r(j) \) occurring on the right we have \( r \neq i(j) \).

Now we are going to successively transform the right-hand member of (31) in the following manner: for some \((i, j) \in C\), substitute for the factor \( y_i(j) \) the expression on the right-hand side of \( S(i, j) \); then use \( S(k) \) to eliminate one of the newly introduced factors \( y_r(k) \).

To analyse this process, for the time being we consider the \( y_i(j) \), \( y_i(j)^{-1} \), \( x_i(j) \), \( x_i(j)^{-1} \) and \( z_j^{-1} \) as abstract symbols (but allowing the automorphisms denoted by * to distribute over the factors in the usual way). If \( U \) is a product of such symbols, possibly decorated with *s, the support \( \text{sup}(U) \) is the multiset of symbols that occur in \( U \), with their multiplicities. For \((i, j) \in C\) let \( Y_{ij} \) denote the right-hand side of \( S(i, j) \), and for \((i, j) \notin C\) set \( Y_{ij} = y_i(j) \).

For \( j = 2, \ldots, n \) put

\[
Z_j = Y_{i(j)+1,j} \cdots Y_{m,j} z_j^{-1} Y_{1,j} \cdots Y_{i(j)-1,j}.
\]

Then

\[
\text{sup}(Y_{ij}) = \{x_i(j)^{-1}, x_i(j), y_i(k)^{-1} \mid k \in \Delta'_j(j)\} \text{ if } (i, j) \in C,
\]

\[
\text{sup}(Y_{ij}) = \{y_i(j)\} \text{ if } (i, j) \notin C
\]

and

\[
\text{sup}(Z_j) = \{z_j^{-1}\} \cup \bigcup_{i \neq i(j)} \text{sup}(Y_{ij})
\]

(disjoint union).

Now set

\[
U_1 = \prod_{i=1}^m Y_{i1}.
\]

Then

\[
\text{sup}(U_1) = \bigcup_i \text{sup}(Y_{i1}) \ni y_{i(2)}(2)^{-1},
\]

because \((i, 1) \in C\) for every \( i \), and \( 2 \in \Delta'_{i(2)}(1) \). Let \( U_2 \) be the expression obtained from \( U_1 \) on replacing \( y_{i(2)}(2)^{-1} \) by \( Z_2 \). Then

\[
\text{sup}(U_2) = \text{sup}(U_1) \cup \text{sup}(Z_2) \setminus \{y_{i(2)}(2)^{-1}\}
\]

\[
= \bigcup_i \text{sup}(Y_{i1}) \cup \{z_2^{-1}\} \cup \bigcup_{i \neq i(2)} \text{sup}(Y_{i2}) \setminus \{y_{i(2)}(2)^{-1}\}.
\]

Iterating this process, suppose that after \( j - 1 < n - 1 \) steps we obtain \( U_j \), where \( \text{sup}(U_j) \) contains

\[
\bigcup_{r=1}^j \left( \bigcup_{i \neq i(r)} \text{sup}(Y_{ir}) \setminus \{y_{i(r)}(r)^{-1}\} \right). \tag{32}
\]
Say $j + 1 = r$, so $r \leq j$ and $j + 1 \in \Delta^*_r(j + 1)$. Then $(i(j + 1), r) \in C$ and (if $r > 1$) $i(j + 1) \neq i(r)$, so $y_{i(j+1)}(j + 1)^{-1} \in \sup(Y_{i(j+1)}, r) \subseteq \sup(U_j)$. Now replace $y_{i(j+1)}(j + 1)^{-1}$ in $U_j$ by $Z_{j+1}$ to obtain $U_{j+1}$. Then the analogue of $\{y\}$ holds with $j + 1$ for $j$.

After $n - 1$ such steps we obtain an expression $U = U_n$ with

$$\sup(U) = \mathcal{X} \cup \mathcal{Y} \cup \mathcal{Z}$$

where

$$\mathcal{X} = \{x_i(j), x_i(j)^{-1} | (i, j) \in C\}, \quad \mathcal{Y} = \{y_i(j), y_i(j)^{-1} | (i, j) \in K'\},$$

$$\mathcal{Z} = \{z_2^{-1}, \ldots, z_n^{-1}\}.$$

To any formal product $V$ of factors $x_i(j)^{\pm*}$, $y_i(j)^{\pm*}$, $z_j^{*-*}$ we assign a numerical sequence $\tau(V)$ as follows: reading $V$ from left to right, ignore all factors $x_i(j)^{\pm*}$ and $z_j^{*-*}$; to each factor $y_i(j)^{*}$ assign the label $i$, and to each maximal product of consecutive terms of the form $y_i(k)^{-*}$ (fixed $i$, varying $k$) assign the label $i$.

**Claim 1:** For each $j = 1, \ldots, n$, $\tau(U_j)$ is a subsequence of

$$S(j) = (1, \ldots, m, 1, \ldots, m, 1, \ldots, m)$$

where $1, \ldots, m$ is repeated $j$ times.

**Proof.** This is clear for $j = 1$. Let $j \geq 1$ and suppose inductively that $\tau(U_j)$ is a subsequence of $S(j)$. Put $i = i(j + 1)$; then $y_i(j + 1)^{-*}$ is a factor in $U_j$, and we obtained $U_{j+1}$ by replacing it with $Z_{j+1}^*$. Thus

$$\tau(U_j) = (I_1, P, i, Q, I_2)$$

where $(P, i, Q)$ is a subsequence of $(1, \ldots, m)$, $I_1$ is a subsequence of $S(p)$ and $I_2$ is a subsequence of $S(q)$ and $p + 1 + q = j$ (here $p$ or $q$ could be 0, with $S(0) = \emptyset$); the displayed $i$ is due to $y_i(j + 1)^{-*}$. Substituting $Z_{j+1}^*$ for $y_i(j + 1)^{-*}$ has the effect of replacing $i$ by $(\underline{i}, i + 1, \ldots, m, 1, \ldots, i - 1, \underline{i})$, where the underlined is may or may not be present (depending on whether $y_i(j + 1)^{-*}$ appears in the middle or at either end of a product of consecutive terms of the form $y_i(k)^{-*}$). In any case,

$$(P, \underline{i}, i + 1, \ldots, m, 1, \ldots, i - 1, \underline{i}, Q)$$

is a subsequence of $S(2)$, and so $\tau(U_{j+1})$ is a subsequence of

$$(S(p), S(2), S(q)) = S(p + 2 + q) = S(j + 1).$$
Claim 2: There exist 2D distinct elements $\xi_1, \eta_1, \ldots, \xi_D, \eta_D$ of $Y$ such that the following holds. There exist $R, A_i, B_i, C_i, D_i$ ($i = 1, \ldots, D$), each of which is a product of factors $t^*$ with $t \in \mathcal{X} \cup \mathcal{Z} \cup Y \setminus \{\xi_1, \eta_1, \ldots, \xi_D, \eta_D\}$, such that

$$U_n \simeq \prod_{i=1}^{D} (A_i \xi_i B_i)^{-*} (C_i \eta_i D_i)^{-*} (A_i \xi_i B_i)^{*} (C_i \eta_i D_i)^{*} \cdot R, \quad (33)$$

meaning that the two sides represent the same element in the free group on all the occurring symbols $t^*$, $t \in \mathcal{X} \cup \mathcal{Y} \cup \mathcal{Z}$.

Proof. This follows from hypothesis (25) and Claim 1 by (the proof of) [NS], Prop. 8.4.

To complete the proof of Proposition 4.29 we need one further lemma:

**Lemma 4.33** Let $K'' \subseteq K'$ be a set of size $|K'| - 2D \geq 0$. Suppose that $W \subseteq N^{(m)}$ satisfies $|W| \geq (1 - \varepsilon/q) |N^{(m)}|$. Let $P$ be the set of elements $z \in S^{(n-1)}$ for which there exist $u \in SC$, $v \in S^{n-1}$ such that

$$|\{w \in W \mid w \Psi_{C \cup K'' \cup \mathcal{Y} \cup \mathcal{Z}} = (u, v, z)\}| \geq (1 - \varepsilon) |S|^{2D}. \quad (33)$$

Then $|P| \geq (1 - \frac{1}{q}) |S|^{n-1}$.

**Proof.** Put $\sigma = |S|$. Recall that $|C| + |K''| = mn - (n - 1) - 2D$, and that $\Psi$ is bijective. Suppose that $|P| = \lambda \sigma^{n-1}$. Then

$$|W| \leq \lambda \sigma^{n-1} \sigma^{mn-n+1-2D} \cdot \sigma^{2D} + (1 - \lambda) \sigma^{n-1} \sigma^{mn-n+1-2D} \cdot (1 - \varepsilon) \sigma^{2D}. \quad (33)$$

It follows that

$$1 - \varepsilon/q \leq \lambda + (1 - \lambda)(1 - \varepsilon),$$

which implies that $\lambda \geq 1 - \frac{1}{q}$. $\blacksquare$

Now, for some subset $L \subseteq K'$ of size 2D we have

$$\{\xi_1, \eta_1, \ldots, \xi_D, \eta_D\} = \{y_i(j) \mid (i, j) \in L\}.$$

Put $K'' = K' \setminus L$. Recall that $W \subseteq N^{(m)}$ satisfies $|W| \geq (1 - \varepsilon/6) |N^{(m)}|$. Let $\mathcal{P} \subseteq S^{(n-1)}$ be the set defined in Lemma 4.33 thus $|\mathcal{P}| \geq \frac{5}{6} |S|^{n-1}$.

By definition, for each $z \in \mathcal{P}$ there exist $u_z \in SC$, $v_z \in S^{K''}$ and $W_z \subseteq W$ with $|W_z| \geq (1 - \varepsilon) |S|^{2D}$ such that

$$W_z \Psi_{C \cup K'' \cup \mathcal{Y} \cup \mathcal{Z}} = \{(u_z, v_z, z)\}.$$

As $\Psi$ is a bijection this implies that $|W_z \Psi_L| = |W_z| \geq (1 - \varepsilon) |S|^{2D}$. 

53
Now let $x \in W_z$. Then

$$x\phi = (x\phi_1, \ldots, x\phi_n) = (x\phi_1, z)$$

and

$$x\phi_1 = U(\mathcal{X}, \mathcal{Y}, \mathcal{Z}).$$

In the expression (33) for $U(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$, each of the factors $R, A_i, B_i, C_i, D_i$ is a product of terms $t^{z_i}$ where $t$ is a component of $x\Psi \pi_{C_\cup K''\cup [n-1]} = (u_z, v_z, z)$. Therefore

$$x\phi_1 = \prod_{i=1}^{D} (a_i \xi_i b_i)^{-\alpha_i} (c_i \eta_i d_i)^{-\beta_i} (a_i \xi_i b_i)^{\gamma_i} (c_i \eta_i d_i)^{\delta_i} \cdot r$$

(34)

$$= \prod_{i=1}^{D} T_{\sigma_i, \tau_i} (\xi_i, \eta_i) \cdot r$$

where $a_i, b_i, c_i, d_i$ and $r$ depend only on $z$ and $\alpha_i, \beta_i, \gamma_i, \delta_i$ are certain automorphisms of $S$, independent of everything else, and

$$\xi_i = (a_i \xi_i b_i)^{-\alpha_i}, \eta_i = (c_i \eta_i d_i)^{-\beta_i}$$

$$\sigma_i = a_i^{-1} \gamma_i, \tau_i = \beta_i^{-1} \delta_i.$$

Now $(\xi_1, \eta_1, \ldots, \xi_D, \eta_D) = x\Psi \pi_L$ takes $|W_z x\Psi \pi_L| \geq (1 - \varepsilon)|S|^{2D}$ values as $x$ ranges over $W_z$; hence so does the tuple $(\xi_1, \eta_1, \ldots, \xi_D, \eta_D)$. According to Theorem 1.1.2 this implies that $\prod_{i=1}^{D} T_{\sigma_i, \tau_i} (\xi_i, \eta_i)$ takes at least $\lambda|S|$ values, where

$$\lambda = l(S)^{-3/5} \text{ if } l(S) \geq 3, \lambda = 1 \text{ if } l(S) = 2; \text{ therefore so does } x\phi_1, \text{ by (34).}$$

It follows that

$$|W\phi| \geq \sum_{x \in \mathcal{P}} |W_z x\phi_1| \geq \frac{5}{6} |S|^{n-1} \cdot \lambda|S| \geq l(S)^{-4/5}|N|$$

since $|S|^n = |N|$ and $(\frac{5}{6})^5 < 3$. This completes the proof of (1).

To prove (2), we replace $D$ by $D_1 = 5D$ in the above. Let $(u, v_0, z)$ be an arbitrary element of $S^C \times S^{K''} \times S^{(n-1)}$, and let $z \in S$. For each $\xi = (\xi_1, \eta_1, \ldots, \xi_D, \eta_D) \in S^L$ there exists $x \in N^{(m)}$ such that

$$x\Psi \pi_{C_\cup K''\cup [n-1]} = (u, v_0, z)$$

$$x\Psi \pi_L = \xi.$$ 

Take $u_z = u$ and $v_z = v_0$ in the above discussion. Then $x\phi_1$ is given by (34). Now Corollary 1.2 says that

$$S = \prod_{i=1}^{D_1} T_{\sigma_i, \tau_i} (S, S).$$

54
We may therefore choose $\xi$ so that
\[
\prod_{i=1}^{D_1} T_{\sigma_i, \tau_i}(\xi_i, \eta_i) = z r^{-1},
\]
and so ensure that $x \phi = (z, z)$. It follows that
\[
| (z, z)^{-1} | \geq | S^{SC} \times S^{K''} | = | S |^{mn-(n-1)-2D_1} \\
> | N |^{-2D_1/n} | N |^{m-1}.
\]

5 Applications

5.1 Subgroups of finite index

Here we re-prove the main result of [NS]:

**Theorem 5.1** If $G$ is a finitely generated profinite group then every subgroup of finite index in $G$ is open.

**Proof.** Let $H$ be a subgroup of finite index in $G$. Then $H$ contains a normal subgroup $N$ of finite index in $G$. The closure $M = \overline{N}$ of $N$ is open in $G$, so $M$ is again a finitely generated profinite group. If $N = M$ then $N$ is open and so $H$ is open.

Suppose that $N < M$. Then Corollary [LS] shows that at least one of
\[
NM' < M, \\
NM_0 < M
\]
holds. Put $q = |M/N|$, so we have $M^q \leq N$. Note that $M'$ is closed, by Theorem [L6]

Now $M/M'M^q$ is a finitely generated abelian profinite group of finite exponent, so it is finite, hence discrete; as $NM'/M'M^q$ is a dense subgroup it follows that $NM' = M$.

To derive a contradiction it remains to show that $NM_0 = M$; to this end we may as well replace $G$ by $G/M_0$, and so assume that $M_0 = 1$. Then $M$ has a closed semisimple normal subgroup $T$ such that $M/T$ is soluble. It follows from the preceding paragraph that $NT = M$.

A theorem of Martinez-Zelmanov [MZ] and Saxl-Wilson [SW] shows that $T^q$ is closed in $T$ (because the word $x^q$ has bounded width in all finite simple groups). As $T^q \leq N$ we may factor it out and assume further that $T^q = 1$. Now the definition of $M_0$ ensures that in fact $T$ is a product of finite simple groups each of which is normal in $M$; and these simple groups have bounded orders [J]. Therefore $M/C_M(T)$ is finite, and so $T$ is finite. Hence $N \cap T$ is closed. Thus
\[
T = [T, M] = [T, \overline{N}] \leq [T, N] \leq T \cap N
\]
whence $M = NT = N$, as required. \[\square\]
5.2 Finite to profinite

Here we recall some standard compactness arguments. We refer to subsection 1.2.1 for the statements of the following theorems, concerning a finitely generated profinite group $G$ with closed normal subgroups $K$ and $H$.

Proof of Theorem 1.5. Write $I = \{(i,j) \mid 1 \leq i \leq r, 1 \leq j \leq f_0\}$. For each open normal subgroup $N$ of $G$ let

$$X(N) = \{x = (x_{ij}) \in K^{(r_f)} \mid G = N \langle y_i^{x_{ij}} \mid (i,j) \in I \rangle \}.$$

Theorem 1.1, applied to the finite group $G/N$, shows that each set $X(N)$ is non-empty. Also $X(N)$ is closed in $K^{(r_f)}$, being a union of cosets of $(N \cap K)^{(r_f)}$, and if $N > M$ then $X(N) \supseteq X(M)$. It follows by compactness that $\bigcap_N X(N)$ is non-empty, taking the intersection over all open normal subgroups $N$ of $G$. Let $x$ be in this intersection. Then

$$G = \bigcap_N N \langle y_i^{x_{ij}} \mid (i,j) \in I \rangle = \langle y_i^{x_{ij}} \mid (i,j) \in I \rangle.$$

Proof of Theorems 1.6 and 1.7. Let $R$ denote the right-hand side of equation (†) or equation (‡) (see Subsection 1.2.1). Then $R$ is a closed subset of $G$. Now let $N$ be an open normal subgroup of $G$. Then Theorem 1.2, respectively Theorem 1.3, applied to the finite group $G/N$ shows that $[H,G]N = RN$. As $R$ is closed, intersecting over all open normal subgroups $N$ of $G$ we get

$$R = \bigcap_N RN \supseteq [H,G],$$

and the results follow since $R \subseteq [H,G]$.

5.3 Verbal subgroups

Here we show how the main results of [NSP] may be quickly derived from Theorems 1.1 and 1.2.

Let $w$ be a group word in $k$ variables, and $G$ a group. The corresponding verbal subgroup is $w(G) = \langle G_w \rangle$, where

$$G_w = \{w(g)^{\pm 1} \mid g \in G^{(k)} \}$$

denotes the (symmetrized) set of $w$-values in $G$. We say that $w$ has width $m$ in $G$ if

$$w(G) = G^{*m},$$

if this holds for some finite $m$ we denote the least such $m$ by $m_w(G)$, and say that $w$ has finite width in $G$.

The following elementary result is Proposition 2.1.2 of [S2]:

56
Lemma 5.2 If \( G \) is abelian-by-finite then \( m_w(G) \) is finite.

Now define
\[
\beta(w,G) = |G : w(G)|.
\]

Let us call the word \( w \) \( d \)-bounded if there exists \( \beta_w = \beta_w(d) \in \mathbb{N} \) such that \( \beta(w, G) \leq \beta_w(d) \) whenever \( G \) is a \( d \)-generator finite group. The positive solution of the Restricted Burnside Problem [Z] asserts that the word \( w = x^q \) is \( d \)-bounded, for all natural numbers \( d \) and \( q \). This implies that every non-commutator word \( w \) is \( d \)-bounded, since for any group \( G \) we have \( w(G) \geq G^q \) where \( q = |\mathbb{Z}/w(\mathbb{Z})| \). (In fact it is easy to see that, conversely, every \( d \)-bounded word is a non-commutator word.)

Proposition 5.3 Suppose that \( w \) is \( d \)-bounded (for some \( d \geq 1 \)). Then there exists \( m_0 = m_0(w) \) such that \( m_w \) with \( m_0 \) in every finite semisimple group.

Proof. It suffices to prove this for a simple group \( G \). Take \( q = |\mathbb{Z}/w(\mathbb{Z})| \). We consider three cases.

(i) Where \( w(G) = 1 \). Then \( m_w(G) = 1 \).

(ii) Where \( w(G) \neq 1 \) but \( G^q = 1 \). There are only finitely many possibilities for \( G \) in this case [J]. Since \( G_w \) generates \( G \) it follows that \( G = G_w^n \) where \( n = n(q) \) is the maximal order of any such group \( G \).

(iii) Where \( G^q \neq 1 \). In this case, the theorem of Martinez-Zelmanov [MZ] and Saxl-Wilson [SW] shows that every element of \( G \) is a product of \( h(q) \) \( q \)-th powers; as each \( q \)-th power is a \( w \)-value it follows that \( m_w(G) \leq h(q) \).

The main result is now

Theorem 5.4 Let \( w \) be a \( d \)-bounded word and \( G \) a finite \( d \)-generator group. Then \( m_w(G) \leq f(w,d) \) where \( f(w,d) \) depends only on \( w \) and \( d \).

Proof. Let \( \mathcal{M} \) denote the (finite) set of (non-abelian) simple groups \( M \) such that \( w(M) = 1 \). For \( n \in \mathbb{N} \) set
\[
\mu(n) = |F_n : K(w)|
\]
where \( F_n \) is free of rank \( n \) and \( K(w) \) is the intersection of all \( \ker \theta \) where \( \theta \) ranges over homomorphisms \( F_n \to \text{Aut}(M) \) with \( M \in \mathcal{M} \).

Put \( W = w(G) \) and set \( \beta = |G : W| \), so that \( \beta \leq \beta_w(d) \). By Schreier’s formula we then have
\[
d(W) \leq d_0 := 1 + \beta(d-1).
\]

Set
\[
H = \bigcap C_W(M)
\]
where \( M \) ranges over all chief factors of \( W \) that belong to \( \mathcal{M} \). Then \( W/H \) is an image of \( F_{d_0}/K(w) \), so \( |W : H| \leq \mu(d_0) \) and \( |G : H| \leq \beta_1 := \beta \mu(d_0) \). It follows that \( d(H) \leq d_1 := 1 + \beta_1(d-1) \).
Now let $K$ be the intersection of the kernels of all homomorphisms $F_d \to \text{Sym}(\beta_1)$. Lemma 2.5 shows that $w$ has finite width $m_1 = m_1(d, \beta)$ in the group $F_d/K'$. As $G/H'$ is an image of $F_d/K'$ it follows that $m_w(G/H') \leq m_1$, so

$$W = H' \cdot G_w^{*m_1}. \quad (35)$$

Put $X_1 = G_w^{*m_1}$, so $W = H'X_1$. Then $H = H'X_2$ where $X_2 = H \cap X_1$. There exist $Y_0 \subseteq X_1$ and $Y_1 \subseteq X_2$ such that

$$W = H' \langle Y_0 \rangle, \quad |Y_0| \leq d_0,$$

$$H = H' \langle Y_1 \rangle, \quad |Y_1| \leq d_1.$$

Now recall Theorem 1.1: this associates to $W$ a characteristic subgroup $W_0$, contained in $H$, such that $W^{(3)}W_0/W_0$ is semisimple. Put $D = W^{(3)}W_0 \cap H$. Applying Theorem 1.1 to the soluble group $H/D$, we find a set $Y_2 \subseteq Y_1^H$ such that

$$H = D \langle Y_2 \rangle, \quad |Y_2| \leq h_1 = f_0(d_1, d_1).$$

By the definition of $H$, the semisimple group $D/W_0$ is a product of simple groups $S$ such that $w(S) = S$. Hence by Proposition 5.3 we have $D = W_0D_w^{*m_0}$. Using Lemma 2.5 we may therefore lift $Y_2$ to a set $Y_3 \subseteq D_w^{*m_0}Y_2$ so that

$$H = W_0 \langle Y_3 \rangle, \quad |Y_3| \leq h_1.$$

Now put $Y_4 = Y_3 \cup Y_0$. Then $W = W' \langle Y_4 \rangle = W_0 \langle Y_4 \rangle$ and $|Y_4| \leq h_1 + d_0$; and $Y_4 \subseteq G_w^{*m_1+m_0}$.

A further application of Theorem 1.1 now provides a set $Y$ such that $W = \langle Y \rangle$, $|Y| \leq h_2 := (h_1 + d_0)f_0(h_1 + d_0, d_0)$, and each element of $Y$ is conjugate to one of $Y_4$.

Put $Y = Y \cup Y^{-1}$. It then follows by Theorem 1.2 that

$$W' = \left( \prod_{y \in Y} [W, y] \right)^{*f_1(2h_2, d_0)} \subseteq G_w^{*m_2},$$

where $m_2 = 4h_2f_1(2h_2, d_0)(m_1 + m_0)$. With (35) this shows that $w$ has width $f(w, d) := m_2 + m_1$ in $G$. $\blacksquare$

Now let $G$ be a $d$-generator profinite group. Suppose that $m_w(Q) \leq m < \infty$ for every continuous finite quotient $Q$ of $G$. Then

$$w(G)N = G_w^{*m} \cdot N$$

for every open normal subgroup $N$ of $G$. But $G_w^{*m}$ is a closed subset of $G$, because $w : G^{(k)} \to G$ is continuous; therefore

$$w(G) \subseteq \bigcap_N G_w^{*m} \cdot N = G_w^{*m},$$

58
so \( w(G) = G^\text{w} \) is a closed subgroup of \( G \).

If also \( w \) is \( d \)-bounded, then \( \beta(w, Q) \leq \beta_w(d) = \beta \), say, for every continuous finite quotient \( Q \) of \( G \). Thus

\[
|G : w(G)N| \leq \beta
\]

for each open normal subgroup \( N \) of \( G \). Choosing an open normal subgroup \( M \) for which \( |G : w(G)M| \) is maximal we infer (given that \( w(G) \) is closed) that

\[
w(G) = \bigcap_N w(G)N = w(G)M.
\]

Thus \( w(G) \) is an open subgroup of \( G \).

Theorem 5.4 now gives

**Theorem 5.5** Let \( G \) be a \( d \)-generator profinite group and \( w \) a \( d \)-bounded word. Then the verbal subgroup \( w(G) \) is open in \( G \).

This shows that \( w(G) \) is open whenever \( G \) is a finitely generated profinite group and \( w \) is any non-commutator word, the main result of [NSP]. The point of our clumsier formulation is that the theorem as stated is independent of the Restricted Burnside Problem.

### 5.4 Verbal subgroups in compact groups

Throughout this subsection, we suppose that \( G \) is a compact group and that the profinite quotient \( G/G^0 \) is finitely generated. Several of the preceding results can be generalized.

**Corollary 5.6** If \( w \) is a non-commutator word then \( w(G) \) is open in \( G \).

**Proof.** Set \( q = |\mathbb{Z}/w(\mathbb{Z})| \). Every element of \( G^0 \) is a \( q \)-th power ([HM], Theorem 9.35), so \( G^0 \leq w(G) \), and so \( w(G)/G^0 \). The result follows by the remark following Theorem 5.5.

The next corollary follows likewise from Theorem 5.1.

**Corollary 5.7** Every subgroup of finite index in \( G \) is open.

**Lemma 5.8** Suppose that \( A \leq Z(G^0) \) is closed and normal in \( G \). Then \([A,G]\) is closed in \( G \).

**Proof.** Let us consider \( A \) as an additively-written \( \Gamma = G/G^0 \)-module. By hypothesis, \( \Gamma \) has a dense finitely generated (abstract) subgroup \( X = \langle x_1, \ldots, x_d \rangle \). Now

\[
[A,X] = A(x_1-1) + \cdots + A(x_d - 1);
\]

this is an \( X \)-submodule of \( A \), and it is closed in \( A \) because \( A \) is compact. Therefore \([A,X]\) is a \( \Gamma \)-submodule, because \( X \) is dense in \( \Gamma \). Therefore \( C := C_\Gamma(A/[A,X]) \) is closed in \( \Gamma \), and as \( X \leq C \) it follows that \( C = \Gamma \). Hence \([A,G] = [A,\Gamma] = [A,X] \) is closed.
Corollary 5.9 The derived group $G'$ is closed in $G$.

Proof. Let $P = (G^0)'$ denote the derived group of $G^0$. Then $P$ is closed, by [HM], Theorem 9.2. So replacing $G$ by $G/P$ we may suppose that $G^0$ is abelian. Then $[G^0, G]$ is closed by the preceding Lemma, so we may factor it out and reduce to the case where $G^0$ is central in $G$. Now according to [HM], theorem 9.41, we have $G = G^0 D$ for some closed profinite subgroup $D$. Since $D/(D \cap G^0) \cong G/G^0$ is finitely generated, $D = (D \cap G^0) H$ for some finitely generated profinite group $H$. Then $G' = H'$ is closed by the remark following Theorem 1.6.

Remark. More generally, we can show that $[H, G]$ is closed for every closed normal subgroup $H$ of $G$. When $G^0 = 1$ this follows from Theorem 1.6 and when $G = G^0$ it follows from the known structure of connected compact groups. The general case depends on a modified form of the ‘Key Theorem’, Theorem 3.10, in which $d = d(G)$ is replaced by $d(G/C_G(H))$; the proof will appear elsewhere.

5.5 Quotients of semisimple compact groups

In this subsection we consider a topological group

\begin{equation}
G = \prod_{i \in I} S_i,
\end{equation}

where $I$ is an index set, and either:

(a) each $S_i$ is a nonabelian finite simple group, and for each $n$ the set

$$I(n) = \{i \in I \mid |S_i| \leq n\}$$

is finite; or

(b) each $S_i$ is a compact connected simple Lie group

(here, by a ‘simple Lie group’ we mean the analogue of a quasisimple finite group: i.e. it may have a non-trivial centre, but is simple modulo the centre and perfect).

Remarks: i. (a) holds in particular when $G$ is a semisimple finitely generated profinite group.

ii. Hofmann and Morris [HM] call a compact connected group $G$ ‘semisimple’ if it is perfect, i.e. if $G = G'$, equivalently if $G = G'$ (loc. cit. Theorem 9.2). However, this holds if and only if $G = \tilde{G}/C$ where $\tilde{G}$ is a product of compact connected simply-connected simple Lie groups and $C$ is a totally disconnected normal subgroup (loc. cit. Theorem 9.19); thus any quotient of $G$ is also a quotient of a group of the form (36).
Theorem 5.10 Let $Q$ be an infinite quotient of (the underlying abstract group) $G$. Then $|Q| \geq 2^\aleph_0$.

This depends on the following technical device:

Proposition 5.11 Let $L$ be (a) a nonabelian finite simple group or (b) a compact connected simple Lie group. In Case (b), let $T$ be a maximal torus of $L$, in Case (a) let $T = L$. There is a function $\lambda = \lambda_L : T \to [0, 1]$ with the following properties:

(i) $\lambda(s) = 0 \iff s \in Z(L)$;

(ii) $\lambda(s^{-1}) = \lambda(s^t) = \lambda(s)$ and $\lambda(st) \leq \lambda(s) + \lambda(t)$ for all $s, t \in T$;

(iii) if $t \in T$ and $\lambda(t) \geq \varepsilon > 0$ then

$$L = (t^L \cup t^{-L})^{f(\varepsilon)}$$

where $f(\varepsilon) \in \mathbb{N}$ depends only on $\varepsilon$;

(iv) in Case (a), $1 \neq s \in L$ implies $\lambda_L(s) \geq \varepsilon(r)$ where $\varepsilon(r) > 0$ depends only on $r = \text{rank}(L)$;

(va) in Case (a): given $\beta, \varepsilon \in (0, 1)$, there exists $s \in L$ with

$$|\lambda_L(s) - \beta| < \varepsilon,$$

provided that $\text{rank}(L) \geq n(\varepsilon)$, where $n(\varepsilon)$ depends only on $\varepsilon$;

(vb) in Case (b): for each $\beta \in [0, 1]$ there exists $s \in T$ with $\lambda_L(s) = \beta$.

Recall that $\text{rank}(L)$ means the (untwisted) Lie rank of $L$ if $L$ is of Lie type, $n$ if $L \cong \text{Alt}(n)$, and 0 otherwise. The proof is postponed to the following subsections.

Given an ultrafilter $\mathcal{U}$ on $I$, one defines the ultralimit of a bounded family $(a_i)_{i \in I}$ of real numbers to be the unique number $\alpha = \lim_{\mathcal{U}} a_i$ such that

$$\varepsilon > 0 \implies \{i \in I \mid |a_i - \alpha| < \varepsilon\} \in \mathcal{U}$$

(cf. [KL], Section 3.1). We remark that if $\mathcal{U}$ is the principal ultrafilter $\mathcal{U}(j)$ over some element $j \in I$, then $\lim_{\mathcal{U}} a_i = a_j$.

In Case (a), set $T_i = S_i$ for each $i$; in Case (b), we choose a maximal torus $T_i$ in $S_i$. In either case, let $G_\bullet = \prod_{i \in I} T_i$. Now define a function $h_\mathcal{U} : G_\bullet \to [0, 1]$ by

$$h_\mathcal{U}(g) = \lim_{\mathcal{U}} \lambda_{S_i}(g_i) \text{ for } g = (g_i).$$

The analogue of property (ii) obviously holds for the function $h_\mathcal{U}$. This implies that the set

$$K_\mathcal{U} := h_\mathcal{U}^{-1}(0)$$

is a normal subgroup of $G_\bullet$, and that $h_\mathcal{U}$ is constant on the cosets of $K_\mathcal{U}$.
For a subset $J$ of $I$ we set
\[ N(J) = \prod_{i \in J} Z_j \times \prod_{i \in I \setminus J} S_i, \]
the kernel of the projection $G \to \prod_{j \in J} S_j / Z_j$, where $Z_j = Z(S_j)$. Each $N(J)$ is a closed normal subgroup of $G$.

Now we can prove Theorem 5.10. Let $Q = G/H$ where $H$ is a normal subgroup of infinite index in $G$. Suppose we are in Case (b) (Lie groups); if $H \leq N(j)$ then $Q$ maps onto $S_j / Z_j$ and the result is clear. Suppose we are in Case (a), and let $J$ be the set of indices $j$ such that $H \leq N(j)$. Then $|Q| = |G/N(j)| / |G_1/H_1|$ where $G_1 = \prod_{j \in I \setminus J} S_j$ and $H_1$ denotes the projection of $H$ into $G_1$. If $J$ is infinite, then $G/N(J)$ is an infinite profinite group and again the result is clear. If $J$ is finite, then $H_1$ has infinite index in $G_1$, and we can replace $G$ by $G_1$.

Thus in any case, we may assume that $H \neq N(j)$ for every $j \in I$. We shall show that in this case,
\[ (*) \quad \text{There exists a non-principal ultrafilter } \mathcal{U} \text{ on } I \text{ such that } H_\bullet := H \cap G_\bullet \leq K_\mathcal{U}; \]
\[ (**) \quad |G_\bullet / K_\mathcal{U}| \geq 2^{\aleph_0}. \]
(Recall that $G_\bullet = G$ in Case (a).)

**Proof of (**)**.

**Case 1.** The $S_i$ are finite simple groups, and for some $m \in \mathbb{N}$, the set
\[ D = D(m) = \{ i \mid \rank(S_i) \leq m \} \]
belongs to $\mathcal{U}$. Then $N(D) \leq K_\mathcal{U}$, so $G/K_\mathcal{U} \cong G_1/K_\mathcal{U}$ where $G_1 = \prod_{i \in D} S_i$ and $\mathcal{U}_1$ is the restriction of $\mathcal{U}$ to $D$. Now property (iv) of the functions $\lambda_{S_i}$ implies that $g \in K_\mathcal{U}$ precisely when the set $\{ i \in D \mid g_i = 1 \}$ belongs to $\mathcal{U}_1$. Therefore the quotient $G_1/K_\mathcal{U}$ coincides with the ultraproduct $\prod_{i \in D} S_i / \mathcal{U}_1$. But an ultraproduct of finite sets is either finite or has cardinality at least $2^{\aleph_0}$ ([FMS], Theorem 1.31). The first possibility is excluded since each of the sets $I(n)$ is finite, hence cannot belong to $\mathcal{U}_1$, and (**) follows.

**Case 2.** The $S_i$ are finite simple groups and $D(m) \notin \mathcal{U}$ for each $m \in \mathbb{N}$. Let $\beta \in (0,1)$. For each $i \in I$ we choose $g_i \in S_i$ so as to minimize
\[ |\lambda_{S_i}(g_i) - \beta| = \varepsilon_i, \]
say. Property (va) ensures that for any $\varepsilon > 0$, we have $\varepsilon_i < \varepsilon$ whenever $\rank(S_i) \geq n(\varepsilon)$. We claim that $h_\mathcal{U}(g) = \beta$. Indeed, suppose that $h_\mathcal{U}(g) = \beta' \neq \beta$, and put $\varepsilon = |\beta' - \beta|$. Then
\[ |\lambda_{S_i}(g_i) - \beta'| < \varepsilon/2 \implies \varepsilon_i = |\lambda_{S_i}(g_i) - \beta| > \varepsilon/2 \implies i \in D(m) \]
where \( m = n(\varepsilon/2) \); thus \( D(m) \) contains a member of \( U \) and so \( D(m) \in U \), a contradiction.

It follows that \( h_U(G) = [0, 1] \). Since \( h_U \) is constant on cosets of \( K_U \) this now implies that \( G/K_U \) has the cardinality of \([0, 1]\), and (**) follows.

**Case 3.** The \( S_i \) are connected simple Lie groups. Let \( \beta \in (0, 1) \). Using Property (vb), choose \( g_i \in T_i \) with \( \lambda_{S_i}(g_i) = \beta \) for each \( i \). Then \( g = (g_i) \in G \), and \( h_U(g) = \beta \); and (**) follows as in the preceding case.

**Proof of (\ast).**

\( H \) is a normal subgroup of infinite index in \( G \), and \( H \subseteq N(j) \) for any \( j \in I \).

For \( t = (t_i)_i \in H \) and \( \epsilon > 0 \) put

\[ A(t, \epsilon) = \{ i \in I \mid \lambda_{S_i}(t_i) < \epsilon \} , \]

and let \( U \) be the collection of all subsets \( A(t, \epsilon) \) with \( t \in H \) and \( \epsilon > 0 \).

We claim that every finite subset of \( U \) has nonempty intersection. Indeed, suppose that

\[ A(t_1, \epsilon_1) \cap A(t_2, \epsilon_2) \cap \ldots \cap A(t_k, \epsilon_k) = \emptyset . \]

Put \( \epsilon = \min_i \{ \epsilon_i \} \) and suppose that \( t_i = (t_{i,j})_j \) with \( t_{i,j} \in T_j \).

Then for each index \( j \in I \) there is some \( i \leq k \) such that \( j \not\in A(t_i, \epsilon) \), so \( \lambda_{S_j}(t_{i,j}) \geq \epsilon \). Now (iii) gives

\[ S_j = \left( t_{i,j}^{S_j} \cup t_{i,j}^{-S_j} \right)^* n , \]

where \( n = f(\epsilon) \). Considering independently each coordinate \( j \in I \) we see that

\[ G = \prod_{i=1}^k \left( t_i^G \cup t_i^{-G} \right)^* n \subseteq H , \]
a contradiction.

On the other hand, the intersection of the collection \( U \) is empty. Let \( T_j^* \) denote the projection of \( H \) into \( S_j \). If \( j \) belongs to every member of \( U \) then \( \lambda_{S_j}(t) = 0 \) for every \( t \in T_j^* \), whence \( T_j^* \leq Z(S_j) \) by property (i). Since the conjugates of \( T_j^* \) generate the projection of \( H \) into \( S_j \), this implies that \( H \leq N(j) \), contrary to hypothesis.

Now a standard application of Zorn’s lemma establishes the existence of a non-principal ultrafilter \( U \) on \( I \) containing \( U \). From the definition of \( U \) it follows that \( h_U(t) = 0 \) for all \( t \in H \), and (\ast) follows.

**5.5.1 The profinite case**

In Case (a) we can say rather more:

**Theorem 5.12** Suppose that \( G = \prod_{i \in I} S_i \) where each \( S_i \) is a finite (non-abelian) simple group and \( \{ i \in I \mid |S_i| \leq n \} \) is finite for each \( n \). Then
• every proper normal subgroup of \( G \) is contained in a maximal one;

• the maximal proper normal subgroups of \( G \) are precisely the subsets \( K_U \)
  for ultrafilters \( U \) on \( I \);

• the normal subgroup \( K_U \) is closed in \( G \) if and only if \( U \) is principal.

**Proof.** If \( U = U(j) \) is principal then \( K_U = N(j) \) is a closed maximal normal
subgroup. If \( U \) is non-principal, then \( K_U \) has infinite index in \( G \), by (**). We
claim that in this case too, \( K_U \) is a maximal normal subgroup. Suppose that
\( g = (g_i)_i \in G \) is not in \( K = K_U \). This means that \( h_U(g) > 0 \), which in turn
implies that for some \( \alpha > 0 \) the set
\[
A = \{ i \in I \mid \lambda_S(g_i) > \alpha \}
\]
belongs to \( U \).

Now if \( i \in A \), we see from (iii) in Proposition 5.11 that
\[
S_i = (g_i^{S_i} \cup g_i^{-S_i})^n
\]
where \( n = f(\alpha) \). It follows that
\[
G = N(A) \cdot (g^{G} \cup g^{-G})^n.
\]
As \( U \) is a filter and \( A \in U \) it is easy to see that \( N(A) \leq K \), and so
\[
G = K (g^{G} \cup g^{-G})^n \subseteq K (g^{G} \cup g^{-G}).
\]
Since \( g \) was an arbitrary element of \( G \setminus K \) it follows that \( G/K \) is simple.

Now suppose that \( H \) is any proper normal subgroup of \( G \). Then either
\( H \leq N(j) = K_U(j) \) for some \( j \in I \), or (*) provides a non-principal ultrafilter \( U \)
such that \( H \leq K_U \).

It remains only to observe that if \( U \) is a non-principal ultrafilter then \( K_U \)
contains the restricted direct product of the \( S_i \), which is dense in \( G \), and so \( K_U \)
cannot be closed. \( \blacksquare \)

### 5.5.2 The connected case: automorphisms

The material in this subsection will only be needed for the proof of Theorem
5.26 in Subsection 5.7. We consider \( G = \prod_{i \in I} S_i \) where \( I \) is an infinite set and
each \( S_i \) is a compact connected simple Lie group. In this case, our functions \( h_U \)
were only defined on \( G_\bullet = \prod_{i \in I} T_i \), which depends on a choice of maximal torus
\( T_i \) in each \( S_i \). Suppose that in each \( S_i \) we choose maximal tori \( T^{(l)}_i \), \( l = 1, \ldots, d \).
Let \( \lambda^{(l)}_S : T^{(l)}_i \rightarrow [0, 1] \) be as in Proposition 5.11, put \( T^{(l)} = \prod_{i \in I} T^{(l)}_i \), and
define \( h^{(l)}_U : T^{(l)} \rightarrow [0, 1] \) and \( K^{(l)}_U := h^{−1}_U(0) \leq T^{(l)} \) as before, using the maps
\( \lambda^{(l)}_S \). A subgroup of the form \( T^{(l)} \) will be called a ‘maximal pro-torus’ of \( G \) (cf.
[HM]). We will write \( \lambda_i \) for \( \lambda_{S_i} \) where the meaning is clear.
Lemma 5.13 Let $H$ be a proper normal subgroup of $G$ with $H \not \leq N(j)$ for all $j \in I$. Then there exists a non-principal ultrafilter $\mathcal{U}$ on $I$ such that $H_{(i)} := H \cap T^{(l)} \leq K^{(l)}_{(i)}$ for $l = 1, \ldots, d$.

Proof. For $t \in H_{(i)}$ and $\epsilon > 0$ define $A^{(l)}(t, \epsilon)$ as in the proof of (*), above, using $\lambda$ in place of $\lambda$. Let $U^{(l)}$ be the collection of all subsets $A^{(l)}(t, \epsilon)$ with $t \in H_{(i)}$ and $\epsilon > 0$. As above, it will suffice to show that every finite subcollection of $U^{(1)} \cup \ldots \cup U^{(d)}$ has non-empty intersection. Arguing as before, we see that if

$$\bigcap_{l=1}^{d} \left( A^{(l)}(t_1^{(l)}) \cap A^{(l)}(t_2^{(l)}) \cap \ldots \cap A^{(l)}(t_k^{(l)}) \right) = \emptyset,$$

then for each $j \in I$ there exist $l \leq d$ and $i \leq k$ such that $\lambda_{S_j}(t_i^{(l)}) \geq \epsilon$ where $\epsilon = \min \epsilon_{i,j}$. As before this yields the contradiction

$$G = \bigcap_{l=1}^{d} \prod_{i=1}^{k} \left( t_i^{(l)G} \cup t_i^{(l)-G} \right)^{\ast n} \subseteq H.$$

Now let $y$ be a continuous automorphism of $G$. The action of $y$ induces a permutation $y^i$ on the index set $I$, so that $S_i^y = S_{iy}$ for each $i$. Let $\mathcal{C}$ denote the set of orbits of $\langle y \rangle$ on $I$, and for each $J \in \mathcal{C}$ pick $i(J) \in J$. Then

$$\prod_{i \in J} S_i = \begin{cases} \prod_{n \in \mathbb{Z}} S_i^{y^n} & (J \text{ infinite}) \\ \prod_{n=0}^{e-1} S_i^{y^n} & (|J| = e < \infty) \end{cases}$$

where $S = S_{i(J)}$. Choose a maximal torus $T_{i(J)}$ in $S_{i(J)}$, and for $i \in \{i(J)\}^{\mathbb{Z}_+}$ set $T_i = T_{i(J)}^{y^n}$. Thus $T = \prod_{i \in J} T_i$ becomes a maximal pro-torus in $G$, and $T$ is ‘almost’ $y$-invariant, in the following sense. For each $J \in \mathcal{C}$ with $|J| < \infty$ put $l(J) = i(J)\mathbb{Z}$, and set

$$Z = \{1(J) \mid J \in \mathcal{C}, \ : \ |J| < \infty \},$$

$$\mathcal{T}(Z) = \{ t = (t_i) \in T \mid t_i = 1 \ \forall i \in I \};$$

for $i \notin Z$ we may identify $S_{iy}$ with $S_i$ via the action of $y$, and then for $t = (t_i) \in \mathcal{T}(Z)$ we have

$$(t^y)_i = t_i \ \forall i \in I,$$  \hspace{1cm} (38)

so $\mathcal{T}(Z)^y \leq \mathcal{T}$.

Set $Z^c = I \setminus Z$. For $\alpha \in [0, 1]$ and $\epsilon > 0$ define

$$A(t, \alpha, \epsilon) = \{ i \in I \mid |\lambda_{S_i}(t_i) - \alpha| < \epsilon \}.$$  

Lemma 5.14 Let $\mathcal{U}$ be a non-principal ultrafilter on $I$ with $Z^c \in \mathcal{U}$, and put $\mathcal{U}' = \mathcal{U} |_{Z^c}$. Then

$$\mathcal{U}' = \{ A(t, 1/2, 1/4) \mid t \in \mathcal{T}(Z), \ h_{\mathcal{U}}(t) = 1/2 \}.$$
Proof. Let $V$ denote the family of sets on the right-hand side of the equation. Then $V \subseteq U'$ by the definition of $h_U(t)$.

Now suppose that $Y \subseteq Z^c$ and $Y \in U$. Choose $t_i \in T_i$ so that

$$
t_i = 1 \text{ for } i \in Z
\lambda_i(t_i) = 1/2 \text{ for } i \in Y
\lambda_i(t_i) = 1 \text{ for } i \notin Y \cup Z.
$$

Then $t = (t_i) \in T(Z)$ and $A(t, \frac{1}{2}, \epsilon) = Y$ for every $\epsilon \in (0, \frac{1}{2}]$, so $h_U(t) = \frac{1}{2}$. Therefore $Y \in V$. Thus $U' \subseteq V$. ■

Lemma 5.15 Suppose that $Z^c \in U$ and that $t^{-1}t^y \in K_U$ for all $t \in T(Z)$. Then $U^y = U$.

Proof. Let $X \in U$. Then $X \supseteq X \cap Z^c = A(t, \frac{1}{2}, \frac{1}{4})$ for some $t \in T(Z)$ with $h_U(t) = \frac{1}{2}$. Now

$$
h_U(t^y) = h_U(t, t^{-1}t^y) \leq h_U(t) + h_U(t^{-1}t^y) = h_U(t),
$$

$$
h_U(t) = h_U(t^{-1}) = h_U(t^{-1}t^y, t^{-y}) \leq h_U(t^{-1}t^y) + h_U(t^{-y}) = h_U(t^y),
$$

so $h_U(t^y) = \frac{1}{2}$. Now it follows from (bS) that

$$
A(t, 1/2, 1/4)^y = A(t^y, 1/2, 1/4) = B,
$$
say, and $B \in U$ since $h_U(t^y) = \frac{1}{2}$. Therefore $X^y \supseteq B \in U$ and so $X^y \in U$. Thus $U^y \subseteq U$, and the result follows since $U^y$ is an ultrafilter. ■

Lemma 5.16 If $U^y = U$ then $\text{fix}(y^y) \in U$.

Proof. Here $\text{fix}(y^y)$ denotes the set of fixed points of $y^y$. We can partition $I$ as

$$
I = A_1 \cup A_2 \cup A_3 \cup \text{fix}(y^y)
$$

where $A_i^y \cap A_i = \emptyset$ for $i = 1, 2, 3$. To see this, it suffices to partition each $\langle y^y \rangle$-orbit $J$ of length at least 2 into three pieces $J_i$ such that $J_i^y \cap J_i = \emptyset$.

Identifying $J$ with $\mathbb{Z}$ or with $(1, 2, \ldots, e)$ where $y^y$ takes $i$ to $i + 1 \pmod{e}$, let

$$
J_1 = 2\mathbb{Z}, \ J_2 = 2\mathbb{Z} + 1, \ J_3 = \emptyset \text{ if } |J| = \infty;
J_1 = 2\mathbb{Z} \cap J, \ J_2 = (2\mathbb{Z} + 1) \cap J, \ J_3 = \emptyset \text{ if } |J| \text{ is even};
J_1 = \{2, \ldots, 2n\}, \ J_2 = \{1, \ldots, 2n - 1\}, \ J_3 = \{2n + 1\} \text{ if } |J| = 2n + 1.
$$

Then set $A_i = \cup_{J_i \in C} J_i$ for $i = 1, 2, 3$.

If $U^y = U$ then $A_i \notin U$ for each $i$, since $\emptyset \notin U$. Therefore $A_i^y \in U$ for each $i$, whence

$$
\text{fix}(y^y) = A_1^y \cap A_2^y \cap A_3^y \in U.
$$

(We are grateful to Martin Kassabov for pointing us to this lemma, which suggested the possibility of Proposition 5.18 below.) ■
Lemma 5.17 Suppose that $Z \in \mathcal{U}$ and that $t^{-1}t^y \in K_\mathcal{U}$ for all $t \in T(Z)$. Then \( \text{fix}(y) \in \mathcal{U} \).

**Proof.** If $J$ is an orbit of $\langle y \rangle$ of length at least 2, choose $t_J \in T_i(J)$ with $\lambda_i(J)(t_J) = 1$. Then set

\[
t_i(J) = \begin{cases} t_J^n & \text{if } J \text{ is infinite}, \\
t_J^n & (0 \leq n \leq e - 2), \\
1 & (|J| = e < \infty)
\end{cases}
\]

and set $t_i = 1$ for each $i \in \text{fix}(y)$ (recall that $l(J) = i(J) y^{(e-1)}$). Then $t = (t_i) \in T(Z)$, and whenever $\infty > |J| \geq 2$ we have

\[
A(t^{-1}t^y) \subseteq K_\mathcal{U} \cap Z \subseteq \text{fix}(y) \subseteq \text{fix}(y).
\]

Proposition 5.18 Let $y_1, \ldots, y_d$ be continuous automorphisms of $G$ and let $H$ be a proper normal subgroup of $G$ with $[G, y_l] \subseteq H$ for each $l$. Suppose that $H \not\subseteq N(j)$ for all $j \in I$. Then there exists a non-principal ultrafilter $\mathcal{U}$ on $I$ such that

\[
\bigcap_{i=1}^d \text{fix}(y_l) \subseteq \mathcal{U}.
\]

Hence $\bigcap_{i=1}^d \text{fix}(y_l)$ is infinite.

**Proof.** For each $l$ choose a maximal pro-torus $T^{(l)}$ corresponding to $y_l$ as above, and apply Lemma 5.13 to find a non-principal ultrafilter $\mathcal{U}$ such that $H \cap T^{(l)} \not\subseteq K_\mathcal{U}$ for $l = 1, \ldots, d$. Now the last three lemmas show that $\text{fix}(y_l) \subseteq \mathcal{U}$ for each $l$, and the result follows.

5.5.3 Proposition 5.11 finite case

Now $L$ is a finite simple group. We define

\[
\lambda(s) = \frac{\log |s^L|}{\log |L|}.
\]

Properties (i) and (ii) are clear, and (iii) follows from Proposition 1.23. (iv) follows from Proposition 1.24.

It remains to establish property (v). Given $\beta, \varepsilon \in (0, 1)$, we have to show that provided rank($L$) is sufficiently large, there exists $g \in L$ such that

\[
\frac{\log |C_L(g)|}{\log |L|} \in (\alpha - \varepsilon, \alpha + \varepsilon)
\]

67
where $\alpha = 1 - \beta$. As we only need to consider groups of large rank, we may suppose that $L$ is either alternating or a classical group.

If $L = \text{Alt}(n)$, take $g$ to be an even cycle of length $t \sim \beta n$ in $\text{Alt}(n)$. Note that $|C_L(g)|$ is roughly $t \cdot \bar{t}/2$ where $\bar{t} \sim \alpha n$. By Stirling’s formula, $\log(n!) \sim n \log n$ and hence $\log(t \cdot \bar{t}/2) \sim \alpha \log(n!/2)$ as $n \to \infty$.

If $L$ is a simple classical group, consider the corresponding universal quasi-simple classical group $\bar{L}$ acting on its natural module $V$ over a finite field of size $q$ equipped with a bilinear form $f$ (symmetric, sesquilinear, alternating or just equal to 0 in case $L$ has type $\text{PSL}_n$). Note that $\dim(V) \to \infty$ as $\text{rank}(L) \to \infty$.

We have $L = \bar{L}/Z$ where $Z$ is the centre of $\bar{L}$; and if $g = \bar{g}Z \in L$ with $\bar{g} \in \bar{L}$ then

$$|g^L| \leq |\bar{g}^\bar{L}| \leq |Z| |g^L|.$$  

Since $Z$ has asymptotically negligible size compared to $L$ it is enough to find an element $\bar{g} \in \bar{L}$ with $\log |C_{\bar{L}}(\bar{g})| \sim \alpha \log |\bar{L}|$.

We can decompose $V$ as $V_0 \oplus V_1 \oplus V_2$ so that:

- $\dim V_0$ is about $\sqrt{\alpha} \dim V$, and $\dim V_1 = \dim V_2,$
- $V_1 \oplus V_2$ is orthogonal to $V_0$, and
- The form $f$ is nondegenerate on both $V_0$ and $V_1 \oplus V_2$ and is isotropic on $V_1$ and on $V_2$

Let $\bar{g} \in \bar{L}$ be equal to the identity on $V_0$ and act on each of $V_1$ and $V_2$ as a cyclic transformation without fixed vectors. In other words there is a vector $v_i \in V_i, (i = 1, 2)$ such that $v_i, \bar{g}v_i, \bar{g}^2 v_i, \ldots$ is a basis for $V_i$.

Now $C_{\bar{L}}(\bar{g})$ contains the classical group $H$ on $V_0$ preserving $f$, and by the choice of $\dim V_0$ we have $\log |H|/\log |\bar{L}| \sim (\dim V_0/\dim V)^2$ which tends to $\alpha$ as $\dim V \to \infty$.

On the other hand if $s \in \bar{L}$ commutes with $\bar{g}$ then $s$ must stabilize $V_0$, the fixed space of $\bar{g}$. Since $V_1$ and $V_2$ are cyclic modules for $\bar{g}$, the action of $s$ on $V_1$ and $V_2$ is determined by $s \cdot v_1$ and $s \cdot v_2$. Hence $s$ is completely known from its restriction to $V_0$ and from the two vectors $sv_1, sv_2 \in V$. Denote by $\text{Gl}(V_0)$ the subgroup of $\text{GL}(V_0)$ which preserves $f$. We have $|\text{Gl}(V_0)| \leq q |H|$.

Therefore

$$|H| \leq C_{\bar{L}}(\bar{g}) \leq |\text{Gl}(V_0)||V|^2 \leq q^{1 + 2 \dim V} |H|$$  

which gives

$$\log |C_{\bar{L}}(\bar{g})|/\log |\bar{L}| \sim \log |H|/\log |\bar{L}| \to \alpha$$

as $\dim V$ tends to infinity.
5.5.4 Proposition 5.11 connected case

We shall need some information about the tori and roots of compact simple Lie groups; see for example [Bu], Chapter 19, [HM], Chapter 6. By $S^1$ we shall denote the group of complex numbers of absolute value 1 under multiplication. It is a compact torus of dimension 1.

Let $L$ be a compact simple Lie group with centre $Z$ (possibly nontrivial). Let $T$ be a maximal torus of $L$ (this is unique up to conjugacy). Every element of $L$ is conjugate to an element of $T$. Let $\Phi$ be a set of roots with respect to $T$. We choose and fix a set of fundamental roots $\Pi = \{\beta_1, \ldots, \beta_r\}$; $r$ is the rank of $L$. Every root $\alpha \in \Phi$ corresponds to a character $T \to S^1$ which we will also denote by $\alpha$. We have $\bigcap_{i=1}^{r} \ker \beta_i = Z$.

There is also a cocharacter $h_\alpha : S^1 \to T$ such that $\alpha(h_\alpha(\mu)) = \mu^2$ for all $\mu \in S^1$. For every pair $(\pm \alpha)$ of opposite roots of $\Phi$ there is a homomorphism $f_\alpha : SU(2) \to L$ such that $h_\alpha$ is the restriction of $f_\alpha$ to the diagonal subgroup $\text{diag}(\mu, \mu^{-1})$ of $SU(2)$ (and $h_{-\alpha} = h_\alpha^{-1}$). Let $S_\alpha = S_{-\alpha}$ be the image of $SU(2)$ in $L$ under $f_\alpha$. Then $S_\alpha$ is either $SU(2)$ or $PSU(2) \cong SO(3)$. Moreover $S_\alpha$ commutes elementwise with the closed subgroup $T_\alpha := \{g \in T \mid \alpha(g) = 1\}$ of $T$, and the central product $S_\alpha T_\alpha$ contains $T$.

Now we have to define $\lambda : T \to [0, 1]$ so that properties (i) – (iii) and (v) of Proposition 5.11 hold.

We can write a complex number $\mu \in S^1$ in a unique way as $\mu = e^{i\theta}$ with $\theta \in (-\pi, \pi]$. Set $l(\mu) := |\theta|$. We shall refer to $l(\mu)$ as the angle of $\mu$. 

**Definition.** For an element $g \in T$ define

$$\lambda(g) = \frac{1}{\pi r} \sum_{i=1}^{r} l(\beta_i(g))$$

Clearly $\lambda(g)$ is the same as $\lambda(\bar{g})$ for $\bar{g} = gZ$, if $\lambda$ is defined taken with respect to the torus $T/Z$ of $L/Z$.

It is also clear that (i) $\lambda(g) = 0$ if and only if $g \in Z$, and (ii) $\lambda(h_1) = \lambda(h_1^{-1})$ and $\lambda(h_1 h_2) \leq \lambda(h_1) + \lambda(h_2)$ for any $h_1, h_2 \in T$. Since $l(\mu)$ takes all values in $[0, \pi]$ and $T$ is a torus, we see that $\lambda(T) = [0, 1]$, which is property (v).

**Example:** If $L = SU(2)$ and $g$ is an element of the diagonal subgroup of $L$ with eigenvalues $\mu$ and $\mu^{-1}$ then $l(g)$ is the angle of $\mu^2$. From here and the isomorphism $PSU(2) \cong SO(3)$ we see that if $g \in SU(2)$ then $\lambda(g)$ is $|\theta|/\pi$ where $\theta$ is the angle of the image $\bar{g} \in PSU(2) = SO(3)$ considered as a rotation of $\mathbb{R}^3$.

Property (iii) follows from

**Lemma 5.19** There is an absolute constant $C > 0$ such that if $g \in T$ and $C/(\lambda(g))^2 < M \in \mathbb{N}$ then $K^*M = L$, where $K = g^L \cup g^{-L}$.
Lemma 5.20  If $L = \text{SU}(2)$ and $g \in L$ with $\lambda(g) = \epsilon > 0$ then every element of $L$ is a product of $N = [2/\epsilon]$ conjugates of $g$. Moreover $L = [L, g]^s N$.

Proof. Consider the realization of $\text{SU}(2) < \text{GL}_2(\mathbb{C})$ by unitary matrices:

$$\text{SU}(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1 \right\}.$$ 

The conjugacy class of an element $h \in \text{SU}(2)$ is uniquely determined by its trace $\text{tr}(h) \in [-2, 2]$. Write $\text{tr}(g) = 2 \cos \gamma$ with $\gamma \in [0, \pi]$; then for fundamental roots $\alpha$ we have $\alpha(g) = e^{\pm 2i\gamma}$, and so $\lambda(g) = 2\gamma/\pi$ if $\gamma \in [0, \pi/2]$, $\lambda(g) = 2(\pi - \gamma)/\pi$ otherwise. Of course $\lambda(g) = \lambda(-g)$ and $(g^L)^sN = L$ is equivalent to $((-g)^L)^sN = L$. So by replacing $g$ with $-g$ if necessary we may assume that $\lambda(g) = 2\gamma/\pi = \epsilon > 0$ and $\gamma = \epsilon \pi/2 \in (0, \pi/2]$. Now a direct computation shows that if $h \in L$ is a diagonal element with $\text{tr}(h) = 2\cos \theta$ then for any $\theta_1 \in [\theta - \gamma, \theta + \gamma]$ we may find a matrix $g' \in g^L$. (i.e. such that $\text{tr}(g') = 2\cos \gamma$) with $\text{tr}(hg') = 2\cos \theta_1$. This shows that for any integer $m > 1$, any element of $L$ with trace $2\cos \theta_2$ with $\theta_2 \in [0, m\gamma]$ is a product of $m$ conjugates of $g$. Taking $N = [2/\epsilon]$ we have $N\gamma \geq \pi$ and so $(g^L)^sN = L$.

This proves the first claim of the Lemma. The second claim follows since $[L, g] = (g^{-1})^L g = g^L \cdot g$ and

$$[L, g]^sN = (g^L \cdot g)^sN = (g^L)^sN g^N = L g^N = L.$$ 

We now consider the general case of Lemma 5.19. It is enough to prove it when $L$ is simply connected, since the definition of $\lambda$ was the same for $L$ and $L/\mathbb{Z}(L)$. We shall assume this from now on.

Let us write $H_{\alpha}$ for the one-parameter torus $\{h_{\alpha}(t) \mid t \in S_1\}$ given by the image of the cocharacter $h_{\alpha}$. Thus we have $T = H_{\beta_1} \times \cdots \times H_{\beta_r}$. Take an element $g \in L$ with $\lambda(g) = \epsilon > 0$. Then for at least one fundamental root $\beta_j$ we have $l(\beta_j(g)) \geq \epsilon \pi$. Fix $N = [2/\epsilon]\pi$ as above.

Case 1: Assume that the rank $r$ of $L$ satisfies $r \leq \max\{10, 4/\epsilon\}$.

In the central product $S_\beta T_\beta$ we can write $g$ as $g = g_1 g_2$ where $g_1 \in S_\beta$ and $g_2 \in T_\beta$. Now $S_\beta$ is a copy of $\text{SU}(2)$ (not $\text{PSU}(2)$ since $L$ is simply connected), and by Lemma 5.20 we can express any $h \in S_\beta$ as $h = \prod_{i=1}^N [s_i, g_1]$ for some $s_i \in S_\beta$. Then

$$h = \prod_{i=1}^N [s_i, g]$$

and in particular the subgroup $H_\beta \leq S_\beta$ is contained in $K^{sN}$. Recall that the Weyl group $W$ acts on $H$. For a pair of roots $\gamma_1, \gamma_2$ of $\Phi$ of the same length there is some element $v \in W$ such that $\gamma_1^v = \gamma_2$ and consequently $H_{\gamma_2}^v = H_{\gamma_2}$.}

70
Moreover, if $\gamma, \delta$ are two roots of different lengths in $\Phi$ then $\gamma$ is in the linear span of roots $\delta_1$ and $\delta_2$ in the orbit of $\delta$ under $W$, and then

$$H_\gamma \leq H_{\delta_1} H_{\delta_2} = H_{u_1} H_{u_2}$$

for some $u_1, u_2 \in W$.

Therefore each of the groups $H_{\beta_i}$ is contained in $K^{*4N}$. But $T$ is a product of all the $H_{\beta_i}$ for $i = 1, \ldots, r$ and hence

$$T \subseteq K^{*4rN}.$$  

Now the right-hand side is a union of conjugacy classes of $L$; since every conjugacy class intersects $T$ we have $L = K^{*4rN} = K^M$ as long as $M \geq 4rN = O(\epsilon^{-2})$, since $r \leq \max\{10, 4/\epsilon\}$.

**Case 2**: The Lie rank of $L$ exceeds both 10 and $4/\epsilon$. This means that $L$ is a classical Lie group of type $A_r, B_r, C_r$ or $D_r$. In all these cases we can label the fundamental roots in $\Pi$ so that $\beta_1, \ldots, \beta_{r-1}$ span a root system of type $A_{r-1}$ and the angle between $\beta_i$ and $\beta_{i+1}$ is $2\pi/3$ for $i = 1, \ldots, r-2$. (This is the labelling on the vertices of the Dynkin diagram of $L$ where we number the vertices on the $A_{r-1}$ part of the diagram consecutively.) The last root $\beta_r$ may have different length from the others.

Put $\eta = \epsilon/8$. It is immediate that for a subset $\Delta \subseteq \Pi$ of size at least $4\eta r$ we must have $l(\beta_i(g)) \geq \epsilon r/2$ for all $i \in \Delta$: otherwise, the average on $\Pi$ could not be $\epsilon r$ since each $l(\beta_i(g)) \leq \pi$. Define

$$\Pi_1 = \{\beta_i | 1 \leq i \leq r-1 \text{ and } i \text{ even}\}, \quad \Pi_2 = \Pi \setminus (\Pi_1 \cup \{\beta_r\}).$$

Then each $\Pi_i$ consists of pairwise orthogonal roots and their union is $\Pi \setminus \{\beta_r\}$.

Observe that $|\Delta| \geq \epsilon r/2 \geq 2$ since $r \geq 4/\epsilon$. Put $\Delta_i = \Pi_i \cap \Delta$. Since $|\Delta_1| + |\Delta_2| \geq |\Delta| - 1 \geq |\Delta|/2 \geq 2\eta r$ we have either $|\Delta_1| \geq \eta r$ or $|\Delta_2| \geq \eta r$. Without loss of generality assume that $|\Delta_1| \geq \eta r$.

The roots in $\Delta_1$ are pairwise orthogonal. The group $Q := \langle T, S_\beta \mid \beta \in \Delta_1 \rangle$ is therefore isomorphic to the central product

$$\left( \prod_{\beta \in \Delta_1} S_\beta \right) \circ T_{\Delta_1},$$

where $T_{\Delta_1} = \{h \in T \mid \beta(h) = 1 \forall \beta \in \Delta_1\}$ and $\prod_{\beta \in \Delta_1} S_\beta$ is the direct product of the $S_\beta$.

Now if $\beta \in \Delta_1$ we have $l(\beta(g)) \geq \epsilon r/2$. Just as in Case 1, working independently in each $S_\beta$ and using Lemma 5.20 we deduce that

$$\prod_{\beta \in \Delta_1} H_\beta \subseteq [Q, g]^{N_1} \subseteq K^{2N_1}$$

(39)

where $N_1 = \lceil 4/\epsilon \rceil$. We now refer to the following straightforward
Lemma 5.21 Let \( \Psi \) be the set of roots in the root system of type \( A_n \). For an integer \( m \leq n/2 \) let \( X, Y \in \Psi^{(m)} \) be two \( m \)-tuples of elements of \( \Psi \) each consisting of pairwise orthogonal roots. Then \( X = Y^w \) for an element \( w \) in the Weyl group of \( \Psi \).

Proof. This can be done directly from the realization of \( \Psi \) and the fact that \( W = \text{Sym}(n+1) \). Alternatively it follows by induction on \( m \) and using that for any root \( \alpha \in \Psi \), the orthogonal complement \( \Psi \cap \alpha^\perp \) is a root system of type \( A_{n-2} \). □

Now the set \( \Pi_1 \) is a union of at most \( |\Pi_1|/\eta r + 1 \) subsets of size \( |\Delta_1| \) and the same holds for \( \Pi_2 \). Altogether \( \Pi_1 \cup \Pi_2 \) is a union of at most \( r/\eta r + 2 = 1/\eta + 2 \) subsets of size \( |\Delta_1| \). Using Lemma 5.21 and (39) we see that

\[
\prod_{i=1}^{r-1} H_{\beta_i} \subseteq K^{*N_2}
\]

where \( N_2 = 2 \left[ 1/\eta + 2 \right] N_1 \). Finally \( H_{\beta_r} \subseteq K^{*4N} \), and hence \( T \subseteq K^{N_2+4N} \). Again, it follows that \( L = K^{*M} \) as long as \( M \geq N_2 + 4N = O(\epsilon^{-2}) \).

5.6 Countable quotients of compact groups

In this subsection, by a quotient of a topological group \( G \) we mean a quotient of the underlying abstract group, unless stated otherwise. We will be interested in countable quotients: in this subsection, one can always replace ‘countable’ with ‘of cardinality strictly less than \( 2^{\aleph_0} \).’

Until further notice, we assume that \( G \) is a compact group such that the profinite quotient \( G/G^0 \) is finitely generated (topologically). Recall (Corollary 5.7) that the derived group \( G' \) of \( G \) is closed; this applies likewise if \( G \) is replaced by any open subgroup of \( G \).

The following observation is an immediate consequence of Corollary 5.7.

Corollary 5.22 If \( M \) is a normal subgroup of \( G \) and \( G/M \) is residually finite then \( M \) is closed.

Indeed, \( M \) is an intersection of normal subgroups of finite index, each of which is open.

Suppose to begin with that \( G \) is infinite and abelian. If \( G/G^0 \) has \( \mathbb{Z}_p \) as a quotient for some prime \( p \) then, as observed in the introduction, we obtain a homomorphism

\[
G \to \mathbb{Z}_p \to \mathbb{Q}_p \to \mathbb{Q}
\]

with countably infinite image. If \( G/G^0 \) is infinite but does not have any quotient of type \( \mathbb{Z}_p \), then \( G/G^0 \) must have infinitely many Sylow subgroups, and so has a quotient \( Q = \prod_{p \in \pi} C_p \) where \( \pi \) is an infinite set of primes. We may identify
Q with the additive group of \( S = \prod_{p \in \pi} F_p \), which maps onto a non-principal ultraproduct \( \tilde{S} \) of the \( F_p \). Now \( \tilde{S} \) is a field of characteristic zero, hence admits an additive epimorphism to \( \mathbb{Q} \); thus \( G \) admits an epimorphism to \( \mathbb{Q} \). (We are indebted to J. Kiehlmann for pointing out a gap in our original argument.)

If \( G/G^0 \) is finite then \( G^0 \neq 1 \), and then \( G^0 \) maps onto a torus \( T \). Let \( D \) be the torsion subgroup of \( T \). Then \( T/D \) is a divisible torsion-free abelian group, so a vector space over \( \mathbb{Q} \); choosing an epimorphism \( T/D \to \mathbb{Q} \) we obtain an epimorphism (of abstract groups) \( G^0 \to \mathbb{Q} \).

Now suppose that \( G \) has an open normal subgroup \( K \) such that \( K/K' \) is infinite. The preceding remarks shows that \( K \), and therefore also \( G \), has a countably infinite quotient.

A group \( Q \) is said to be \( FAb \) if every virtually-abelian quotient of \( Q \) is finite; when \( Q \) is a topological group, this refers to continuous quotients.

**Theorem 5.23** Let \( G \) be a compact group such that \( G/G^0 \) is (topologically) finitely generated. Then every countable \( FAb \) quotient of \( G \) is finite.

Before giving the proof, let us deduce

**Corollary 5.24** \( G \) has a countably infinite quotient if and only if \( G \) is not \( FAb \).

We remark that many familiar compact groups are \( FAb \): among connected groups, these are just the semisimple ones; among profinite groups, examples include \( \mathfrak{S}(\mathbb{Z}_p) \) for Chevalley groups \( \mathfrak{S} \).

**Proof.** The remarks above show that if \( G \) is not \( FAb \) then \( G \) has a countably infinite quotient. Suppose conversely that \( G \) has a countably infinite quotient \( G/N \). By Theorem 5.23 we may suppose that \( G/N \) is virtually abelian, so \( G \) has a normal subgroup \( K \) of finite index with \( K' \leq N \leq K \). Now \( K \) is open by Corollary 5.7 and so \( K' \) is closed. Thus \( G/K' \) is an infinite virtually-abelian continuous quotient of \( G \), so \( G \) is not \( FAb \).

**Proof of Theorem 5.23.** Let \( H \) be a normal subgroup of \( G \) such that \( G/H \) is countable and \( FAb \), and suppose that \( G/H \) is infinite.

Set \( P = (G^0)' \). Then \( P \) is closed in \( G \) and \( P \) is a semisimple connected compact group, hence has no proper countable quotient, by Theorem 5.10 (and the remark preceding it). So \( H \geq P \), and replacing \( G \) by \( G/P \) we may suppose that \( G^0 \) is abelian.

Since \( G^0 H/H \) is abelian, \( G/G^0 H \) must be infinite. Replacing \( G \) by \( G/G^0 \) and \( H \) by \( G^0 H/G^0 \), we may suppose that \( G \) is a finitely generated profinite group. Put \( K = \overline{H} \); then \( K \) is open in \( G \), so \( K \) is again a finitely generated profinite group. Now \( G/K' H \) is virtually abelian and therefore finite. Thus \( K' H \) is open by Theorem 5.1 and so \( K' H = K \).

Now recall the definition of \( K_0 \) (see the Introduction). This is a characteristic closed subgroup of \( K \) such that \( K^{(3)} K_0/K_0 \) is semisimple, where \( K/K^{(3)} \) is
soluble of derived length at most 3. Since any soluble FAb group is finite, we infer that \( G/K(3)K_0H \) is finite, and as before conclude that \( K(3)K_0H = K \). Thus \( K/HK_0 \) is a countable image of the finitely generated semisimple group \( K(3)K_0/K_0 \); so \( K/HK_0 \) is finite by Theorem 5.10, and as above it follows that \( HK_0 = K \).

Now Corollary 1.8 shows that \( H = K \). Hence \( G/H \) is finite, a contradiction.

Now we consider arbitrary compact groups:

**Theorem 5.25** Let \( G \) be a compact group and \( N \) a normal subgroup of (the underlying abstract group) \( G \). If \( G/N \) is finitely generated then \( G/N \) is finite.

**Proof.** Suppose that \( G/N \) is finitely generated and infinite. Then \( G = N \langle X \rangle \) for some finite subset \( X \). Let \( K = \langle X \rangle \) be the subgroup topologically generated by \( X \). Then \( G/N \cong K/(K \cap N) \), so replacing \( G \) by \( K \) we may suppose that \( G \) is topologically finitely generated. Now \( G/N \) is countable, hence by Theorem 5.23 there exists \( M \triangleleft G \) with \( M \geq N \) such that \( G/M \) is infinite and virtually abelian. But a finitely generated virtually abelian group is residually finite; hence \( M \) is closed in \( G \), by Corollary 5.22. Thus \( G/M \) is both countably infinite and compact, a contradiction.

### 5.7 Dense normal subgroups

Let \( G \) be a compact group such that \( G/G^0 \) is (topologically) finitely generated. If \( N \triangleleft G \) and \( G/N \) is countable then the closure \( \overline{N} \) of \( N \) is open in \( G \); in this case, we say that \( N \) is virtually dense. Generalizing the preceding subsection, we can ask: under what conditions does \( G \) have a virtually dense normal subgroup \( N \) of infinite index? Note that \( N \) has infinite index if and only \( N \) is not closed, in view of Corollary 5.7.

Suppose that \( G \) is abelian. If \( G/G^0 \) is infinite, then \( G/G^0 \) contains a dense (abstractly) finitely generated subgroup. If \( G^0 \) is infinite, then \( G^0 \) has a dense proper subgroup (necessarily of infinite index), because it maps onto a torus.

A group of the form \( \prod_{i \in I} S_i \) is said to be strictly infinite semisimple if the index set \( I \) is infinite and either each \( S_i \) is a finite (non-abelian) simple group or each \( S_i \) is a connected compact simple Lie group. Such a group has a characteristic dense subgroup of infinite index, namely the restricted direct product \( N \) of the \( S_i \). Note that \( N \) is countable if \( I \) is countable and the \( S_i \) are finite groups.

It turns out that these examples essentially account for all possibilities:

**Theorem 5.26** Let \( G \) be a compact group such that \( G/G^0 \) is (topologically) finitely generated. Then \( G \) has a virtually dense normal subgroup of infinite index if and only if \( G \) has an open normal subgroup \( H \) and a closed normal subgroup \( K < H \) such that \( H/K \) is either infinite and abelian or strictly infinite semisimple.
In one direction, this follows quickly from the preceding observations. Supposing that \( H \) and \( K \) exist as indicated, we may as well assume that \( K = 1 \). Necessarily dense subgroup \( N \) of infinite index, and then \( N \) is normal in \( G \).

Now suppose that \( H \) is abelian. If \( G/G^0 \) is finite then \( H/G^0 \) has a countable dense subgroup \( M/G^0 \). Then \( N := \langle M^G \rangle = M^{g_1} \cdots M^{g_n} \) is virtually dense and normal in \( G \), where \( \{g_1, \ldots, g_n\} \) is a set of coset representatives for \( G/H \), and \( N/G^0 \) is countable, so \( N \) has infinite index in \( G \). Suppose finally that \( G/G^0 \) is finite. As \( G^0 \) is a compact connected abelian group, it has a subgroup \( T \) such that \( G^0/T \) is a one-dimensional torus. Put \( S = T^{g_1} \cap \ldots \cap T^{g_n} \) where \( \{g_1, \ldots, g_n\} \) is a set of coset representatives for \( G/G^0 \). Then \( G^0/S \) is a torus, so has a countable dense subgroup \( M/S \) (in fact we can choose \( M/S \) to be cyclic). Now take \( N = \langle M^G \rangle = M^{g_1} \cdots M^{g_n} \) as before.

For the converse, let \( N \) be a normal subgroup of infinite index in (the abstract group) \( G \) such that \( L = N \) is open in \( G \). Note that \( L \geq G^0 \) and that \( L/G^0 \) is a finitely generated profinite group. It will suffice to find an open normal subgroup \( H \) of \( G \) and a closed normal subgroup \( K \) of \( H \) such that \( H/K \) is either infinite and abelian or strictly infinite semisimple; for if \( \{g_1, \ldots, g_n\} \) is a set of coset representatives for \( G/H \) then \( K = K^{g_1} \cap \ldots \cap K^{g_n} \) is closed and normal in \( G \), and \( H/K \) is a subdirect product of copies of \( H/K \), hence shares the given property of \( H/K \).

Now we separate cases.

Case 1: where \( G^0 = 1 \), i.e. \( G \) is profinite.

Recall that \( L' \) is closed, by Corollary 8.12. Suppose that both \( L/L' \) and \( L/L_0 \) are finite. Then both \( L' \) and \( L_0 \) are open in \( L \), so \( NL' = NL_0 = L \). It follows by Corollary 8.13 (applied to the finitely generated profinite group \( L \)) that \( N = L \), a contradiction. Therefore at least one of \( L/L' \), \( L/L_0 \) is infinite.

If \( L/L' \) is infinite we set \( H = L \) and \( K = L' \). Suppose finally that \( L/L_0 \) is infinite, and put \( T = L^{(0)}L_0 \); recall that \( T/L_0 \) is semisimple (a consequence of Proposition 1.18). If \( T/L \) is finite, then \( T/L_0 \) is infinite; in this case, set \( H = T \) and \( K = L_0 \). If \( L/T \) is infinite, then some term \( S \) of the derived series of \( L \) must satisfy: \( L/S \) is finite and \( S/S' \) is infinite. In this case, we take \( H = S \) and \( K = S' \).

Case 2: where \( G \) is connected.

In this case, \( N \) is dense in \( G \). According to [HM, Theorem 9.24], \( G \) is a quotient \((A \times P)/Z\) where \( A \) is a connected compact abelian group, \( P = \prod_{i \in I} S_i \) is a connected compact semisimple group, and \( Z \leq Z(P) \). If we assume that \( G \) has no infinite abelian image, it follows that \( G \cong P/(P \cap Z) \). If \( G \) has a proper dense normal subgroup, then so does \( P \). Now the claim (*) in Subsection 5.5 above, shows that there exists a non-principal ultrafilter on the index set \( I \): but this implies that \( I \) is infinite. Thus \( G \cong P/(P \cap Z) \) has a strictly infinite semisimple quotient \( G/K \) isomorphic to the product \( \prod_{i \in I} S_i/Z(S_i) \).
The General Case.

If \( NG^0 < L \) the result follows by Case 1 applied to \( G/G^0 \). So we may assume that \( NG^0 = L \). Let \( Z = Z(G^0) \). If \( L/ZN \) is finite then \( H := ZN \) is open in \( G \) and \( K := H' \) is closed (Corollary 5.9); and \( K \) has infinite index in \( H \) because \( K \leq N \).

So replacing \( G \) by \( G/Z \) and \( N \) by \( ZN/N \) we may assume that \( Z(G^0) = 1 \).

In this case, \( G^0 = \prod_{i \in I} S_i \) where each \( S_i \) is a connected (and centreless) simple Lie group (\([HM]\), loc. cit.). Put \( D = G^0 \cap N \). Then \( [G^0, N] \leq D \). It follows that

\[
G^0 = G'^0 \leq [G^0, L] = [G^0, N] \leq D,
\]

so \( D \) is dense in \( G^0 \). In particular, in view of Case 2 above, the index set \( I \) must be infinite.

Since \( G/G^0 \) is finitely generated, so is \( L/G^0 \); thus \( L = G^0(y_1, \ldots, y_{67}) \) for some \( y_i \in N \). Then \( [G^0, y_i] \leq D \) for each \( i \). Applying Proposition 5.13 we deduce that there exists an infinite subset \( J \) of \( I \) such that each \( y_i \) normalizes \( S_i \) for every \( i \in J \). As \( N_L(S_i) \) is closed and contains \( G^0 \), it follows that \( S_i \) is normal in \( L \) for every \( i \in J \). Put \( C_i = C_L(S_i) \). Then \( L/C_i \) embeds in the outer automorphism group of \( S_i \), which embeds in \( \text{Sym}(3) \) (cf. \([HM]\), page 256). As the finitely generated profinite group \( L/G^0 \) admits only finitely many homomorphisms into \( \text{Sym}(3) \) and \( C_i S_i \geq G^0 \), it follows that \( L \) has a characteristic open subgroup \( H \geq G^0 \) such that \( C_i S_i \geq H \) for all \( i \in J \).

Thus putting \( X = \prod_{i \in J} S_i \) we have \( H = C_H(X) \times X \); indeed, if \( h \in H \) then \( h = c_i s_i \) \((c_i \in C_i, s_i \in S_i)\) for each \( i \in J \), and if \( x = (s_j)_{j \in J} \) then \([hx^{-1}, s_j] = 1\) for every \( j \in J \), so \( hx^{-1} \in C_H(X) \). To complete the proof we may therefore take \( K = C_H(X) \).

Remark. It might be more natural to ask: when does \( G \) have a virtually normal virtually dense subgroup? \((N \text{ is virtually normal if the normalizer } N_G(N) \text{ has finite index in } G)\).

**Corollary 5.27** \( G \) has a virtually normal virtually dense subgroup of infinite index if and only if \( G \) has a normal virtually dense subgroup of infinite index.

This follows from the theorem: suppose that \( R \) is a subgroup of finite index in \( G \), that \( H \) is open and normal in \( R \), and that \( K \leq H \) is a closed normal subgroup of \( R \). Then as above we can replace \( K \) by a closed normal subgroup \( K^* \) of \( G \) such that \( H/K^* \) is a subdirect product of \(|G:R| \) copies of \( H/K \), and replace \( H \) by \( H^* \), where \( H^* \) is normal of finite index in \( G \). Then \( H^* \) is open by Corollary 5.7 whence \( H^* / K^* \) is again an infinite abelian or semisimple group of the same type as \( H/K \).

The conditions for the existence of a proper dense normal subgroup are more delicate, and we merely state the result. The proof, which depends on Corollary 1.8 and further arguments in the spirit of Subsection 4.4 will appear elsewhere.
**Definition.** (a) Let $S$ be a finite simple group. Then $Q(S)$ denotes the following subgroup of $\text{Aut}(S)$:

- $\text{InnDiag}(S) \langle \tau \rangle$ if $S = D_n(q)$, $n \geq 5$
- $\text{InnDiag}(S) \langle [q] \rangle$ if $S = 2D_n(q)$
- $\text{InnDiag}(S)$ if $S$ is of another Lie type
- $\text{Aut}(S)$ in all other cases

where $\tau$ is the non-trivial graph automorphism of $D_n(q)$ and $[q]$ denotes the field automorphism of order 2 of $2D_n(q)$.

(b) Let $S$ be a connected simple Lie group. Then

$Q(S) = \begin{cases}
\text{Aut}(S) & \text{if } S = \text{PSO}(2n), \ n \geq 3 \\
\text{Inn}(S) & \text{else}
\end{cases}$

(c) A topological group $H$ is $Q$-almost-simple if $S \triangleleft H \leq Q(S)$ where $S$ is a finite simple group or a connected simple Lie group.

If $H$ is $Q$-almost-simple as above, the rank of $H$ is then the rank of $S$, namely the (untwisted) Lie rank if $S$ is of Lie type, $n$ if $S \cong \text{Alt}(n)$, and zero otherwise.

**Theorem 5.28** Let $G$ be a compact group with $G/G^0$ finitely generated. Then $G$ has a proper dense normal subgroup if and only if one of the following holds:

- $G^{\text{ab}}$ is infinite, or
- $G$ has a strictly infinite semisimple quotient, or
- $G$ has $Q$-almost-simple quotients of unbounded ranks.

**References**

[AG] M. Aschbacher and R. M. Guralnick, Some applications of the first cohomology group, *J. Algebra* 90 (1984), 446-460.

[As] M. Aschbacher, *Finite group theory*, Cambridge Univ. Press, Cambridge, 1988.

[B] H. Blau, A fixed-point theorem for central elements in quasisimple groups, *Proc. AMS* 122 (1994), 79-84.

[BCP] L. Babai, P. J. Cameron and P. Pálfy, On the orders of primitive groups with restricted non-abelian composition factors, *J. Algebra* 79 (1982), 161-168.

[BNP] L. Babai, N. Nikolov and L. Pyber, Product Growth and Mixing in Finite Groups, *19th ACM-SIAM Symposium on Discrete Algorithms*, SIAM, 2008, Pages 248-257.
[Bu] D. Bump, *Lie groups*, Springer-Verlag, New York, 2004.

[C] R. W. Carter, *Finite groups of Lie type: conjugacy classes and complex characters*, Wiley and Sons, London, 1985.

[DM] J. D. Dixon and B. Mortimer, *Permutation groups*, Springer-Verlag, New York, 1996.

[FG] J. Fulman, R. Guralnick, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups with applications to derangements, [http://arxiv.org/abs/0902.2238](http://arxiv.org/abs/0902.2238)

[FMS] T. Frayne, A. Morel and D. Scott, Reduced direct products, *Fund. Math.* 51 (1962), 195-228.

[FJ] M. Fried and M. Jarden, *Field arithmetic*, Springer-Verlag, Berlin – Heidelberg, 1986.

[GaSh] S. Garion and A. Shalev, Commutator maps, measure preservation, and T-systems, *Trans. Amer. Math. Soc.* 361 (2009), 4631–4651.

[Gch] W. Gaschütz, Zu einem von B. H. und H. Neumann gestellten Problem, *Math. Nachrichten* 14 (1955), 249-252.

[GL] R. M. Guralnick and F. Lübeck, On p-singular elements in Chevalley groups in characteristic p, in *Groups and computation III*, 169-182, Ohio State Univ. Math. Res. Inst. Publ. 8, de Gruyter, Berlin, 2001.

[GLS] D. Gorenstein, R. Lyons and R. Solomon, *The classification of the finite simple groups*, no.3, American Math. Soc., Providence, Rhode Island, 1998.

[Go] D. Gorenstein, *Finite Groups*, 2nd ed., Chelsea, New York, 1980.

[Gt] M. Goto, A theorem on compact semisimple groups. *J. Math. Soc. Japan* 1 (1949), 270-272.

[GFSG] D. Gorenstein, *Finite simple groups*, Plenum Press, New York and London, 1982.

[GSS] D. Gluck, A. Seress and A. Shalev, Bases for primitive permutation groups and a conjecture of Babai, *J. Algebra* 199 (1998), 367–378.

[HM] K. H. Hofmann and S. A. Morris, *The structure of compact groups*. 2nd edn., de Gruyter Studies in Mathematics, 25. Walter de Gruyter & Co., Berlin, 2006.

[J] G. A. Jones, Varieties and simple groups, *J. Austral. Math. Soc.* 17 (1974), 163–173.

[JZ] A. Jaikin-Zapirain, On linear just infinite pro-p groups, *J. Algebra* 255 (2002), 392-404.

78
[KIL] P. Kleidman and M. Liebeck, *The subgroup structure of the finite classical groups*, LMS Lect. Notes 129, Cambridge Univ. Press, Cambridge, 1990.

[KL] M. Kapovich and B. Leeb, On asymptotic cones and quasi-isometry of fundamental groups of 3-manifolds, *GAFA* 5 (1995), 582-603.

[LaS] V. Landazuri and G. M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* 32 (1974), 418–443.

[LiSh] M. W. Liebeck and A. Shalev, Diameters of finite simple groups: sharp bounds and applications, *Annals of Math.* 154 (2001), 383-406.

[LiSh2] M. W. Liebeck and A. Shalev, Fuchsian groups, finite simple groups and representation varieties. *Invent. Math.* 159 (2005), 317–367.

[LOST] M. Liebeck, E. O’Brien, A. Shalev and P. Tiep, The Ore conjecture, *J. European Math. Soc.* 12 (2010), 939-1008.

[LOST2] M. Liebeck, E. O’Brien, A. Shalev and P. Tiep, Commutators in finite quasisimple groups, *to appear*

[MZ] C. Martinez and E. Zelmanov, Products of powers in finite simple groups, *Israel J. Math.* 96 (1996), 469–479.

[NS] N. Nikolov and D. Segal, On finitely generated profinite groups, I: strong completeness and uniform bounds, *Annals of Math.* 165 (2007), 171–238.

[NS2] N. Nikolov and D. Segal, On finitely generated profinite groups, II: products in quasisimple groups, *Annals of Math.* 165 (2007), 239–273.

[NSP] N. Nikolov and D. Segal, Powers in finite groups, *Groups, Geometry and Dynamics*, *to appear*; arXiv:0909.6439

[SW] J. Saxl and J. S. Wilson, A note on powers in simple groups, *Math. Proc. Cambridge Philos. Soc.* 122 (1997), 91–94.

[S1] D. Segal, Closed subgroups of profinite groups, *Proc. London Math. Soc.* 81 (2000), 29–54.

[S2] D. Segal, *Words: notes on verbal width in groups*, London Math. Soc. Lecture Notes Series 361, Cambridge Univ. Press, Cambridge, 2009.

[SGT] J.-P. Serre, *Topics in Galois Theory*, Res. Notes Math. 1, Jones and Bartlett, Boston – London, 1992.

[W] J. S. Wilson, On simple pseudofinite groups, *J. London Math. Soc.* 51 (1995), 471–490.
[Z] E. I. Zelmanov, On the restricted Burnside problem, *Proc. Internl. Congress Math. Kyoto 1990*, Math. Soc. Japan, Tokyo, 1991, pp. 395-402.