TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Mathematics
Institute of Geometry

On ultraproducts of compact quasisimple groups

Doctoral Thesis

To Receive the Degree

Doctor rerum naturalium (Dr.rer.nat.)

in Mathematics

Degree: summa cum laude

Author:
M. Sc. Jakob Schneider
(* January 26, 1994, Peine, Germany)

Referees:
Prof. Dr. Andreas Thom (TU Dresden)
Prof. Dr. Martin Liebeck (Imperial College, London)

November 09, 2020
(date of defense)

Statutory Declaration

I declare on oath that I completed this work on my own and that information which has been directly or indirectly taken from other sources has been noted as such. Neither this, nor a similar work, has been published or presented to an examination committee.

Dresden, November 09, 2020 Jakob Schneider

Contents

Introduction **1**

0 Notation, basic definitions, and facts **7**

0.1 Group theory . . . 7

0.2 Some ring and field theory . . . 13

0.3 Ultraproducts and norms . . . 13

1 The normal subgroup lattice of an algebraic ultraproduct **17**

1.1 Introduction . . . 17

1.2 Auxiliary geometric results . . . 19

1.3 Relative bounded normal generation in universal finite quasisimple groups . 21

1.4 The lattice of normal subgroups . . . 24

2 Metric approximation of groups by finite groups **31**

2.1 Introduction . . . 31

2.2 Preliminaries . . . 33

2.2.1 On $\mathcal{C}$-approximable abstract groups . . . 33

2.2.2 On $\mathcal{C}$-approximable topological groups . . . 34

2.3 On **Sol**-approximable groups . . . 37

2.4 On **Fin**-approximable groups . . . 43

2.5 On the approximability of Lie groups . . . 47

3 Word maps are surjective on metric ultraproducts **53**

3.1 Introduction . . . 53

3.2 Symmetric groups . . . 57

3.2.1 Power words . . . 57

3.2.2 The cycle structure of elements from $\mathrm{PSL}_2(q)$ . . . 61

3.2.3 Effective surjectivity of word maps over finite fields . . . 62

3.2.4 Proof of Theorem 3.1 . . . 64

3.3 Unitary groups . . . 67

3.3.1 Proof of Theorem 3.3 . . . 67

3.3.2 Further implications . . . 72

3.3.3 Concluding remarks . . . 73

3.4 Finite groups of Lie type . . . 74
3.4.1 The linear case . . . 75
3.4.2 The case of quasisimple groups of Lie type stabilizing a form . . . 77
3.4.3 An alternative way of proving Theorem 3.1 using wreath products . 89

4 Isomorphism questions for metric ultraproducts 91
4.1 Introduction . . . 91
4.2 Description of conjugacy classes in $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q)$, and $\mathrm{PGL}_{\mathcal{U}}(q)$ . . . 93
4.3 Characterization of torsion elements in $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q)$, and $\mathrm{PGL}_{\mathcal{U}}(q)$ . . . 96
4.4 Faithful action of $\mathrm{S}_{\mathcal{U}}$ and $\mathrm{PGL}_{\mathcal{U}}(q)$ . . . 97
4.5 Centralizers in $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q)$, $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, and $\mathrm{GU}_{\mathcal{U}}(q)$ . . . 99
4.6 Centralizers in $\mathrm{PGL}_{\mathcal{U}}(q)$, $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, and $\mathrm{PGU}_{\mathcal{U}}(q)$ . . . 102
4.7 Double centralizers of torsion elements . . . 103
4.7.1 The case $\mathrm{S}_{\mathcal{U}}$ . . . 103
4.7.2 The case $\mathrm{PGL}_{\mathcal{U}}(q)$, $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, and $\mathrm{PGU}_{\mathcal{U}}(q)$ . . . 103
4.8 Distinction of metric ultraproducts . . . 108
4.8.1 Computation of $e_{\overline{G}}(o)$ when $\gcd\{o,p\} = \gcd\{o,|Z|\} = 1$ . . . 108
4.8.2 Proof of Theorem 4.1 . . . 111

Index of Symbols 113

Index 125

Bibliography 129

Introduction

Recently, the class of sofic groups has attracted much interest [1, 6, 7, 13, 14, 15, 24, **?**, 38, 40, 58, 69, 76]. A group is called *sofic* if it can be approximated by finite symmetric groups equipped with the normalized Hamming length function (see Theorem 3.5 of [58]). An alternative characterization of soficity is that the group embeds into a metric ultraproduct of finite symmetric groups equipped with the normalized Hamming length function. Amenable and residually finite groups are known to be sofic (see Examples 4.2 and 4.4 of [58]). However, there are more sofic groups than this (see [6, 40]). Some applications of the notion of soficity are the following: Gromov showed in [29] that Gottschalk's surjunctivity conjecture holds for every sofic group. Elek and Szabó [13] settled Kaplansky's direct finiteness conjecture for every sofic group. They also proved in [14] that Connes' embedding conjecture and Lück's determinant conjecture hold for such groups. Recently, Nitsche and Thom proved in [56] that certain systems of group equations can be solved over every group satisfying Connes' embedding conjecture and hence over every sofic group.

The question whether every group is sofic remains open until now.

In this thesis, motivated by the notion of soficity, more generally, we study various properties of metric ultraproducts of finite groups, each of them being equipped with a 'natural' norm. Here, by a *norm* on a group G we mean a function $\ell\colon G \to [0, \infty]$ which satisfies $\ell(g) = \ell(g^{-1}) = \ell(g^h)$, $\ell(gh) \leq \ell(g) + \ell(h)$, and $\ell(g) = 0$ iff $g = 1_G$ for all $g, h \in G$. Such a metric ultraproduct of a family of normed finite groups, i.e., groups equipped with a norm, along a non-principal ultrafilter reflects 'asymptotic properties' of the family. It can be seen as a limit object associated to the family, which makes it easier to express such properties. E.g., that a word map on such an ultraproduct is surjective means that it is almost surjective in a metric sense on the groups of the family (see Chapter 3).

Subsequently, we give a brief outline of the structure of this thesis.

Chapter 0 provides the reader with the necessary prerequisites. In Sections 0.1 and 0.2, we introduce the necessary group and ring theoretic notation. In Section 0.3, we define the notions of an *algebraic* resp. *metric ultraproduct* of a family of abstract resp. normed groups (see Definitions 0.3 resp. 0.7). Later we remark that an algebraic ultraproduct can be seen as a metric ultraproduct (see Remark 0.11). We also introduce the notion of an *invariant length function* (or *norm*) on a group (see Definition 0.5) and provide definitions of the most common length functions on finite, linear, and projective linear groups (see

Definition 0.12), which we will use in the subsequent chapters. We point out the *Lipschitz equivalence* of some of them on the family of finite symmetric or alternating groups and classical finite simple groups of Lie type (see Fact 0.13).

In Chapter 1, we consider the lattice of normal subgroups of the algebraic ultraproduct $\mathcal{H}_{\mathcal{U}}$ of a family $\mathcal{H} = (H_i)_{i\in I}$ of universal finite quasisimple groups (cf. the beginning of Section 1.2). The corresponding lattice, when the H_i ($i \in I$) are finite alternating groups A_n of degree n tending to ∞ along $\mathcal{U}$, is linearly ordered and was described completely in [17]. Lemma 1.8 shows that this result still holds when the H_i ($i \in I$) are double covers $2\cdot \mathrm{A}_n = \widetilde{\mathrm{A}}_n$ of alternating groups A_n ($n \geq 8$); essentially we only add a two-element center compared to the previous situation. Regarding this, in [67] it was claimed that, more generally, the lattice of normal subgroups of $\mathcal{H}_{\mathcal{U}}$, when the H_i ($i \in I$) are non-abelian finite simple groups, is also linearly ordered. However, this is false in this form in most cases when the H_i ($i \in I$) are classical groups of Lie type of unbounded rank (i.e., of type $\mathrm{PSL}_n(q)$, $\mathrm{PSp}_{2m}(q)$, $\mathrm{P}\Omega_{2m+1}(q)$, $\mathrm{P}\Omega^{\pm}_{2m}(q)$, or $\mathrm{PSU}_n(q)$ for suitable $m, n \in \mathbb{Z}_+$, q a prime power) of rank tending to infinity along $\mathcal{U}$. E.g., see Example 1.2. In Lemmas 1.8 and 1.13, we point out a version of 'relative' bounded normal generation for universal finite quasisimple groups, generalizing the results from [49]. From Lemma 1.13 we derive Theorem 1.3, the main result of Chapter 1, which provides a complete description of the structure of the lattice of normal subgroups of an algebraic ultraproduct of universal finite quasisimple groups which are not double covers of alternating groups (as mentioned above, this case was already settled), correcting the above mentioned misstatement of [67]. It turns out that considering the universal finite quasisimple groups rather than finite simple groups is more natural here. We state here Theorem 1.3.

Theorem 1.3. *Let $(H_i)_{i\in I}$ be a sequence of universal finite quasisimple groups of Lie type (i.e., H_i is of type $\mathrm{SL}_n(q)$, $\mathrm{Sp}_{2m}(q)$, the double cover of $\Omega_{2m+1}(q)$ or $\Omega^{\pm}_{2m}(q)$ for q odd, $\Omega^{\pm}_{2m}(q)$ for q even, or $\mathrm{SU}_n(q)$ for suitable prime power q and $m, n \in \mathbb{Z}_+$; cf. the list at the beginning of Section 1.2), each endowed with the norms ℓ_{rk}, and $\overline{H}_i \coloneqq H_i/\mathbf{Z}(H_i)$ endowed with ℓ_{pr} (see Definition 0.12). Write V_i for the natural module of H_i ($i \in I$; see Section 0.1(f)). Assume that the dimensions $n_i \coloneqq n(H_i) \coloneqq \dim(V_i)$ tend to infinity along a fixed ultrafilter $\mathcal{U}$ on I. Set $G \coloneqq \prod_{i\in I} H_i$ and define the normal subgroups $N_{\mathrm{rk}} \coloneqq \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_{\mathrm{rk}}(h_i) = 0\}$ and $N_{\mathrm{pr}} \coloneqq \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_{\mathrm{pr}}(\overline{h}_i) = 0\}$ of G (where $\overline{h}_i$ means the image of $h_i \in H_i$ in $\overline{H}_i$). We also define $N_0 \coloneqq \{(h_i)_{i\in I} \in G \mid h_i = 1_{H_i} \text{ along } \mathcal{U}\}$ and suitable subgroups $N_1, A_0, A_1 \leq G$ defined in Section 1.4. Then the following hold.*

(i) The subgroup N_{pr} contains all proper normal subgroups of G containing N_0. In particular, G/N_{pr} is non-abelian simple and N_{pr}/N_0 is a characteristic subgroup of G/N_0.

(ii) The normal subgroups of G lying between N_0 and N_{rk} are linearly ordered. Any such normal subgroup is perfect.

(iii) Define maps between the following two sets

$$\begin{array}{c} \{N \trianglelefteq G \mid N_1 \leq N \leq N_{\mathrm{pr}}\} \\ {\scriptstyle\Psi}\big\uparrow\big\downarrow{\scriptstyle\Phi} \\ \{M \mid M \trianglelefteq G,\ N_1 \leq M \leq N_{\mathrm{rk}}\} \times \{A \mid A \trianglelefteq G,\ N_1 \leq A \leq A_1\} \end{array}$$

by $\Phi\colon N \mapsto (N \cap N_{\mathrm{rk}}, N \cap A_1)$ *and* $\Psi\colon (M, A) \mapsto MA$*. Then* Φ *and* Ψ *are isomorphisms of posets and mutually inverse to each other.*

(iv) If N *is normal in* G *containing* N_0 *and* $N_1 \not\leq N$*, then* $N \leq A_0$*.*

(v) A normal subgroup N *containing* N_0 *but not* N_1 *is contained in a normal subgroup* $K \leq N_{\mathrm{pr}}$ *containing* N_1 *if and only if* $N \leq A$*, where* $\Phi(K) = (M, A)$ *is the image of* K *under the map* Φ *from above.*

In this context, one well-known observation (due to Liebeck and Shalev [49]) is that $\mathcal{H}_{\mathcal{U}}$ contains a unique maximal normal subgroup $N_{\mathrm{pr}}/N_0 = \overline{N}_{\mathrm{pr}} \trianglelefteq \mathcal{H}_{\mathcal{U}}$ (cf. Theorem 1.3(i)). Hence $\mathcal{H}_{\mathcal{U}}^{\mathrm{met}} := \mathcal{H}_{\mathcal{U}}/\overline{N}_{\mathrm{pr}}$ is simple.

It turns out that this group $\mathcal{H}_{\mathcal{U}}^{\mathrm{met}}$ is isomorphic to the underlying abstract group of the metric ultraproduct of the groups $\overline{H}_i$ (see Theorem 1.3; $i \in I$) with respect to the (up to Lipschitz equivalence) unique minimal norm ℓ_i ($i \in I$). In Chapters 2–4, we will only deal with such metric ultraproducts rather than algebraic ones.

In Chapter 2, we discuss the question which abstract and topological groups embed into a metric ultraproduct of normed finite groups (with the norms chosen arbitrarily). We introduce the concepts of a $\mathcal{C}$*-approximable abstract* and *topological group* for a class $\mathcal{C}$ of finite groups (which was first introduced by Holt and Rees in [34]). This is a common generalization of the concepts of a sofic, weakly sofic, and linear sofic group (see Definitions 2.1 and 2.8; note that we allow arbitrary norms on the groups from the class $\mathcal{C}$ compared to Definition 1.1 of [34]).

Glebsky [24] raised the question whether all groups are approximable by finite solvable groups with arbitrary norm. We answer this in the negative by proving the following theorem.

Theorem 2.17. *Any non-trivial finitely generated and perfect group is not approximable by finite solvable groups.*

This generalizes a counterexample of Howie [35]. On a related note, we establish the following result, which is based on a private note of Nikolov.

Theorem 2.25. *Any non-trivial finitely generated group which can be approximated by finite groups has a non-trivial quotient which can be approximated by finite projective special linear groups. In particular, every simple such group can be approximated by finite projective special linear groups.*

Moreover, we discuss the question which connected Lie groups can be embedded into a metric ultraproduct of normed finite groups. Regarding this, we prove the following result.

Theorem 2.33. *A connected Lie group is approximable by finite groups as a topological group if and only if it is abelian.*

This provides a negative answer to a question of Doucha [10, Question 2.11]. Referring to a problem of Zilber [79, page 17] (also Question 1.1 of Pillay [59]), we show the following.

Theorem 2.37. *A Lie group equipped with a norm generating its topology that is an abstract quotient of a product of finite groups has abelian identity component.*

We point out that both, Theorem 2.33 and Theorem 2.37, give an alternative proof of a result of Turing [74]. Finally, we solve a conjecture of Pillay [59, Conjecture 1.7] by proving the following.

Theorem 2.38. *Let G be a pseudofinite group. Then the identity component of any compactification C of G is abelian.*

All results of Chapter 2 are applications of theorems on generators and commutators in finite groups by Nikolov and Segal [**?**, 63, 65]. In Section 2.4, we also use results of Liebeck and Shalev [49] on bounded normal generation in finite simple groups.

In Chapter 3, we study the behavior of *word maps* on metric ultraproducts of finite quasisimple and (complex) unitary groups. Let $w \in \mathbf{F}_r$ be a non-trivial *word*, where $\mathbf{F}_r = \langle x_1, \ldots, x_r \rangle$ denotes the free group of rank $r \in \mathbb{N}$. Denote by $w(G) \subseteq G$ the image of the associated word map $w\colon G^r \to G$, i.e., $w(G) \coloneqq \{\varphi(w) \,|\, \varphi\colon \mathbf{F}_r \to G \text{ is a homomorphism}\}$. Let G be one of the finite groups S_n, $\mathrm{GL}_n(q)$, $\mathrm{Sp}_{2m}(q)$, $\mathrm{GO}^{\pm}_{2m}(q)$, $\mathrm{GO}_{2m+1}(q)$, $\mathrm{GU}_n(q)$ (q a prime power, $n \geq 2$, $m \geq 1$), or the unitary group U_n (over $\mathbb{C}$). Let d_G be the normalized Hamming metric resp. the normalized rank metric on G when G is a symmetric group resp. one of the other classical groups (see Definition 0.12) and write $n(G)$ for the permutation degree resp. the dimension of the natural module of G. We show the following density result.

Summary of Theorems 3.1, 3.2, and 3.3. *For $\varepsilon > 0$ there exists an integer $N(\varepsilon, w)$ such that $w(G)$ is ε-dense in G with respect to the metric d_G if $n(G) \geq N(\varepsilon, w)$, i.e., for all $g \in G$ we find $h \in w(G)$ such that $d_G(g, h) \leq \varepsilon$.*

This confirms metric versions of a conjecture of Shalev [3, Conjecture 8.3] and a conjecture posed by Larsen at the 2008 Meeting of the AMS in Bloomington. Equivalently, we prove that any non-trivial word map is surjective on a metric ultraproduct of groups G from above such that $n(G) \to \infty$ along the chosen ultrafilter.

As a consequence of our methods, we also obtain an alternative proof of the result of Hui, Larsen, and Shalev [36, Theorem 2.3] that $w_1(\mathrm{SU}_n)w_2(\mathrm{SU}_n) = \mathrm{SU}_n$ for non-trivial words $w_1, w_2 \in \mathbf{F}_r$ and n sufficiently large.

In Chapter 4, we discuss isomorphism questions for simple metric ultraproducts of certain families of finite groups which are 'close' to finite simple groups. This is motivated by the article [73] of Thom and Wilson.

Let $\mathcal{U}$ be an ultrafilter on an index set I. Denote by $\mathrm{S}_{\mathcal{U}}$ a metric ultraproduct of symmetric groups S_{n_i} $(i \in I)$ equipped with the normalized Hamming length function such that $n_i \to_{\mathcal{U}} \infty$. Note that $\mathrm{S}_{\mathcal{U}}$ is always simple by Lemma 3.2 of [49]. For X being one of the Lie types GL, Sp, GO, or GU denote by $X_{\mathcal{U}}$ a metric ultraproduct of groups $X_{n_i}(q_i)$ $(n_i \in \mathbb{Z}_+,\ q_i$ a prime power; $i \in I$; whether the groups are of plus or minus type in the orthogonal case in even dimension will not be important) equipped with the normalized rank length function such that $n_i \to_{\mathcal{U}} \infty$. Write $X_{\mathcal{U}}(q)$ for such an ultraproduct when q_i is eventually constant along $\mathcal{U}$, i.e., $\{i \in I \,|\, q_i = q\} \in \mathcal{U}$. Denote by $\overline{X}_{\mathcal{U}}$ resp. $\overline{X}_{\mathcal{U}}(q)$ the unique simple quotient of $X_{\mathcal{U}}$ resp. $X_{\mathcal{U}}(q)$. By Theorem 1.3(i) this is precisely the metric ultraproduct of the groups $\overline{X}_{n_i}(q_i)$ $(i \in I)$ equipped with the projective rank length function (see Definition 0.12). Here $\overline{X}$ denotes the projective Lie type associated to X, i.e., $\overline{X}$ equals PGL, PSp, PGO, or PGU when X is GL, Sp, GO, or GU, respectively.

In Theorem 2.2 of [73] Thom and Wilson prove that in this situation always $\mathrm{S}_{\mathcal{U}_1} \not\cong \overline{X}_{\mathcal{U}_2}$ (as abstract groups), where $\mathcal{U}_1$ and $\mathcal{U}_2$ are ultrafilters on different index sets I_1 and I_2. They also show that one can extract the limit characteristic $p := \lim_{\mathcal{U}} \mathrm{char}(\mathbb{F}_{q_i})$ of the finite fields $\mathbb{F}_{q_i}$ $(i \in I)$ out of the group $\overline{X}_{\mathcal{U}}$. In this chapter, we extend their results. The main result of Chapter 4 is the following.

Theorem 4.1. *Let $\overline{G}_1 \cong \overline{G}_2$ with $\overline{G}_j = \overline{X}_{j\mathcal{U}_j}(q_j)$, where $\overline{X}_j \in \{\mathrm{PGL}, \mathrm{PSp}, \mathrm{PGO}, \mathrm{PGU}\}$ $(j = 1, 2)$. Then it holds that $q_1 = q_2$. Also we must have $\overline{X}_1 = \overline{X}_2$ or $\{\overline{X}_1, \overline{X}_2\} = \{\mathrm{PSp}, \mathrm{PGO}\}$. Moreover, an ultraproduct $\overline{X}_{1\mathcal{U}_1}$ where the sizes q_i of the finite fields $\mathbb{F}_{q_i}$ $(i \in I_1)$ tend to infinity along $\mathcal{U}_1$ cannot be isomorphic to an ultraproduct $\overline{X}_{2\mathcal{U}_2}(q)$.*

We prove Theorem 4.1 by computing double centralizers of semisimple torsion elements in the above groups. It remains an open problem whether a group $\mathrm{PSp}_{\mathcal{U}_1}(q)$ can be isomorphic to a group $\mathrm{PGO}_{\mathcal{U}_2}(q)$ for an odd prime power q. However, due to the isomorphism $\mathrm{Sp}_{2m}(q) \cong \mathrm{GO}_{2m+1}(q)$ for $q = 2^e$, this is possible in characteristic two.

Chapters 1–3 are completely independent and only depend on the conventions and definitions of Chapter 0. They can be read in any order. Chapter 4 depends on Chapter 0, but also on Subsection 3.4.2 of Chapter 3. All mathematical symbols that occur are explained once again in the Index of Symbols. A version of Chapter 1 was published as [61]. A version of Chapter 2 was published as [55]. After Andreas Thom and I finished a first version of this article and circulated it among some experts, it was pointed out that (independently and slightly earlier) Lev Glebsky found a solution to Zilber's problem along the same lines. Chapter 3 was published as [62]. Chapter 4 is not yet published.

Acknowledgments

This research was supported by ERC Consolidator Grant No. 681207. I wish to thank my supervisor Andreas Thom for numerous valuable discussions. Andreas Thom and I thank Alessandro Carderi for interesting discussions regarding Chapter 2. We also thank Vadim Alekseev and Sebastian Manecke for nice conversations regarding Chapter 3.

Chapter 0

Notation, basic definitions, and facts

In this introductory chapter, we fix our notation and provide the reader with some basic definitions and facts, which are needed and used in the subsequent Chapters 1–4.

0.1 Group theory

For this section let G be a group and $S \subseteq G$ be a subset of G. We call S a *normal subset* if it is invariant under conjugation. The *neutral element* of G is denoted by 1_G, the *inverse* of $g \in G$ is written as g^{-1}, the *product* of $g, h \in G$ is written as gh, and the *conjugate* of g by h is written as $g^h := h^{-1}gh$. Write $g^G := \{g^h \mid h \in G\}$ for the *conjugacy class* of the element $g \in G$. Denote by $\mathrm{ord}(g)$ resp. $|G|$ the *order* of $g \in G$ resp. of the group G. Write $\exp(G)$ for the *exponent* of G. Denote by $\mathbf{1}$ the *trivial group*. Denote by C_k the *cyclic group of order* $k \in \mathbb{Z}_+$. Write $H \leq G$ for the statement 'H *is a subgroup of* G' and $N \trianglelefteq G$ for the statement 'N *is a normal subgroup of* G'. Write $H \cap I$ for the *intersection of subgroups* H and I of G. Let H be another group. Denote by $G \times H$ the *direct product* of the groups G and H. Write $\varphi\colon G \to H$ for a *group homomorphism* φ from G to H. Denote by $\mathrm{im}(\varphi)$ resp. $\ker(\varphi)$ the *image* resp. *kernel* of φ. Write $\varphi\colon G \hookrightarrow H$ resp. $\varphi\colon G \twoheadrightarrow H$ if φ is *injective* resp. *surjective*. Write $\overline{G} = G/N$ for the *quotient* of G by a normal subgroup N. Let $\overline{g}$ denote the image of $g \in G$ under the *natural homomorphism* $\pi\colon G \to \overline{G} = G/N$. Write $G \cong H$ if the groups G and H are *isomorphic*. Write $\mathrm{Aut}(G)$ for the *automorphism group* of G. Set $\langle S \rangle$ resp. $\langle\!\langle S \rangle\!\rangle_G$ to be the subgroup resp. normal subgroup of G *generated* by S. If the ambient group G is understood, the subscript in the second case is omitted. For $k \in \mathbb{N}$ set $S^{*k} := \{s_1 \cdots s_k \mid s_i \in S\}$ to be the k-fold product of the subset S, where the empty product S^{*0} equals the singleton containing the neutral element $\{1_G\}$. When $T \subseteq G$ is another subset of G, set $ST := \{st \mid s \in S, t \in T\}$. The symbol $\mathbf{F}$ always denotes a *free group*. Write $\mathrm{rk}(\mathbf{F})$ for the *rank* of $\mathbf{F}$. The symbol $\mathbf{F}_r$ denotes the free group of rank r freely generated by $x_1, \ldots, x_r$. Elements of a free group are also called *words*. A word $w \in \mathbf{F}$ is said to be *non-trivial* if $w \neq 1_{\mathbf{F}}$. We write $\langle S \mid R \rangle$

for the group *presented* by the *generators* S and *relations* R. If G is a permutation group acting on a set Ω, write $H \wr G = H^{\Omega} \rtimes G$ for the *wreath product* of the group H with G. Write $(h_\omega)_{\omega\in\Omega}.g \in H^{\Omega} \rtimes G$ for an element of this group. Write $N.H$ for an *extension* of the group N by the group H.

(a) Group actions. All *group actions* that occur are right actions. When G acts on the set X, we write $x.g$ for the image of $x \in X$ under the map associated to $g \in G$. However, when X is a ring and α is a *ring automorphism*, we write x^α instead of $x.\alpha$ (for aesthetic reasons; see Section 0.2 and Subsection 3.4.2). We write $\mathrm{stab}_G(x) \coloneqq \{g \in G \mid x.g = x\}$ for the *stabilizer* of x in G and $\mathrm{orb}_G(x) \coloneqq \{x.g \mid g \in G\}$ for the *orbit* of x under G.

(b) Commutators. For $g, h \in G$ we write $[g, h] \coloneqq g^{-1}h^{-1}gh$ for their *commutator*. For $g \in G$ set $[S, g] \coloneqq \{[s, g] \mid s \in S\}$ and $[g, S] \coloneqq \{[g, s] \mid s \in S\}$. For subgroups $H, L \leq G$ write $[H, L] \coloneqq \langle [h, l] \mid h \in H, l \in L\rangle$.

(c) Subgroups. We write $\mathbf{Z}(G)$ for the *center* of the group G and $\mathbf{C}_G(S)$ for the *centralizer* of the set S in G. We also write $\mathbf{C}(S)$ if the ambient group G is clear from the context. Write $G' \coloneqq [G, G]$ for the *commutator subgroup* and, more generally, for $i \in \mathbb{N}$ set $G^{(i+1)} \coloneqq [G^{(i)}, G^{(i)}]$ to be the $(i + 1)$st term in the *derived series* of G, where $G^{(0)} \coloneqq G$. For $i \in \mathbb{Z}_+$ write $\gamma_i(G)$ for the ith term in the *lower central series* of G, i.e., $\gamma_1(G) \coloneqq G$ and $\gamma_{i+1}(G) \coloneqq [\gamma_i(G), G]$. We set $\gamma_\omega(G) \coloneqq \bigcap_{i\in\mathbb{Z}_+} \gamma_i(G)$.

(d) Symmetric groups, alternating groups, and permutations. Write $\mathrm{Sym}(\Omega)$ resp. $\mathrm{Alt}(\Omega)$ for the *symmetric* resp. *alternating group* on the finite set Ω. Write id_Ω for the *identity permutation* on Ω. If Ω is clear from the context, we drop the subscript. Set $\mathrm{S}_n \coloneqq \mathrm{Sym}(\underline{n})$ resp. $\mathrm{A}_n \coloneqq \mathrm{Alt}(\underline{n})$ to be the symmetric resp. alternating group on the set $\underline{n} \coloneqq \{1, \ldots, n\}$. Write $\widetilde{\mathrm{A}}_n$ for the *Schur covering group* of A_n (see Section 2.7 of [78]; this symbol only appears in Section 1.3). Fix a permutation $\sigma \in \mathrm{S}_n$. Define the *support* $\mathrm{supp}(\sigma) \subseteq \underline{n}$ of σ to be the set $\{x \in \underline{n} \mid x.\sigma \neq x\}$. Similarly, define the *fixed point set* of σ by $\mathrm{fix}(\sigma) \coloneqq \{x \in \underline{n} \mid x.\sigma = x\}$. For $k \in \mathbb{Z}_+$ define $c_k(\sigma)$ to be the number of *k-cycles* of σ. Setting $c_k \coloneqq c_k(\sigma)$ for $k \in \mathbb{Z}_+$, we say that σ has *cycle type* $(k^{c_k})_{k\in\mathbb{Z}_+}$ (cf. Section 2.3.1 of [78]). Let $C_k(\sigma) \subseteq \mathrm{S}_n$ resp. $\Omega_k(\sigma) \subseteq \underline{n}$ denote the set of k-cycles resp. the support of the k-cycles of σ. In this situation we have that $\bigsqcup_{c\in C_k(\sigma)} \mathrm{supp}(c) = \Omega_k(\sigma)$ and $\bigsqcup_{k\in\mathbb{Z}_+} \Omega_k(\sigma) = \underline{n}$. Set $n_k(\sigma) \coloneqq |\Omega_k(\sigma)| = kc_k(\sigma)$ for $k \in \mathbb{Z}_+$. We call σ *k-isotypic* if $\Omega_k(\sigma) = \underline{n}$, or equivalently, $c_i(\sigma) = 0$ for all $i \neq k$. If σ is k-isotypic for one k, we call it *isotypic*.

(e) Vector spaces, linear maps, general linear groups, and special linear groups. Let k be a field. Fix a k-vector space V. The *zero vector* of V is written as 0, the *additive inverse* of $v \in V$ is written as $-v$, the *sum* of $u, v \in V$ is written as $u + v$, and the vector $v \in V$ *scaled* by $\lambda \in k$ is written as λv. Let $\dim(V)$ denote the *dimension* of V. Write $U \leq V$ resp. $U < V$ for the statement '*U is a k-vector subspace of V*' resp. '*U is a proper k-vector subspace of V*'. Write $\mathrm{Sub}(V)$ for the set of all k-vector subspaces of V. Let

codim(U) denote the *codimension* of the k-vector subspace U in V. Write $U \cap W$ for the *intersection of vector subspaces* U and W of V. Let U be another k-vector space. Denote by $U \oplus V$ the *direct sum* of U and V. We use the same notation for (injective or surjective) *k-linear maps* between k-vector spaces as for group homomorphisms (see above). We also use the same notation for the *image* and the *kernel* of a linear map as in the case of a group homomorphism. Write $\mathrm{rk}(\varphi) := \dim(\mathrm{im}(\varphi))$ for the *rank* of the linear map φ. For linear maps φ and ψ let $\varphi \oplus \psi$ be their direct sum. Write V/W for the *quotient* of V by the k-vector subspace W. Write $U \cong V$ if U and V are *isomorphic* as k-vector spaces. For vectors $v_1, \ldots, v_n \in V$ write $\langle v_1, \ldots, v_n \rangle \leq V$ for the k-vector subspace *generated* by $v_1, \ldots, v_n$.

Write $\mathbf{M}(V)$ for the ring of k-linear maps on V. Fix a linear map $g \in \mathbf{M}(V)$. Denote by $\det(g)$ resp. $\mathrm{tr}(g)$ the *determinant* resp. *trace* of g. Write $k[X]$ for the *polynomial ring* over k with one variable X. All *polynomials* in $k[X]$ that occur in the text are meant to be monic ones, i.e., the coefficient of the largest power of X is one. For such a polynomial $\chi = a_0 + a_1 X + \cdots + a_{k-1}X^{k-1} + X^k \in k[X]$ of degree k write $F(\chi)$ for the *Frobenius block* with characteristic polynomial χ, i.e., multiplication by $\overline{X}$ in the quotient ring $k[X]/(\chi)$ represented in the basis $1, \overline{X}, \ldots, \overline{X}^{k-1}$:

$$F(\chi) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & \cdots & \cdots & \cdots & -a_{k-1} \end{pmatrix}.$$

For a scalar $\lambda \in k$ write $J_e(\lambda)$ for the *Jordan block* of size $e \geq 1$ with eigenvalue λ, i.e., multiplication by $\lambda + \overline{X}$ in $k[X]/(X^e)$ in the basis $1, \overline{X}, \ldots, \overline{X}^{e-1}$:

$$J_e(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

Call the polynomial χ *primary* if it is a power i^e of an irreducible polynomial $i \in k[X]$ (or equivalently, the ideal $(\chi) \subseteq k[X]$ is a *primary ideal*). Recall that by the existence of the *generalized Jordan normal form* of g (sometimes also called *primary rational canonical form*), we can write g in the form

$$g \cong \bigoplus_{\chi \text{ primary}} F(\chi)^{\oplus c_\chi},$$

where the c_χ are uniquely determined. In this situation, for $\chi \in k[X]$ primary set $c_\chi(g) :=$

c_χ and write $V_\chi(g)$ for the g-invariant subspace U on which $g|_U$ by the above normal form acts as $F(\chi)^{\oplus c_\chi}$. Note that $V_\chi(g)$ is not uniquely determined. However, when we use the symbol $V_\chi(g)$ for $\chi \in k[X]$ primary, we mean that $V = \bigoplus_{\chi \text{ primary}} V_\chi(g)$ is some decomposition such that g acts $F(\chi)$-isotypically on each $V_\chi(g)$ for all $\chi \in k[X]$ primary. Also set $n_\chi(g) \coloneqq \dim(V_\chi(g)) = k_\chi c_\chi(g)$, where $k_\chi = \deg(\chi)$. Now drop the assumption that χ is primary. We call g *$F(\chi)$-isotypic* if $g \cong F(\chi)^{\oplus c_\chi}$ and *isotypic* if it is $F(\chi)$-isotypic for some (not necessarily primary) polynomial $\chi \in k[X]$.

Denote by $\mathrm{id}_V \in \mathbf{M}(V)$ the *identity map* on V. If the space V is understood, we drop the subscript. Write $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ for a *diagonal matrix* on $V = k^n$ with $\lambda_1, \dots, \lambda_n \in k$ on the main diagonal. For a k-algebra R write $R^\times$ for its *units*; so $k^\times = k \setminus \{0\}$. Write $\mathrm{P}\mathbf{M}(V) \coloneqq (\mathbf{M}(V) \setminus \{0\})/k^\times$ for the *projective space* associated to $\mathbf{M}(V)$. Set $\mathbf{M}_n(k) \coloneqq \mathbf{M}(k^n)$. Write $\mathrm{GL}(V) \coloneqq \mathbf{M}(V)^\times$ resp. $\mathrm{SL}(V) \coloneqq \ker(\det\colon \mathrm{GL}(V) \to k^\times)$ for the *general* resp. *special linear group* of V. Denote by $\mathrm{PGL}(V) \coloneqq \mathrm{GL}(V)/\mathbf{Z}(\mathrm{GL}(V)) = \mathrm{GL}(V)/k^\times$ resp. $\mathrm{PSL}(V) \coloneqq \mathrm{SL}(V)/\mathbf{Z}(\mathrm{SL}(V)) = \mathrm{SL}(V)/\{\lambda \in k^\times \mid \lambda^{\dim(V)} = 1\}$ the *projective general* resp. *projective special linear group* of V. Set $\mathrm{GL}_n(k) \coloneqq \mathrm{GL}(k^n)$, $\mathrm{SL}_n(k) \coloneqq \mathrm{SL}(k^n)$, $\mathrm{PGL}_n(k) \coloneqq \mathrm{PGL}(k^n)$, and $\mathrm{PSL}_n(k) \coloneqq \mathrm{PSL}(k^n)$.

(f) Classical groups of Lie type and vector spaces with form. Denote by $\mathbb{F}_q$ the *finite field* with q elements. Set $\mathrm{GL}_n(q) \coloneqq \mathrm{GL}_n(\mathbb{F}_q)$, $\mathrm{SL}_n(q) \coloneqq \mathrm{SL}_n(\mathbb{F}_q)$, $\mathrm{PGL}_n(q) \coloneqq \mathrm{PGL}_n(\mathbb{F}_q)$, and $\mathrm{PSL}_n(q) \coloneqq \mathrm{PSL}_n(\mathbb{F}_q)$ (cf. Section 3.3 of [78]).

Fix a vector space $V = k^n$ over the field k (which we specify in the subsequent cases). Set $p \coloneqq \mathrm{char}(\mathbb{F}_q)$ to be the *characteristic* of the field $\mathbb{F}_q$. Now we distinguish three cases (cf. Section 3.4.1 of [78]).

(i) The symplectic case. We set $k \coloneqq \mathbb{F}_q$ and let $f\colon V \times V \to k$ be an *alternating bilinear form*, i.e., f is bilinear and $f(v,v) = 0$ for all $v \in V$.

(ii) The orthogonal case. We set $k \coloneqq \mathbb{F}_q$ and let $f\colon V \times V \to k$ be a *symmetric bilinear form*, i.e., f is bilinear and $f(u,v) = f(v,u)$ for all $u, v \in V$. $(*)$ If $p = 2$, we require additionally that f comes from a *quadratic form* $Q\colon V \to k$, i.e., $Q(\lambda v) = \lambda^2 Q(v)$ and $Q(u+v) = Q(u) + f(u,v) + Q(v)$ for all $\lambda \in k$, $u, v \in V$. Note that this condition forces f to be alternating, since $Q(2v) = 0 = 2Q(v) + f(v,v) = f(v,v)$ for all $v \in V$.

(iii) The unitary case. We set $k \coloneqq \mathbb{F}_{q^2}$ and let $\sigma\colon k \to k$ denote the *q-Frobenius map* $x \mapsto x^q$ (which is an involution). Let $f\colon V \times V \to k$ be a *σ-conjugate-symmetric sesquilinear*, i.e., f is linear in the first entry and $f(u,v) = f(v,u)^\sigma$ for all $u, v \in V$.

For two vectors $u, v \in V$ write $u \perp v$ if $f(u,v) = 0$ and say that u and v are *perpendicular* in this case. Subsequently, let $U, W, Z \leq V$ be subspaces. Write $U \perp W$ and say that U and W are *perpendicular*, if $u \perp w = 0$ for all $u \in U$, $w \in W$. Write $U \perp W = Z$ if $U \perp W$ and $U \oplus W = Z$. In this situation Z is called the *orthogonal direct sum* of U and W. If we are not in Case $(*)$ from above, a vector $v \in V$ is called *isotropic* if $v \perp v$. In Case $(*)$, a vector $v \in V$ is called isotropic if $Q(v) = 0$ (as remarked above, here every vector fulfills $v \perp v$, since f is alternating). Write $U^\perp \coloneqq \{v \in V \mid u \perp v \text{ for all } u \in U\}$

for the *perpendicular space* of U. Define the *radical* $\mathrm{rad}(f)$ of the form f to be $V^\perp$. In Case $(*)$, the radical of Q is defined as $\mathrm{rad}(Q) := \mathrm{rad}(f) \cap \{v \in V \mid v \text{ is isotropic}\}$. Note that $\mathrm{rad}(Q)$ is a subspace, since Q is semilinear on $\mathrm{rad}(f)$ with respect to the 2-Frobenius map $k \to k$; $x \mapsto x^2$ (as it is a quadratic form). If we are not in Case $(*)$, we say that U is *non-singular* if $\mathrm{rad}(f|_U) = U \cap U^\perp = 0$ and U is called *totally isotropic* if $U \leq U^\perp$. Note that in this case, if U is non-singular, then $U \perp U^\perp = V$, so $U^\perp$ is actually a complement to U. In Case $(*)$, we call U non-singular if $\mathrm{rad}(Q|_U) = 0$ and totally singular if $Q|_U = 0$. A maximal totally isotropic subspace of V is called *Witt subspace* of V.

We say that f is *non-singular* if $\mathrm{rad}(f) = 0$. In Case $(*)$, we call Q non-singular if $\mathrm{rad}(Q) = 0$ and *non-degenerate* if even $\mathrm{rad}(f) = 0$. If g is a form like f (resp. R a form like Q in Case $(*)$), we say that f and g (resp. Q and R in Case $(*)$) are *linearly equivalent* if there is a linear map $h \in \mathrm{GL}(V)$ such that $f(u, v) = g(u.h, v.h)$ (resp. $Q(v) = R(v.h)$ in Case $(*)$) for all $u, v \in V$.

Let

$$\begin{aligned} \mathrm{GI}(V, f) &:= \{g \in \mathrm{GL}(V) \mid f(u.g, v.g) = f(u, v) \text{ for all } u, v \in V\} \\ &\text{resp.} \\ \mathrm{GI}(V, Q) &:= \{g \in \mathrm{GL}(V) \mid Q(v.g) = Q(v) \text{ for all } v \in V\} \end{aligned}$$

denote the *full isometry group* of (V, f) (resp. (V, Q) in Case $(*)$). Note that, when f and g (resp. Q and R) are linearly equivalent via $h \in \mathrm{GL}(V)$, we have an isomorphism $\varphi\colon \mathrm{GI}(V, f) \to \mathrm{GI}(V, g)$ (resp. $\varphi\colon \mathrm{GI}(V, Q) \to \mathrm{GI}(V, R)$); $a \mapsto a^h$. Subsequently, set $G := \mathrm{GI}(V, f)$ (resp. $G := \mathrm{GI}(V, Q)$ in Case $(*)$) and assume f (resp. Q in Case $(*)$) to be non-singular.

In Case (i), f being non-singular forces $n = 2m$ to be even. Also f is uniquely determined up to linear equivalence. Hence G is determined up to isomorphism and we call it the *symplectic group* of degree $2m$ over $\mathbb{F}_q$ and denote it by $\mathrm{Sp}_{2m}(q)$ (cf. Section 3.5 of [78]).

In Case (iii), f is again uniquely determined up to linear equivalence. Here we call G the *general unitary group* of degree n over $\mathbb{F}_{q^2}$ and denote it by $\mathrm{GU}_n(q)$ (cf. Section 3.6 of [78]). In Chapter 3, we shall also speak about the general resp. special unitary group over $\mathbb{C}$, which we denote by U_n resp. SU_n.

In Case (ii), assume first that we are not in Case $(*)$ (cf. Section 3.7 of [78]). There exist two equivalence classes of non-singular symmetric bilinear forms (cf. Section 3.4.6 of [78]). Let f be such a form. When $n = 2m + 1$ is odd and $\alpha \in \mathbb{F}_q^\times$ is a non-square, then f and αf are inequivalent, but apparently $G = \mathrm{GI}(V, f) = \mathrm{GI}(V, \alpha f)$. Denote this group G by $\mathrm{GO}_{2m+1}(q)$ and call it the *general orthogonal group* of degree $2m + 1$ over $\mathbb{F}_q$. When $n = 2m$ is even, there is a form f^+ of *plus type* (i.e., a Witt subspace of V is of dimension m) and a form f^- of *minus type* (i.e., a Witt subspace is of dimension

$m-1$). We set $\mathrm{GO}_{2m}^{+}(q) \coloneqq \mathrm{GI}(V, f^{+})$ resp. $\mathrm{GO}_{2m}^{-}(q) \coloneqq \mathrm{GI}(V, f^{-})$ and call these groups the *general orthogonal group of plus type* resp. *minus type* of degree $2m$ over $\mathbb{F}_q$.

Now assume that we are in Case $(*)$ (cf. Section 3.8 of [78]). If $n = 2m+1$ is odd, f must be singular, since it is alternating. Hence $\mathrm{rad}(f)$ has dimension one and we have a homomorphism $\pi\colon G \to \mathrm{GI}(V/\mathrm{rad}(f), \overline{f}) = \mathrm{Sp}_{2m}(q)$. It turns out that every map $g \in \mathrm{GI}(V/\mathrm{rad}(f), \overline{f}) = \mathrm{Sp}_{2m}(q)$ has a unique lift $\widetilde{g} \in G$, so that π is actually an isomorphism (cf. Section 3.4.7 of [78]) witnessing that $\mathrm{GO}_{2m+1}(q) \cong \mathrm{Sp}_{2m}(q)$. Hence we neglect this case and assume that $n = 2m$ is even and that $\mathrm{rad}(f) = 0$. In this case, there are again two equivalence classes of quadratic forms Q. One of plus type (i.e., a Witt subspace of V is of dimension m), which we call Q^{+}, and one of minus type (i.e., a Witt subspace is of dimension $m-1$), which we call Q^{-}. Set $\mathrm{GO}_{2m}^{+}(q) \coloneqq \mathrm{GI}(V, Q^{+})$ resp. $\mathrm{GO}_{2m}^{-}(q) \coloneqq \mathrm{GI}(V, Q^{-})$ and call the groups as above.

Now let $\mathrm{SU}_n(q) \coloneqq \mathrm{SL}_n(q) \cap \mathrm{GU}_n(q)$ denote the *special unitary group* of degree n over $\mathbb{F}_{q^2}$, $\mathrm{SO}_{2m+1}(q) \coloneqq \mathrm{SL}_{2m+1}(q) \cap \mathrm{GO}_{2m+1}(q)$ denote the *special orthogonal group* of degree $2m+1$ over $\mathbb{F}_q$, and $\mathrm{SO}_{2m}^{+}(q) \coloneqq \mathrm{SL}_{2m}(q) \cap \mathrm{GO}_{2m}^{+}(q)$ resp. $\mathrm{SO}_{2m}^{-}(q) \coloneqq \mathrm{SL}_{2m}(q) \cap \mathrm{GO}_{2m}^{-}(q)$ denote the *special orthogonal group of plus type* resp. *minus type* of degree $2m$ over $\mathbb{F}_q$. Note that $\mathrm{Sp}_{2m}(q) \leq \mathrm{SL}_{2m}(q)$, whence there is no special symplectic group. Write $\Omega_{2m+1}(q)$ resp. $\Omega_{2m}^{\varepsilon}(q)$ ($\varepsilon = \pm$) for the kernel of the *spinor norm* $\mathrm{SO}_{2m+1}(q) \to \mathbb{F}_q^{\times}/(\mathbb{F}_q^{\times})^2$ resp. $\mathrm{SO}_{2m}^{\varepsilon}(q) \to \mathbb{F}_q^{\times}/(\mathbb{F}_q^{\times})^2$ if p is odd (cf. Section 3.7.1 of [78]). In characteristic $p = 2$, use the same notation $\Omega_{2m}^{\varepsilon}(q)$ for the kernel of the *quasideterminant* $\mathrm{SO}_{2m}^{\varepsilon}(q) \to \{\pm 1\}$ (cf. Section 3.8.1 of [78]).

Let G denote any of the groups $\mathrm{Sp}_{2m}(q)$, $\mathrm{GO}_{2m+1}(q)$ (q odd), $\mathrm{SO}_{2m+1}(q)$ (q odd), $\mathrm{GO}_{2m}^{\varepsilon}(q)$, $\mathrm{SO}_{2m}^{\varepsilon}(q)$, $\Omega_{2m+1}(q)$ (q odd), $\Omega_{2m}^{\varepsilon}(q)$, $\mathrm{GU}_n(q)$, $\mathrm{SU}_n(q)$ ($m \geq 1$, $n \geq 1$, q a prime power, $\varepsilon = \pm$). In each of the above cases, call V the *natural module* of G. Set $\overline{G} \coloneqq G/\mathbf{Z}(G)$. Note that $G \leq \mathrm{GL}(V)$ and $\overline{G} \leq \mathrm{PGL}(V)$.

Then, if $G = \mathrm{Sp}_{2m}(q)$, call $\overline{G}$ the *projective symplectic group* of degree $2m$ over $\mathbb{F}_q$ and denote it by $\mathrm{PSp}_{2m}(q)$. If $G = \mathrm{GO}_{2m+1}(q)$ resp. $G = \mathrm{SO}_{2m+1}(q)$ (q odd) call $\overline{G}$ the *projective general* resp. *projective special orthogonal group* of degree $2m+1$ over $\mathbb{F}_q$ and denote it by $\mathrm{PGO}_{2m+1}(q)$ resp. $\mathrm{PSO}_{2m+1}(q)$. If $G = \mathrm{GO}_{2m}^{\varepsilon}(q)$ resp. $G = \mathrm{SO}_{2m}^{\varepsilon}(q)$ ($\varepsilon = \pm$) call $\overline{G}$ the *projective general* resp. *projective special orthogonal group of ε type* of degree $2m$ over $\mathbb{F}_q$ and denote it by $\mathrm{PGO}_{2m}^{\varepsilon}(q)$ resp. $\mathrm{PSO}_{2m}^{\varepsilon}(q)$. If $G = \Omega_{2m+1}(q)$ (q odd) resp. $G = \Omega_{2m}^{\varepsilon}(q)$ write $\mathrm{P}\Omega_{2m+1}(q)$ resp. $\mathrm{P}\Omega_{2m}^{\varepsilon}(q)$ for $\overline{G}$. Lastly, if $G = \mathrm{GU}_n(q)$ resp. $G = \mathrm{SU}_n(q)$, call $\overline{G}$ the *projective general* resp. *projective special unitary group* of degree n over $\mathbb{F}_{q^2}$ and denote it by $\mathrm{PGU}_n(q)$ resp. $\mathrm{PSU}_n(q)$.

We end up this section by stating an important result which we will use frequently in Chapters 1–4.

Lemma 0.1 (Witt's lemma)**.** *Assume f (resp. Q in Case $(*)$) is non-singular (resp. Q is even non-degenerate in Case $(*)$). Let $U, W \leq V$ be subspaces such that there is a partial isometry $\alpha\colon (U, f|_U) \to (W, f|_W)$ (resp. $\alpha\colon (U, Q|_U) \to (W, Q|_W)$ in Case $(*)$). Then α*

extends to a full isometry $\beta\colon (V,f) \to (V,f)$ (resp. $\beta\colon (V,Q) \to (V,Q)$ in Case $(*)$).

Remark 0.2. Lemma 0.1 (cf. Section 3.4.8 of [78]) implies that, if f is non-singular (resp. Q is non-degenerate in Case $(*)$), any totally isotropic subspace is contained in a Witt subspace and all Witt subspaces are conjugate (and hence have the same dimension).

0.2 Some ring and field theory

For Chapters 3 and 4 we need some basic ring and field theory.

Fix a ring R. Write 0 resp. 1 for the *neutral element* with respect to addition resp. multiplication in R. Write $a+b$ resp. ab for the *sum* resp. *product* of ring elements $a, b \in R$. Write a^{-1} for the *inverse element* of the unit $a \in R^\times$. Write (r) for the *principal ideal* generated by $r \in R$. Write rR for the *right ideal* generated by $r \in R$. Use the same notation for *ring homomorphisms* resp. *quotient rings* as for group homomorphisms resp. quotient groups. Write $\mathrm{Aut}(R)$ for the *automorphism group* of R. For a *ring automorphism* $\alpha \in \mathrm{Aut}(R)$ write r^α for the image of $r \in R$ under α. Sometimes we use the same notation when α is just an isomorphism between two different rings. Let $R_\alpha \coloneqq \{r \in R \mid r^\alpha = r\}$ denote the *fixed ring* of α. If $I \subseteq R$ is an ideal of R which is fixed setwise by α, write $I_\alpha \coloneqq \{i \in I \mid i^\alpha = i\} = I \cap R_\alpha \subseteq R_\alpha$ for the *fixed ideal* of α inside I. For a subset $S \subseteq R$ set $S^{+k} \coloneqq \{s_1 + \cdots + s_k \mid s_1, \ldots, s_k \in S\}$ and define $S^{+0} = \{0\}$. Also write $\mathbf{C}_R(S)$ for the *centralizer* of S in R. When the ring R is understood, we omit the subscript. Write $R[X_1, \ldots, X_n]$ for the *polynomial ring* over R in the commuting variables $X_1, \ldots, X_n$. For $r \in R[X]$, $r = r_0 + r_1X + \cdots r_{n-1}X^{n-1} + r_nX^n$ $(r_n \neq 0)$ write $\deg(r) \coloneqq n$ for its *degree*. Write $\Phi_n(X) \in \mathbb{Z}[X]$ for the nth *cyclotomic polynomial* $(n \in \mathbb{Z}_+)$. For a field k write $\overline{k}$ for its *algebraic closure*. Write $k[G]$ for the *group algebra* of G over the field k. For a finite field extension L/K write $\mathrm{tr}_{L/K}\colon L \to K$ for the *field trace*. Write $\mathrm{Gal}(L/K)$ for the *Galois group* of the *Galois extension* L/K.

0.3 Ultraproducts and norms

For this section fix an *index set* I and a *non-principal ultrafilter* $\mathcal{U}$ on I. We start with the notion of an *algebraic ultraproduct* of a family of groups. Note that this notion will only be used in Chapter 1.

Definition 0.3 (Algebraic ultraproduct)**.** *The* algebraic ultraproduct *of a family of (abstract) groups $\mathcal{H} = (H_i)_{i\in I}$ with respect to $\mathcal{U}$ is defined as the quotient*

$$\mathcal{H}_{\mathcal{U}} \coloneqq G/N_0,$$

where $G \coloneqq \prod_{i\in I} H_i$ and $N_0 \coloneqq \{(h_i)_{i\in I} \in G \mid \{i \in I \mid h_i = 1_{G_i}\} \in \mathcal{U}\}$.

Remark 0.4. The subset N_0 of G is a subgroup, since $\mathcal{U}$ is a filter. It is apparent that N_0 is normal.

In Chapters 2–4, we will switch from algebraic to metric ultraproducts. To introduce this notion, we first need to define what a *length function* (or *norm*) on a group is.

Definition 0.5 (Length function)**.** *A* length function $\ell_H\colon H \to [0,\infty]$ *on the group H is a function which obeys the following three properties:*

(i) $\ell_H(h) = 0$ *if and only if* $h = 1_H$ *(identity of indiscernibles);*

(ii) $\ell_H(h) = \ell_H(h^{-1})$ *for* $h \in H$ *(symmetry);*

(iii) $\ell_H(gh) \leq \ell_H(g) + \ell_H(h)$ *for* $g, h \in H$ *(triangle inequality).*

Call ℓ_H invariant*, if* $\ell_H(gh) = \ell_H(hg)$ *for all* $g, h \in H$*, i.e.,* ℓ_H *is invariant under the conjugation action of H on itself. An invariant length function is also called a* norm.

Remark 0.6. Note that there is a one-to-one correspondence between left-invariant metrics d_H (which may attend the value infinity) on H (i.e., metrics $d_H\colon H^2 \to [0,\infty]$ such that $d_H(fg, fh) = d_H(g,h)$ for all $f, g, h \in H$) and length functions ℓ_H on H via the identity $\ell_H(g^{-1}h) = d_H(g,h)$ for all $g, h \in H$. Throughout the thesis, we will indicate that a length function and a left-invariant metric correspond to each other by equipping them with the same subscript.

A length function $\ell_H\colon H \to [0,\infty]$ is invariant, i.e., a norm, if and only if d_H is right-invariant (and hence bi-invariant), i.e., $d_H(gf, hf) = d_H(g,h)$ for all $f, g, h \in H$. Throughout the thesis, all length functions that occur will be invariant. Also, subsequently, we do not add the group to the subscript, when it is clear from the context. By a *normed group* we mean a group H equipped with a norm.

Now we introduce the notion af a *metric ultraproduct* of normed groups.

Definition 0.7 (Metric ultraproduct)**.** *The* metric ultraproduct *of a family of normed groups* $\mathcal{H} = (H_i, \ell_i)_{i\in I}$ *with respect to* $\mathcal{U}$ *is defined as the quotient*

$$\mathcal{H}_{\mathcal{U}}^{\mathrm{met}} := G/N_\ell,$$

where $G := \prod_{i\in I} H_i$ *and* $N_\ell := \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_i(h_i) = 0\}$*. Then* $\mathcal{H}_{\mathcal{U}}^{\mathrm{met}}$ *is a complete metric group with respect to the norm* $\ell_{\mathcal{U}}(h) := \lim_{\mathcal{U}} \ell_i(h_i)$*, which is the* ultralimit *of the norms* ℓ_i $(i \in I)$*. Here* $h = \overline{(h_i)}_{i\in I}$*.*

Remark 0.8. The function $\ell_{\mathcal{U}}$ is well-defined, since $\mathcal{U}$ is an ultrafilter and $[0,\infty]$ is compact. It is invariant, since all ℓ_i $(i \in I)$ are invariant. Also N_ℓ is well-defined and normal, since all ℓ_i $(i \in I)$ are invariant.

Some authors use a slightly different definition by restricting G to the sequences $(h_i)_{i\in I} \in \prod_{i\in I} H_i$ of uniformly bounded length, i.e., $\sup_{i\in I} \ell_i(h_i) < \infty$. However, we prefer the above definition, since then the metric ultraproduct is always a quotient of the product of the finite groups H_i $(i \in I)$. For more details on the algebraic and geometric structure of such ultraproducts see also [67] and [73, **?**].

Remark 0.9. Another thing we mention here is that the tuple $(\mathcal{H}_\mathcal{U}, \ell_\mathcal{U})$ in Definition 0.7 is always a metric group (i.e., the group operations are continuous), since $\ell_\mathcal{U}$ is invariant.

Remark 0.10. In Chapter 4, we shall also use ultraproducts $(X_\mathcal{U}^{\mathrm{met}}, d_\mathcal{U})$ of metric spaces (X_i, d_i) ($i \in I$; namely for $(X_i, d_i) = (\mathbf{M}_{n_i}(k), d_{\mathrm{rk}})$ or $(X_i, d_i) = (\mathrm{P}\mathbf{M}_{n_i}(k), d_{\mathrm{pr}})$). Such a metric ultraproduct is defined as $X_\mathcal{U}^{\mathrm{met}} := \prod_{i\in I} X_i/\sim$, where $(x_i)_{i\in I} \sim (y_i)_{i\in I}$ if and only if $\lim_\mathcal{U} d_i(x_i, y_i) = 0$. Also the metric $d_\mathcal{U}$ is defined as $d_\mathcal{U}(x, y) = \lim_\mathcal{U} d_i(x_i, y_i)$ (where $x = \overline{(x_i)}_{i\in I}, y = \overline{(y_i)}_{i\in I} \in X_\mathcal{U}^{\mathrm{met}}$).

Remark 0.11. Lastly, we remark that the algebraic ultraproduct of the family $\mathcal{H} = (H_i)_{i\in I}$ of finite groups is isomorphic to the metric ultraproduct of these groups equipped with the *discrete length function* (see Definition 0.12 below) with respect to the same ultrafilter. In this sense we can view every algebraic ultraproduct as a metric ultraproduct.

We end this chapter by introducing the most common length functions on finite, linear, and projective linear groups, and showing that some of them are '*equivalent*' on some infinite families of finite groups.

Definition 0.12 (Common length functions)**.** *On every group H we define the* discrete length function *by*

$$\ell_{\mathrm{d}}(h) := 1 - \delta_{1_H,h} = \begin{cases} 0 & \text{if } 1_H = h \\ 1 & \text{otherwise} \end{cases}.$$

On every finite centerless group H we define the normalized conjugacy length function *by*

$$\ell_{\mathrm{c}}(h) := \frac{\log|h^H|}{\log|H|}.$$

On the finite symmetric group S_n ($n \in \mathbb{Z}_+$) we define the normalized Hamming length function *by*

$$\ell_{\mathrm{H}}(\sigma) := \frac{1}{n}|\mathrm{supp}(\sigma)|.$$

On the general linear group $\mathrm{GL}_n(k) = \mathrm{GL}(V)$, where $V = k^n$ is a k-vector space, we define the normalized rank length function *by*

$$\ell_{\mathrm{rk}}(h) := \frac{1}{n}\,\mathrm{rk}(\mathrm{id}_V - h).$$

On the projective linear group $\mathrm{PGL}_n(k) = \mathrm{PGL}(V)$ we define the normalized projective rank length function *by*

$$\ell_{\mathrm{pr}}(\overline{h}) := \min\{\ell_{\mathrm{rk}}(h) \,|\, h \text{ a lift of } \overline{h}\}.$$

Finally, the Cayley length function *with respect to some subset $S \subseteq H$ is defined by*

$$\ell_{\mathrm{Cay},S}(h) := \min\{n \in \mathbb{N} \,|\, h = s_1 \cdots s_n \text{ for } s_i \in S \cup S^{-1}\} \cup \{\infty\}.$$

It is a simple matter to check that all of the above length functions, apart from the Cayley length function, are indeed norms. The latter is invariant if and only if S is invariant under conjugation. Also note that ℓ_{H} can be pulled back to any subgroup of S_n (e.g., to A_n) and, similarly, ℓ_{rk} resp. ℓ_{pr} can be pulled back to any subgroup of $\mathrm{GL}_n(k)$ resp. $\mathrm{PGL}_n(k)$ (e.g., when $k = \mathbb{F}_q$ or $\mathbb{F}_{q^2}$, to the groups G and $\overline{G}$ in Section 0.1(f)). By abuse of notation, we still write ℓ_{H}, ℓ_{rk}, ℓ_{pr} for these norms restricted to the smaller subgroup (cf. Theorem 1.3), and d_{H}, d_{rk}, d_{pr} for the corresponding metrics (as remarked in Remark 0.6). If we want to emphasize the group on which these norms resp. metrics are defined, we add it to the subscript (e.g., $\ell_{\mathrm{rk},H}\colon H \to [0,\infty]$).

We call a family of norms $(\ell_i)_{i\in I}$ on a sequence of finite groups $\mathcal{H} = (H_i)_{i\in I}$ *Lipschitz continuous* with respect to a second such family $(\ell'_i)_{i\in I}$ on the same groups if there is $L > 0$ such that $\ell_i \leq L\ell'_i$ $(i \in I)$.

For example, since $\ell_{\mathrm{c},H} \leq \ell_{\mathrm{d},H}$ for any finite centerless group H, the normalized conjugacy length function is Lipschitz continuous with respect to the discrete length function on such groups (with $L = 1$).

If $(\ell'_i)_{i\in I}$ is also Lipschitz continuous with respect to $(\ell_i)_{i\in I}$, we call these families *Lipschitz equivalent*.

Note the following fact (which we use in the argument at the end of Section 2.4).

Fact 0.13. *The length functions ℓ_{H}, $\frac{1}{n}\ell_{\mathrm{Cay},\tau^{\mathrm{S}_n}}$, and ℓ_{c} are Lipschitz equivalent on all finite symmetric or alternating groups S_n or A_n $(n \geq 5)$, where $\tau \in \mathrm{S}_n$ is a transposition. The length functions ℓ_{pr} and ℓ_{c} are Lipschitz equivalent on the classical simple groups $\mathrm{PSL}_n(q)$, $\mathrm{PSp}_{2m}(q)$, $\mathrm{P\Omega}_{2m+1}(q)$ $(q$ odd$)$, $\mathrm{P\Omega}^{\pm}_{2m}(q)$, $\mathrm{PSU}_n(q)$ $(n, m \in \mathbb{Z}_+$, q a prime power$)$.*

Proof. The first statement was proven in Proposition 2.7 and Theorem 2.11 of [67].

The second statement follows from Lemmas 5.3, 5.4, 6.3, 6.4, and the end of Section 7 of [49]. Indeed, Lemma 5.4, Lemma 6.4, and the last four lines of Section 7 of this article imply that the normalized projective rank length function on the above families of groups is as small as possible (up to a multiplicative constant). Lemma 5.3, Lemma 6.3, and the last inequality of Section 7 show that the normalized conjugacy length function is Lipschitz continuous with respect to the normalized projective rank length function. These two facts imply that both must be Lipschitz equivalent. The proof is complete. □

Chapter 1

The normal subgroup lattice of an algebraic ultraproduct of classical groups of Lie type

1.1 Introduction

The purpose of this chapter is to correct statements from the work of Stolz and Thom in [67] and to generalize them to a setting of quasisimple groups. As stated, Theorem 3.9 of [67] is not correct and the main result of this chapter (Theorem 1.3) should replace it. Note that already in [12] it was pointed out that some of the techniques and results of [67] were flawed. Some corrections on results about bounded normal generation in the setting of unitary groups on finite-dimensional Hilbert spaces can be found in [12]. The statement of [67, Theorem 4.20] should be considered as an open problem at the time of writing. In this chapter, we focus entirely on the case of finite groups.

Using the results of [49], it is a simple matter to prove the following result about 'relative' bounded normal generation for the alternating groups A_n with $n \geq 5$.

Lemma 1.1. *There exists $c > 0$ such that for any $S, T \subseteq \mathrm{A}_n$ $(n \geq 5)$ normal subsets with $|S|, |T| \geq 1$, $T \neq \{\mathrm{id}\}$ for any integer*

$$k \geq c \max\{\log|S|/\log|T|, 1\}$$

it holds that $S \subseteq (T^{\mathrm{A}_n})^{*k}$.

Proof. Use Corollary 2.4 of [49] to reduce to the case when T is a single conjugacy class. Hence, subsequently, we may assume that T is a single conjugacy class. Now assume the result holds for S being a single conjugacy class C. Then, in the more general case where S is arbitrary normal and T a single conjugacy class, we have for all conjugacy classes $C \subseteq S$ that for $k \geq c \max\{\log|S|/\log|T|, 1\} \geq c \max\{\log|C|/\log|T|, 1\}$ it holds that $C \subseteq (T^{\mathrm{A}_n})^{*k}$ and so $S \subseteq (T^{\mathrm{A}_n})^{*k}$. Hence we may assume that both S and T are conjugacy classes.

The result in this case follows from Lemma 1.8 below and the Lipschitz equivalence of the normalized Hamming length function ℓ_{H} and the normalized conjugacy length function ℓ_{c} on the family of finite alternating groups of degree at least five (see Fact 0.13). □

A straightforward consequence of the previous lemma is that for any $\sigma, \tau \in \mathrm{A}_n$ ($n \geq 5$) either $\sigma \in (\tau^{\mathrm{A}_n})^{*k}$ or $\tau \in (\sigma^{\mathrm{A}_n})^{*k}$ for any integer $k \geq c$, which easily implies that the normal subgroups of an algebraic ultraproduct of alternating groups are linearly ordered (which was first observed in [17]).

However, in the case of classical finite simple groups of Lie type, all of the above is false in general. The prototype of a counterexample is given by the following family of pairs of elements.

Example 1.2. Let $h_1, h_2 \in H \coloneqq \mathrm{SL}_q(q) \cong \mathrm{PSL}_q(q)$ be the elements given by $h_1 = \mathrm{diag}(1, \lambda, \ldots, \lambda)$ and $h_2 = \mathrm{diag}(1, \mu, \ldots, \mu)$, where $\lambda, \mu \in \mathbb{F}_q^\times$ are arbitrary such that $\lambda \notin \langle \mu \rangle$ and $\mu \notin \langle \lambda \rangle$ (e.g., take $\lambda = \zeta^a$ and $\mu = \zeta^b$, where $\langle \zeta \rangle = \mathbb{F}_q^\times$ and $a, b > 1$ are coprime with $q - 1 = ab$). Then it is easy to show by induction that h_1^k has eigenvalue λ^k and h_2^k has eigenvalue μ^k for $|k| < q$. From the assumptions it follows that $h_1 \notin (h_2^H \cup (h_2^{-1})^H)^{*k}$ and $h_2 \notin (h_1^H \cup (h_1^{-1})^H)^{*k}$ for any such k.

The example shows that Lemma 3.12 in [67] is false. The problem in its proof is that the rank and Jordan length (in our notation $n\ell_{\mathrm{rk}}$ and $n\ell_{\mathrm{pr}}$) of g and h do not always coincide. The correct replacement of this lemma is Lemma 1.13 below.

This example already implies that the normal subgroups of an algebraic ultraproduct of finite simple groups of type $\mathrm{SL}_q(q)$ are not linearly ordered. In this chapter, we shall prove the 'best possible' result on relative bounded normal generation in classical finite quasisimple groups of Lie type and fully describe the lattice of normal subgroups of an algebraic ultraproduct of the universal such, i.e., quasisimple groups which are the Schur covering group of a finite non-abelian simple group. We will prove the following theorem.

Theorem 1.3. *Let $(H_i)_{i\in I}$ be a sequence of universal finite quasisimple groups of Lie type (i.e., H_i is of type $\mathrm{SL}_n(q)$, $\mathrm{Sp}_{2m}(q)$, the double cover of $\Omega_{2m+1}(q)$ or $\Omega_{2m}^{\pm}(q)$ for q odd, $\Omega_{2m}^{\pm}(q)$ for q even, or $\mathrm{SU}_n(q)$ for suitable prime power q and $m, n \in \mathbb{Z}_+$; cf. the list at the beginning of Section 1.2), each endowed with the norms ℓ_{rk}, and $\overline{H}_i \coloneqq H_i/\mathbf{Z}(H_i)$ endowed with ℓ_{pr} (see Definition 0.12). Write V_i for the natural module of H_i ($i \in I$; see Section 0.1(f)). Assume that the dimensions $n_i \coloneqq n(H_i) \coloneqq \dim(V_i)$ tend to infinity along a fixed ultrafilter $\mathcal{U}$ on I. Set $G \coloneqq \prod_{i\in I} H_i$ and define the normal subgroups $N_{\mathrm{rk}} \coloneqq \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_{\mathrm{rk}}(h_i) = 0\}$ and $N_{\mathrm{pr}} \coloneqq \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_{\mathrm{pr}}(\overline{h}_i) = 0\}$ of G (where $\overline{h}_i$ means the image of $h_i \in H_i$ in $\overline{H}_i$). We also define $N_0 \coloneqq \{(h_i)_{i\in I} \in G \mid h_i = 1_{H_i}$ along $\mathcal{U}\}$ and suitable subgroups $N_1, A_0, A_1 \leq G$ defined in Section 1.4. Then the following hold.*

(i) The subgroup N_{pr} contains all proper normal subgroups of G containing N_0. In particular, G/N_{pr} is non-abelian simple and N_{pr}/N_0 is a characteristic subgroup of G/N_0.

(ii) *The normal subgroups of* G *lying between* N_0 *and* N_{rk} *are linearly ordered. Any such normal subgroup is perfect.*

(iii) *Define maps between the following two sets*

$$\begin{array}{c} \{N \trianglelefteq G \mid N_1 \leq N \leq N_{\mathrm{pr}}\} \\ {\scriptstyle\Psi}\big\uparrow\big\downarrow{\scriptstyle\Phi} \\ \{M \mid M \trianglelefteq G,\ N_1 \leq M \leq N_{\mathrm{rk}}\} \times \{A \mid A \trianglelefteq G,\ N_1 \leq A \leq A_1\} \end{array}$$

by $\Phi\colon N \mapsto (N \cap N_{\mathrm{rk}}, N \cap A_1)$ *and* $\Psi\colon (M, A) \mapsto MA$. *Then* Φ *and* Ψ *are isomorphisms of posets and mutually inverse to each other.*

(iv) *If* N *is normal in* G *containing* N_0 *and* $N_1 \not\leq N$, *then* $N \leq A_0$.

(v) *A normal subgroup* N *containing* N_0 *but not* N_1 *is contained in a normal subgroup* $K \leq N_{\mathrm{pr}}$ *containing* N_1 *if and only if* $N \leq A$, *where* $\Phi(K) = (M, A)$ *is the image of* K *under the map* Φ *from above.*

1.2 Auxiliary geometric results

In this section, we provide the necessary geometric results for the rest of this chapter. Subsequently, let H be a quasisimple group from the following list:

(i) linear: $\mathrm{SL}_n(q)$, $n \geq 2$, $(n, q) \neq (2, 2), (2, 3)$;

(ii) symplectic: $\mathrm{Sp}_{2m}(q)$, $m \geq 2$, $(m, q) \neq (2, 2)$;

(iii) orthogonal: $\Omega_{2m+1}(q)$, $m \geq 3$, q odd; $\Omega^{\pm}_{2m}(q)$, $m \geq 4$;

(iv) unitary: $\mathrm{SU}_n(q)$, $n \geq 3$, $(n, q) \neq (3, 2)$.

Remark 1.4. Here we omit the groups $\Omega_{2m+1}(q)$ with q even as they are isomorphic to the groups $\mathrm{Sp}_{2m}(q)$ (see Section 0.1(f)).

Use the notation of Section 0.1(f): Let V be the natural module of H (there are two such representations when $H = \Omega_{2m+1}(q)$ for q odd, corresponding to the two equivalence classes of non-singular symmetric bilinear forms in this dimension) and $n := \dim(V)$. For Cases (ii)–(iv) denote by f the corresponding non-singular alternating bilinear, symmetric bilinear, or conjugate-symmetric sesquilinear form. In Case (iii) for q even, denote by Q the corresponding quadratic form inducing the non-singular alternating form f.

As we ignore the case that $H = \mathrm{GO}_{2m+1}(q)$ with q even, a subspace U of V is non-singular if $f|_U$ is non-singular. The following fact will be used in the Section 1.4.

Lemma 1.5. *If* $U \leq V$ *is a non-singular subspace with* $2 \leq \dim(U) < n/2$, *then there exists a perpendicular decomposition* $W_1 \perp W_2 = U^{\perp}$ *and an element* $h \in H$ *such that* h *is the identity on* W_2 *and interchanges* U *and* W_1.

Proof. Explicit computations with standard bases show that $U^\perp$ always contains a subspace W_1 isometric to U via an isometry $\theta\colon U \to W_1$ with respect to f or Q in the orthogonal case in characteristic two (here we need $\dim(U^\perp) > \dim(U)$). Set $W_2 :=(U \perp W_1)^\perp$. Then $h_1\colon V = U \perp W_1 \perp W_2 \to U \perp W_1 \perp W_2 = V$ given by $(u, w_1, w_2) \mapsto (\theta^{-1}(w_1), \theta(u), w_2)$ is an isometric involution of V. Hence we may take $h := h_1$ in the symplectic case.

In the unitary case, if $\det(h_1) = \varepsilon \in \{\pm 1\}$, letting h_2 be the linear map which scales a non-isotropic vector of U by ε and fixes its perpendicular complement, then $h := h_1h_2$ works.

In the orthogonal case for q odd, define h_2 as in the unitary case. If the spinor norm of h_1h_2 is $\varepsilon \in \{\pm 1\}$, find an element $s \in \mathrm{SO}(U)$ of spinor norm ε and set $h_3 := s \oplus \mathrm{id}_{U^\perp}$ (for the existence of s we use that $\dim(U) \geq 2$). Then $h := h_1h_2h_3$ works.

In the orthogonal case for q even (forcing $\dim(V)$ to be even), if the quasideterminant of h_1 is $\varepsilon \in \{\pm 1\}$, find an element $s \in \mathrm{GO}(U)$ of quasideterminant ε and set $h_4 := s \oplus \mathrm{id}_{U^\perp}$ (again using $\dim(U) \geq 2$). Then $h := h_1h_4$ is an appropriate choice. The proof is complete. □

The following simple result about the existence of certain non-singular subspaces will also be used in later sections.

Lemma 1.6. *Let $U \leq V$, $\dim(U) = l$. Then there exists $W \leq U$ non-singular such that $\dim(W) \geq 2l - n$.*

Proof. Choose $W \leq U$ maximal non-singular. If there were two vectors $u, v \in W^\perp \cap U$ with $f(u, v) \neq 0$, then $W \perp \langle u, v\rangle > W$ would still be non-singular. Hence f is zero on this subspace. Moreover, for dimension reasons $U = W \perp (W^\perp \cap U)$. Together this implies $U^\perp \cap U = W^\perp \cap U$. Hence $\dim(W^\perp \cap U) = \dim(U) - \dim(W) \leq \dim(U^\perp) = n - l$, so $\dim(W) \geq 2l - n$, as desired. □

We introduce the group of quasiscalars $S(H)$ of H by

$$\begin{aligned}
S(\mathrm{SL}_n(q)) &:= \mathbb{F}_q^\times \cong \mathrm{C}_{q-1};\\
S(\mathrm{Sp}_{2m}) &:= \{\pm 1\};\\
S(\Omega_{2m+1}(q)) &:= \{\pm 1\} \quad (q \text{ odd});\\
S(\Omega_{2m}^\pm(q)) &:= \{\pm 1\};\\
S(\mathrm{SU}_n(q)) &:= \{x \in \mathbb{F}_{q^2}^\times \mid x^{q+1} = 1\} \cong \mathrm{C}_{q+1}\,.
\end{aligned}$$

Of course $-1 = 1$ if $\mathrm{char}(\mathbb{F}_q) = 2$.

Our last auxiliary result will be used in Section 1.4:

Lemma 1.7. *For any $\lambda \in S(H)$ there is a diagonalizable element $h \in H$ for a suitable basis of V such that all but two of its diagonal entries are equal to λ.*

Proof. Clearly, we may assume that $\lambda \neq 1$, since otherwise we can always take $h = \mathrm{id}_V$.

If $H = \mathrm{SL}_n(q)$, take $h = \mathrm{diag}(\lambda, \ldots, \lambda, \lambda^{-(n-1)})$ with respect to any basis of V.

If $H = \mathrm{Sp}_{2m}(q)$, q odd, and $\lambda = -1$, take $h = -\mathrm{id}_V$.

If $H = \mathrm{SU}_n(q)$, take $h = \mathrm{diag}(\lambda, \ldots, \lambda, \lambda^{-(n-1)})$ with respect to an orthonormal basis $e_1, \ldots, e_n$ for f, i.e., $f(e_i, e_j) = \delta_{ij}$ for $i, j = 1, \ldots, n$.

For $H = \Omega_{2m+1}(q)$, q odd, and $\lambda = -1$ take $h = \mathrm{diag}(-1, \ldots, -1, 1)$ with respect to a basis $e_1, \ldots, e_{2m+1}$ such that $f(e_i, e_j) = 0$ if $i \neq j$ and $f(e_i, e_i) = 1$ for $i = 1, \ldots, 2m$ and $f(e_{2m+1}, e_{2m+1}) = 1$ or a non-square $\alpha \in \mathbb{F}_q^\times$ (there are two equivalence classes of such forms both giving a natural representation of $\Omega_{2m+1}(q)$). Clearly, h has determinant and spinor norm equal to one.

For $H = \Omega^\pm_{2m}(q)$, q odd, and $\lambda = -1$ we find a basis $e_1, \ldots, e_{2m}$ such that either $f(e_i, e_j) = \delta_{ij}$ for all $i, j = 1, \ldots, 2m$, or $f(e_i, e_j) = 0$ for $i \neq j$, $f(e_i, e_i) = 1$ for $i = 1, \ldots, 2m-1$ and $f(e_{2m}, e_{2m}) = \alpha \in \mathbb{F}_q^\times$ a non-square (there are two equivalence classes of non-singular symmetric bilinear forms corresponding to the two non-isomorphic groups $\Omega^+_{2m}(q)$ and $\Omega^-_{2m}(q)$). In either case, we can take $h = \mathrm{diag}(-1, \ldots, -1, 1, 1)$ with respect to this basis, which has determinant and spinor norm one.

In all remaining cases we have $S(H) = \mathbf{1}$, so the proof is complete. □

1.3 Relative bounded normal generation in universal finite quasisimple groups

In this section, we keep the notation from Section 1.2. Recall that our group H carries the norm ℓ_{rk}, and $\overline{H} \coloneqq H/\mathbf{Z}(H)$ carries the Lipschitz equivalent norms ℓ_{pr}, ℓ_{c} (see Definition 0.12 and Fact 0.13).

In the following, we will prove a version of 'relative' bounded normal generation for all universal finite quasisimple groups from families of 'unbounded rank' (where we mean the permutation degree in the alternating case; for the others there is no such notion). We start with the alternating case. Recall that A_n ($n \geq 5$) is equipped with the Lipschitz equivalent norms ℓ_{H} and ℓ_{c} (see Definition 0.12 and Fact 0.13). Recall from Section 0.1(d) that $\widetilde{\mathrm{A}}_n$ denotes the Schur covering group of A_n.

Lemma 1.8. *There exists a constant $c > 0$ such that for any $\sigma \in \widetilde{\mathrm{A}}_n$ ($n \geq 5$), $\tau \notin \mathbf{Z}(\widetilde{\mathrm{A}}_n)$ for any integer $k \geq c \max\{\ell_{\mathrm{H}}(\overline{\sigma})/\ell_{\mathrm{H}}(\overline{\tau}), 1\}$ we have $\sigma \in (\tau^{\widetilde{\mathrm{A}}_n})^{*k}$. Here $\overline{\sigma}, \overline{\tau} \in \mathrm{A}_n$ are the images of $\sigma, \tau \in \widetilde{\mathrm{A}}_n$ under the canonical homomorphism $\widetilde{\mathrm{A}}_n \twoheadrightarrow \mathrm{A}_n$.*

Proof. We prove the lemma for A_n ($n \geq 5$) and derive the corresponding result for its Schur covering group. Let $\sigma, \tau \in \mathrm{A}_n$, $\tau \neq \mathrm{id}$. After conjugating σ, we may assume that either $\mathrm{supp}(\sigma) \subseteq \mathrm{supp}(\tau)$ or the opposite inclusion holds. In the first case, $\sigma \in \mathrm{Alt}(\mathrm{supp}(\tau)) = (\tau^{\mathrm{Alt}(\mathrm{supp}(\tau))})^{*c_1} \subseteq (\tau^{\mathrm{A}_n})^{*c_1}$ for some integral constant $c_1 > 0$, and in the second case, $\sigma \in \mathrm{Alt}(\mathrm{supp}(\sigma)) = (\tau^{\mathrm{Alt}(\mathrm{supp}(\sigma))})^{*k} \subseteq (\tau^{\mathrm{A}_n})^{*k}$ for any integer $k \geq c_2/\ell_{\mathrm{H},\mathrm{Alt}(\mathrm{supp}(\sigma))}(\tau) = c_2 \ell_{\mathrm{H}}(\sigma)/\ell_{\mathrm{H}}(\tau)$ with some constant $c_2 > 0$; both times we use Lemma 3.2 of [49]. So we may take $c \coloneqq \max\{c_1, c_2\}$.

The statement now extends to the Schur cover $\widetilde{\mathrm{A}}_n$ by the following argument. W.l.o.g., $n \geq 8$ (we may neglect this finite data), so $\widetilde{\mathrm{A}}_n$ is a twofold cover of A_n. It follows from the standard construction of the two double covers of S_n that, when a and b are lifts of the transposition (12) and (34), then $(ab)^2 = z$ is the unique non-trivial central element in one of these. This shows that $\mathbf{Z}(\widetilde{\mathrm{A}}_n) \subseteq (\tau^{\widetilde{\mathrm{A}}_n})^{*c_3}$ for any non-central $\tau \in \widetilde{\mathrm{A}}_n$ and some absolute integer constant $c_3 > 0$, which implies the claim for $\widetilde{\mathrm{A}}_n$. □

Remark 1.9. According to Fact 0.13 by Lipschitz equivalence we can replace ℓ_{H} by ℓ_{c} in the lemma.

Now we turn to the universal finite quasisimple groups of Lie type from families of unbounded rank. The proof of 'relative' bounded normal generation for these is actually very similar as for the Schur covering groups of the alternating groups.

We need the following fact which can be deduced by adapting the proof of Lemma 4.1 of [49] to quasisimple groups and looking at Lemmas 5.4 and 6.4, and the end of Section 7 of the same article.

Lemma 1.10. *There is an absolute constant $c > 0$ (independent of H) such that for $h \in H \setminus \mathbf{Z}(H)$, for any $k \geq c/\ell_{\mathrm{pr}}(\overline{h})$ it holds that $H = (h^H)^{*k}$. Here $\overline{h}$ denotes the image of h in $\overline{H}$.*

Remark 1.11. Again by Lipschitz equivalence we can replace ℓ_{pr} by ℓ_{c}.

We will also make use of the following Proposition 2.13 from [67], which we state here without proof.

Lemma 1.12 (Proposition 2.13 of [67])**.** *For $h \in H$, if $\ell_{\mathrm{rk}}(h) = \delta$, then $\ell_{\mathrm{pr}}(\overline{h}) \geq \min\{\delta, 1-\delta\}$. Here $\overline{h}$ is the image of h in $\overline{H}$.*

Here is now the promised result for almost all other universal finite quasisimple groups from families of unbounded rank:

Lemma 1.13. *Let $\varepsilon > 0$ be arbitrary. There exists an absolute constant $C > 0$ and a constant $D > 0$ only depending on ε such that the following holds: Let $h_1 \in H \setminus \mathbf{Z}(H)$, $h_2 \in H$, $\ell_{\mathrm{rk}}(h_1) \leq 1-\varepsilon$. Then $h_2 \in (h_1^H)^{*k}$ for all integers $k \geq \max\{C\ell_{\mathrm{rk}}(h_2)/\ell_{\mathrm{rk}}(h_1), D\}$.*

Proof. First assume that $\varepsilon \leq \ell_{\mathrm{rk}}(h_1) \leq 1-\varepsilon$. Then by Proposition 2.13 of [67] it holds that $\ell_{\mathrm{pr}}(\overline{h}_1) \geq \varepsilon$, so by Lemma 1.10 there is $D \in \mathbb{N}$ only depending on ε such that $(h_1^H)^{*D} = H$.

So we may assume, w.l.o.g., that $\ell_{\mathrm{rk}}(h_1) < \varepsilon \leq 1/8$. Assume additionally that $\ell_{\mathrm{rk}}(h_2) \leq 1/8$ as well (we will remove this assumption at the end). Set $U_i := \ker(1-h_i) \leq V$ and $l_i := \dim(U_i)$ $(i = 1, 2)$.

At first we treat the special linear case, i.e., $H = \mathrm{SL}_n(q)$: Replace $h_1, h_2 \in H$ by conjugates in Jordan normal form (where one off-diagonal entry might be different from 0 or 1) as matrices with respect to a suitable basis $e_1, \ldots, e_n$ with all 1×1 Jordan blocks corresponding to eigenvalue one in the upper left corner. Assuming there are m_i of these for h_i, i.e., $e_1.h_i = e_1, \ldots, e_{m_i}.h_i = e_{m_i}$, it is easy to see that $m_i \geq 2l_i - n$ $(i = 1, 2)$.

Indeed, if k_j ($j \geq 1$) is the number of Jordan blocks $J_j(1)$ of a linear map g on a vector space k^n and $l := \dim(\ker(1-g)) = n(1-\ell_{\mathrm{rk}}(g))$, we have $l = \sum_{j=1}^n k_j$ and $\sum_{j=1}^n jk_j \leq n$. Hence $k_1 = l - \sum_{j=2}^n k_j$ together with $\sum_{j=2}^n k_j \leq \frac{1}{2}\sum_{j=2}^n jk_j \leq \frac{n-k_1}{2}$ imply that $k_1 \geq 2l-n$. Applying this to $g := h_i$ yields the claim $m_i \geq 2l_i - n$ ($i = 1,2$).

Set $m := \min\{m_1, m_2\}$ and $W' := \langle e_{m+1}, \ldots, e_n\rangle$. For $X := \langle e_{2m-n}, \ldots, e_m\rangle$ the space $Y := X \oplus W'$ gives rise to a quasisimple group $K := \mathrm{SL}(Y)$ ($\dim(Y) \geq 3$, since $n - m \geq 1$ as $h_1 \neq \mathrm{id}_V$, so that K is quasisimple). Note that X is well-defined, since by assumption on l_1 and l_2 we have $n \geq m \geq 3n/4 > n/2$. Then $\dim(X) = \dim(W')+1 = n-m+1 \geq 2$. Then set $Y' := \langle e_1, \ldots, e_{2m-n-1}\rangle$. The operators h_1, h_2 write as $h_i = \mathrm{id}_{Y'} \oplus \mathrm{id}_X \oplus A_i = \mathrm{id}_{Y'} \oplus B_i$ ($i = 1,2$) with respect to the decompositions $V = Y' \oplus X \oplus W' = Y' \oplus Y$. Assuming $m_1 \leq m_2$ gives $1/6 \leq \ell_{\mathrm{rk}}(B_1) \leq 1/2$ (B_1 and B_2 are seen as elements of $K = \mathrm{SL}(Y)$; one gets close to the lower bound when A_1 only has Jordan blocks $J_2(1)$), so by Proposition 2.13 of [67] we have $\ell_{\mathrm{pr}}(B_1) \geq 1/6$, implying the existence of a constant $D \in \mathbb{N}$ with $B_2 \in K = (B_1^K)^{*D}$ which yields $h_2 \in (h_1^H)^{*D}$ (by Lemma 1.10). If $m_2 < m_1$, the same argument shows $1/6 \leq \ell_{\mathrm{rk}}(B_2)$ and $\ell_{\mathrm{rk}}(B_1) \leq 1/2$, so as previously $\ell_{\mathrm{pr}}(B_1) = \ell_{\mathrm{rk}}(B_1)$. By Lemma 1.10, there is $c > 0$ such that for all integers $k \geq c/\ell_{\mathrm{pr}}(B_1)$ we have $K = (B_1^K)^{*k}$. Then it holds that $h_2 \in (h_1^H)^{*k}$ and by

$$6c\ell_{\mathrm{rk}}(h_2)/\ell_{\mathrm{rk}}(h_1) = 6c\ell_{\mathrm{rk}}(B_2)/\ell_{\mathrm{rk}}(B_1) \geq c/\ell_{\mathrm{pr}}(B_1)$$

we are done in this case with $C := 6c$.

In the other cases, i.e., $H \neq \mathrm{SL}_n(q)$, the proof is almost identical: Define $U := U_1 \cap U_2$ and $l := \dim(U) \geq l_1 + l_2 - n$. From Lemma 1.6 we get $W \leq U$ non-singular with $\dim(W) \geq 2l-n$. Then we infer $\dim(W^\perp) \leq 2(n-l)$ from $\dim(W) \geq 2l-n$. This implies

$$\dim(W^\perp) \leq 2(n-l) \leq 2(2n-(l_1+l_2)) \leq 4(n-\min\{l_1,l_2\})$$

and

$$4\min\{l_1,l_2\} - 3n \leq 2(l_1+l_2) - 3n \leq 2l-n \leq \dim(W).$$

As by assumption $\ell_{\mathrm{rk}}(h_1), \ell_{\mathrm{rk}}(h_2) \leq 1/8$, implying that $7/8n \leq l_1, l_2$, we obtain $\dim(W^\perp) \leq \dim(W)$. Now take $X \leq W$ a non-singular subspace such that $\dim(X) \geq \dim(W^\perp)$ is as small as possible such that $Y := X \perp W^\perp$ gives rise to a classical quasisimple group K from the beginning of Section 1.2. As $\dim(W^\perp) \geq 1$ (since $h_1 \neq \mathrm{id}_V$), then $\dim(X) \leq d\dim(W^\perp)$ for some absolute $d > 1$. With respect to the decompositions $V = Y^\perp \perp X \perp W^\perp = Y^\perp \perp Y$, the operators h_1, h_2 write as $h_i = \mathrm{id}_{Y^\perp} \oplus \mathrm{id}_X \oplus A_i = \mathrm{id}_{Y^\perp} \oplus B_i$ with isometric automorphism A_i of $W^\perp$ and B_i of Y ($i = 1,2$). Now assume $l_1 \leq l_2$. Then

$$\ell_{\mathrm{rk}}(B_1) = \frac{\mathrm{rk}(\mathrm{id}_{W^\perp} - A_1)}{\dim(W^\perp)+\dim(X)},$$

which can be bounded from above by 1/2 and from below by the chain

$$\frac{n-l_1}{2(1+d)(n-l)} \geq \frac{n-l_1}{2(1+d)(2n-(l_1+l_2))} \geq \frac{n-l_1}{4(1+d)(n-l_1)} = \frac{1}{4(1+d)}.$$

Hence by Proposition 2.13 of [67] we get that $\ell_{\mathrm{pr}}(B_1) \geq 0.25(1+d)^{-1}$, so by Lemma 1.10 there is a constant $D \in \mathbb{N}$ such that $B_2 \in K = (B_1^K)^{*D}$, implying that $h_2 \in (h_1^H)^{*D}$. On the other hand, if $l_2 < l_1$, by the same computation as above $0.25(1+d)^{-1} \leq \ell_{\mathrm{rk}}(B_2)$ and $\ell_{\mathrm{rk}}(B_1) \leq 1/2$, so by Proposition 2.13 of [67] we have $\ell_{\mathrm{pr}}(B_1) = \ell_{\mathrm{rk}}(B_1)$. Applying Lemma 1.10 gives $c > 0$ such that for all integers $k \geq c/\ell_{\mathrm{pr}}(B_1)$ we have $K = (B_1^K)^{*k}$. But then $h_2 \in (h_1^H)^{*k}$ for such k, and, since

$$4(1+d)c\ell_{\mathrm{rk}}(h_2)/\ell_{\mathrm{rk}}(h_1) = 4(1+d)c\ell_{\mathrm{rk}}(B_2)/\ell_{\mathrm{rk}}(B_1) \geq c/\ell_{\mathrm{pr}}(B_1),$$

we are done in this case with $C \coloneqq 4(1+d)c$.

We still need to eliminate the condition $\ell_{\mathrm{rk}}(h_2) \leq 1/8$. This goes as follows: As shown previously, the conjugacy class h_1^H generates all elements $h \in H$ with $\ell_{\mathrm{rk}}(h) \leq 1/8$ 'quickly' and these elements generate the whole group H in constantly many steps (by Lemma 1.10). In total any $h_2 \in H$ with $\ell_{\mathrm{rk}}(h_2) > 1/8$ is generated quickly by h_1^H. □

Remark 1.14. The condition $\ell_{\mathrm{rk}}(h_1) \leq 1 - \varepsilon$ for a fixed $\varepsilon > 0$ cannot be removed by Example 1.2. In that sense, the previous result is best possible.

Remark 1.15. The only universal finite quasisimple groups from families of unbounded rank which are not covered by Lemmas 1.8 and 1.13 are the double covers of the orthogonal groups in odd characteristic.

Defining $\ell_{\mathrm{rk}}(h) \coloneqq \ell_{\mathrm{rk}}(\overline{h})$ for h an element of the twofold cover of $\Omega_{2m}^{\pm}(q)$ or $\Omega_{2m+1}(q)$ (q odd) the statement of Lemma 1.13 also holds for these. This is since $\widetilde{\mathrm{A}}_8$ is embedded into both of them (if m is large enough) as lifts of products of an even number of the reflections with respect to the vectors $e_1 - e_2, e_2 - e_3, \ldots, e_7 - e_8 \in V$, where $e_1, \ldots, e_8$ is an orthonormal system for f. So we can argue as at the end of the proof of Lemma 1.8 to see that the two-element kernel of the covering map is generated 'quickly' by conjugates of any non-central element.

1.4 The lattice of normal subgroups of an algebraic ultraproduct of classical finite quasisimple groups

Let $\mathcal{H} = (H_i)_{i \in I}$ be a sequence of groups from the list at the beginning of Section 1.2 and set $G \coloneqq \prod_{i \in I} H_i$. Let n_i be the dimension of the natural module V_i of H_i. For some ultrafilter $\mathcal{U}$ on I for which $\lim_{\mathcal{U}} n_i = \infty$ define the normal subgroup $N_0 \coloneqq \{(h_i)_{i \in I} \in G \mid h_i = 1_{H_i} \text{ along } \mathcal{U}\}$ of G as in Theorem 1.3. In this section, we give a complete description of the lattice of normal subgroups of the algebraic ultraproduct $\mathcal{H}_{\mathcal{U}} = G/N_0$ of the groups H_i ($i \in I$) with respect to $\mathcal{U}$.

Recall from Theorem 1.3 the definitions of the subgroups $N_{\mathrm{rk}} := \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_{\mathrm{rk}}(h_i) = 0\}$ and $N_{\mathrm{pr}} := \{(h_i)_{i\in I} \in G \mid \lim_{\mathcal{U}} \ell_{\mathrm{pr}}(\overline{h}_i) = 0\}$ of G. Here ℓ_{rk} and ℓ_{pr} are the norms from Definition 0.12 (again we write $\overline{h}_i \in \overline{H}_i = H_i/\mathbf{Z}(H_i)$ for the image of $h_i \in H_i$ under the natural homomorphism). As they are norms, it is clear that N_{rk} and N_{pr} are normal in G and contain N_0. Moreover, as $\ell_{\mathrm{pr}}(\overline{h}_i) \leq \ell_{\mathrm{rk}}(h_i)$ ($h_i \in H_i$; $i \in I$), we get that $N_{\mathrm{rk}} \leq N_{\mathrm{pr}}$.

The following result is an immediate consequence of Lemma 1.10:

Lemma 1.16. *The subgroup N_{pr} contains all proper normal subgroups of G containing N_0. In particular, G/N_{pr} is non-abelian simple and N_{pr}/N_0 is a characteristic subgroup of $\mathcal{H}_{\mathcal{U}} = G/N_0$.*

Proof. When $h = (h_i)_{i\in I} \in G \setminus N_{\mathrm{pr}}$, there exist $\varepsilon > 0$ and $U \in \mathcal{U}$ such that $\ell_{\mathrm{pr}}(\overline{h}_i) \geq \varepsilon$ for $i \in U$. Hence from Lemma 1.10 it follows that there is $k \in \mathbb{N}$ such that $(h_i^{H_i})^{*k} = H_i$ for $i \in U$. This implies $\langle\langle h\rangle\rangle_G N_0 = G$, as wished. $\square$

Hence from now on we may restrict to the normal subgroups of G above N_0 which are contained in N_{pr}. Let us first characterize N_{pr} among the subgroups of G containing N_{rk}. For this we recall the definition of the quasiscalars $S(H)$ from the end of Section 1.2.

Lemma 1.17. *The map $\varphi\colon N_{\mathrm{pr}} \to Z := \prod_{\mathcal{U}} S(H_i)$ defined by $(h_i)_{i\in I} \mapsto \overline{(\lambda_i)_{i\in I}}$, where λ_i is arbitrary (from $S(H_i)$) if $\ell_{\mathrm{pr}}(\overline{h}_i) \geq 1/4$ and λ_i is the unique $\lambda \in \mathbb{F}_{q_i}^{\times}$ for which $\ell_{\mathrm{rk}}(\lambda^{-1}h_i) < 1/4$ otherwise, is a surjective homomorphism with kernel N_{rk}. Moreover, $Z \cong N_{\mathrm{pr}}/N_{\mathrm{rk}} = \mathbf{Z}(G/N_{\mathrm{rk}})$.*

Proof. At first we check that λ_i is always in $S(H_i)$: This is clear when $\ell_{\mathrm{pr}}(\overline{h}_i) \geq 1/4$. So assume the opposite. In the special linear case there is nothing to check, so consider the remaining cases. Set U_i to be the eigenspace corresponding to eigenvalue λ_i of h_i and set $\dim(U_i) =: l_i > 3/4n_i$. Then by Lemma 1.6 there is $W_i \leq U_i$ non-singular such that $\dim(W_i) \geq 2l_i - n_i > n_i/2 \geq 1$. The restriction $h_i|_{W_i}$ is a scalar from the full isometry group $\mathrm{GI}(W_i, f_i|_{W_i})$, so it must lie in $S(H_i)$.

Let us now show that φ is a homomorphism. So let $(g_i)_{i\in I}, (h_i)_{i\in I} \in N_{\mathrm{pr}}$ and pick $U \in \mathcal{U}$ and sequences $(\lambda_i)_{i\in I}, (\mu_i)_{i\in I}$ from $\prod_{i\in I} S(H_i)$ such that

$$\ell_{\mathrm{rk}}(\lambda_i^{-1}g_i), \ell_{\mathrm{rk}}(\mu_i^{-1}h_i) < 1/8$$

for $i \in U$ (this is possible by the definition of N_{pr}). Then it follows that $\ell_{\mathrm{rk}}((\lambda_i\mu_i)^{-1}g_ih_i) < 1/4$ for $i \in U$ by the triangle inequality, establishing that φ is a homomorphism.

Also φ is surjective: Let $\lambda = \overline{(\lambda_i)_{i\in I}} \in Z$. In any case, by Lemma 1.7 there is a diagonal matrix in H_i with all but at most two entries of the diagonal equal to λ_i. This gives the desired preimage of λ.

Moreover, the kernel of φ consists of all sequences $h = (h_i)_{i\in I} \in N_{\mathrm{pr}}$ such that, when $\overline{(\lambda_i)_{i\in I}}$ is the image of h under φ, then $\lambda_i = 1$ for $i \in U$ and some $U \in \mathcal{U}$. But this means precisely that $h \in N_{\mathrm{rk}}$, since then $0 = \lim_{\mathcal{U}} \ell_{\mathrm{pr}}(\overline{h}_i) = \lim_{\mathcal{U}} \ell_{\mathrm{rk}}(\lambda_i^{-1}h_i) = \lim_{\mathcal{U}} \ell_{\mathrm{rk}}(h_i)$.

Lastly, we have to show that $N_{\mathrm{pr}}/N_{\mathrm{rk}}$ is the center of G/N_{rk}. Since by Lemma 1.16 the quotient G/N_{pr} is non-abelian simple, it suffices to show that $[G, N_{\mathrm{pr}}] \subseteq N_{\mathrm{rk}}$. So let $g = (g_i)_{i\in I} \in G$ and $h = (h_i)_{i\in I} \in N_{\mathrm{pr}}$ and let $(\lambda_i)_{i\in I}$ be a sequence in $\prod_{i\in I} \mathbb{F}_{q_i}^\times$ such that $(\lambda_i^{-1}h_i)_{i\in I} \in N_{\mathrm{rk}}$. Let U_i be the eigenspace of h_i corresponding to eigenvalue λ_i and set $W_i := U_i \cap U_i.g_i$. Then for $w \in W_i$ one computes that $w.[g_i, h_i] = w.g_i^{-1}h_i^{-1}g_ih_i = w$. As $\dim(W_i)/n_i$ tends to one along $\mathcal{U}$, this verifies that $[g, h] \in N_{\mathrm{rk}}$. The proof is complete. □

To describe the normal subgroups N of G with $N_0 \leq N \leq N_{\mathrm{pr}}$ accurately, we need some definitions: Define $\mathcal{L}$ to be the set of null sequences $(r_i)_{i\in I}$ along $\mathcal{U}$ with $r_i \in [0, 1]$ $(i \in I)$. Two such sequences $(r_i)_{i\in I}$ and $(s_i)_{i\in I}$ are said to be equivalent (write $(r_i)_{i\in I} \sim (s_i)_{i\in I}$ and $\overline{(r_i)}_{i\in I}$ for the corresponding equivalence class) if $\lim_{\mathcal{U}} r_i/s_i \in (0,\infty)$. Here we set $0/0 := 1$ and $x/0 := \infty$ for $x > 0$. For two elements r and s of the quotient $\mathcal{L}/\!\sim$ write $r \leq s$ if and only if $\lim_{\mathcal{U}} r_i/s_i < \infty$, where $(r_i)_{i\in I}$ and $(s_i)_{i\in I}$ are representatives for r and s, respectively. It is routine to check that this is a definition and turns $(\mathcal{L}/\!\sim, \leq)$ into a linear order.

Now define the function $\mathrm{ct}\colon N_{\mathrm{pr}} \to \mathcal{L}/\!\sim$ by $h = (h_i)_{i\in I} \mapsto \overline{(\ell_{\mathrm{pr}}(\overline{h}_i))}_{i\in I}$ and call $\mathrm{ct}(h)$ the *convergence type* of h. Denote by L the subset of elements r of $\mathcal{L}/\!\sim$ for which either $r \geq \overline{(1/n_i)}_{i\in I}$ or $r = \overline{(0)}_{i\in I}$.

Lemma 1.18. *The image of the function* $\mathrm{ct}\colon N_{\mathrm{pr}} \to \mathcal{L}/\!\sim$ *is equal to* L.

Proof. At first we prove that the image of ct lies in L: Namely when $r = \overline{(r_i)}_{i\in I} \in \mathrm{ct}(N_{\mathrm{pr}})$ and $r \neq \overline{(0)}_{i\in I}$, then $r_i \neq 0$ for all $i \in U$ for some $U \in \mathcal{U}$. But then it follows that $r_i \geq 1/n_i$ for these i and hence $r \geq \overline{(1/n_i)}_{i\in I}$.

The surjectivity is a bit more subtle. We prove it here only for the case that $H_i = \mathrm{SL}_{n_i}(q_i)$; the other cases are analogous. Choose non-central elements $g_i \in \mathrm{SL}_2(q_i)$ $(i \in I)$. Then, if $r = \overline{(r_i)}_{i\in I}$ satisfies $r_i \geq 1/n_i$ along $\mathcal{U}$, set $h_i := g_i^{\oplus\lfloor n_ir_i\rfloor} \oplus \mathrm{id}_{n_i-2\lfloor n_ir_i\rfloor}$ if $n_i \geq 2\lfloor n_ir_i\rfloor$ and choose an arbitrary element otherwise. Letting $h = (h_i)_{i\in I}$, one sees immediately that $\mathrm{ct}(h) = r$. Finally, $\mathrm{ct}(1_G) = \overline{(0)}_{i\in I}$ completes the proof. □

Remark 1.19. The normalized rank length function ℓ_{rk} does not attain any possible value on the classical quasisimple groups even for large ranks: E.g., $2m\ell_{\mathrm{rk}}(g)$ for $g \in \mathrm{GO}_{2m}^{\varepsilon}(q)$ (q even) is even if and only if $g \in \Omega_{2m}^{\varepsilon}(q)$ (see [78, page 77] at the end of Section 3.8.1).

Owing to Lemma 1.18, henceforth we shall consider ct as a function with domain N_{pr} and codomain L. We extend this function to subsets $S \subseteq N_{\mathrm{pr}}$ by setting $\mathrm{ct}(S) := \{\mathrm{ct}(s) \mid s \in S\}$. For a normal subgroup N of G between N_0 and N_{pr} the set $\mathrm{ct}(N)$ is then called the associated order ideal (see Lemma 1.22 below).

Lemma 1.20. *For such a subgroup it holds that* $\mathrm{ct}(N) = \mathrm{ct}(N \cap N_{\mathrm{rk}})$.

Proof. It is clear that $\mathrm{ct}(N) \supseteq \mathrm{ct}(N \cap N_{\mathrm{rk}})$. Let us prove the converse containment to finish the proof: Let $h = (h_i)_{i\in I} \in N \leq N_{\mathrm{pr}}$. Then there is $U \in \mathcal{U}$ and $\lambda_i \in \mathbb{F}_{q_i}^\times$ $(i \in I)$ such that $\ell_{\mathrm{pr}}(\overline{h}_i) = \ell_{\mathrm{rk}}(\lambda_i^{-1}h_i) < 1/8$ for $i \in U$ and so, when U_i is the eigenspace of h_i corresponding to the eigenvalue λ_i for such an i, then $l_i := \dim(U_i) > 7/8n_i$.

At first assume that $H_i = \mathrm{SL}_{n_i}(q_i)$ for all $i \in I$. Then by the same argument as in the proof of Lemma 1.13 we find a subspace $W_i \le U_i$ of dimension $\dim(W_i) \ge 2l_i - n_i > n_i/2$ which has an h_i-invariant complement $W_i^\perp$. Decompose $W_i = W_i^1 \oplus W_i^2$ such that there is a map $g_i \in H_i = \mathrm{SL}_{n_i}(q_i)$ restricting to the identity on W_i^2 and interchanging $W_i^\perp$ and W_i^1 $(i \in U)$.

In the opposite case, i.e., all H_i preserve a form, by Lemma 1.6 there is a non-singular subspace $W_i \le U_i$ with $\dim(W_i) \ge 2l_i - n_i > n_i/2$. Moreover, for all large n_i we may also assume that $n_i - \dim(W_i) = \dim(W_i^\perp) \ge 2$ (by modifying W_i a little). Now Lemma 1.5 implies the existence of $g_i \in H_i$ and $W_i^1, W_i^2 \le W_i$ $(i \in U)$ non-singular such that $V_i = W_i^\perp \perp W_i^1 \perp W_i^2$ and g_i restricts to the identity on W_i^2 while it interchanges $W_i^\perp$ and W_i^1.

In both cases it holds that $u.[g_i, h_i] = u.g_i^{-1}h_i^{-1}g_ih_i = u.g_i^{-1}(\lambda_i^{-1}g_i)h_i = \lambda_i^{-1}u.h_i$ for $u \in W_i^\perp$, $v.[g_i, h_i] = v.g_i^{-1}h_i^{-1}g_ih_i = \lambda_i v.h_i^{-g_i}$ for $v \in W_i^1$, and $w.[g_i, h_i] = w$ for $w \in W_i^2$ $(i \in U)$. Since $\dim(W_i^2) = n_i - \dim(W_i^1) - \dim(W_i^\perp) = n_i - 2\dim(W_i^\perp) = n_i - 2(n_i - \dim(W_i)) \ge 4l_i - 3n_i > n_i/2$, we have that $\ell_{\mathrm{rk}}([g_i, h_i]) = 1 - \dim(\ker(1 - [g_i, h_i]))/n_i < 1/2$ $(i \in U)$. Altogether, this implies that $\ell_{\mathrm{pr}}(\overline{[g_i, h_i]}) = \ell_{\mathrm{rk}}([g_i, h_i]) = \ell_{\mathrm{rk}}(\lambda_i^{-1}h_i) + \ell_{\mathrm{rk}}(\lambda_i h_i^{-1}) + 0 = 2\ell_{\mathrm{rk}}(\lambda_i^{-1}h_i) = 2\ell_{\mathrm{pr}}(\overline{h}_i)$. Here we use essentially Proposition 2.13 of [67] for the first equality.

Therefore, setting $g \coloneqq (g_i)_{i\in I} \in G$, by Lemma 1.17 the commutator $[g, h]$ is in N_{rk} (and of course in N as N is normal) and has the same convergence type as h. $\square$

Remark 1.21. In view of Lemma 1.18, this implies that even the restriction $\mathrm{ct}|_{N_{\mathrm{rk}}} : N_{\mathrm{rk}} \to L$ is surjective.

Now we are ready to justify the name 'associated order ideal' for $\mathrm{ct}(N)$ for N normal in G lying between N_0 and N_{pr}.

Lemma 1.22. *Let N be a normal subgroup of G between N_0 and N_{pr}. Then $\mathrm{ct}(N)$ is an order ideal of L. Moreover, the maps*

$$\{N \trianglelefteq G \mid N_0 \le N \le N_{\mathrm{rk}}\} \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \{\textit{order ideals of } (L, \le)\}$$

defined by $\alpha\colon N \mapsto \mathrm{ct}(N)$ and $\beta\colon J \mapsto \{g \in N_{\mathrm{rk}} \mid \mathrm{ct}(g) \in J\}$ are isomorphisms of posets and mutually inverse to each other.

Proof. By Lemma 1.20, we may restrict to the case $N \le N_{\mathrm{rk}}$. Clearly, $\mathrm{ct}(N)$ is not empty as $1_G \in N$. If $r \in \mathrm{ct}(N)$, Lemma 1.13 implies that, if $s \le r$, then $s \in \mathrm{ct}(N)$. As $(L, \le)$ was a linear order, $\mathrm{ct}(N)$ is an order ideal.

Concerning the second part: At first note that, as shown previously, $\alpha(N) = \mathrm{ct}(N)$ is an order ideal. Also $\beta(J)$, for an order ideal J of $(L, \le)$, is a normal subgroup, since $\mathrm{ct}(g), \mathrm{ct}(h) \in J$ implies that $\mathrm{ct}(gh) \le \max\{\mathrm{ct}(g), \mathrm{ct}(h)\} \in J$ and $\mathrm{ct}(g^{-1}) = \mathrm{ct}(g^h) = \mathrm{ct}(g) \in J$; also $\mathrm{ct}(1_G) = \overline{(0)}_{i\in I} \in J$. Moreover, both maps α and β are inclusion preserving. To show that $\beta \circ \alpha\colon N \mapsto \mathrm{ct}(N) \mapsto \{g \in N_{\mathrm{rk}} \mid \mathrm{ct}(g) \in \mathrm{ct}(N)\}$ is the identity, i.e., that $\mathrm{ct}(g) \in \mathrm{ct}(N)$ iff $g \in N$, is just another straightforward application of Lemma 1.13. The map $\alpha \circ \beta \colon J \mapsto \{g \in N_{\mathrm{rk}} \mid \mathrm{ct}(g) \in J\} \mapsto \{\mathrm{ct}(g) \in L \mid g \in N_{\mathrm{rk}}, \mathrm{ct}(g) \in J\}$ is the identity by Remark 1.21. $\square$

Remark 1.23. Similarly, one can prove that for $g \in N_{\mathrm{rk}}$ and $S \subseteq N_{\mathrm{rk}}$ it holds that $\mathrm{ct}(g)$ is in the order ideal of $(L, \leq)$ generated by $\mathrm{ct}(S)$ if and only if $g \in \langle\!\langle S \rangle\!\rangle_G N_0$.

Corollary 1.24. *The normal subgroups of G lying between N_0 and N_{rk} are linearly ordered.*

Proof. This holds by the correspondence of Lemma 1.22, since $(L, \leq)$ is linear and so its order ideals are linearly ordered by inclusion. $\square$

Corollary 1.25. *For $N \leq N_{\mathrm{rk}}$ normal in G containing N_0 it holds that N is perfect, i.e., $N = N'$.*

Proof. Pick $h \in N$ and proceed as in the proof of Lemma 1.20. The construction of g shows that $\mathrm{ct}(g) \leq \mathrm{ct}(h)$, so Remark 1.23 implies $g \in \langle\!\langle h \rangle\!\rangle_G N_0 \leq N$. Hence the commutator $[g, h]$ lies in N' and $\mathrm{ct}([g, h]) = \mathrm{ct}(h)$, so $\mathrm{ct}(N') = \mathrm{ct}(N)$. To apply Lemma 1.22 to deduce that $N = N'$, we still need that $N_0 \leq N'$. It is therefore enough to show that N_0 is perfect. This follows from the following simple application of Lemma 1.10: Let $h \in H$ be an element with $\ell_{\mathrm{pr}}(\overline{h}) \geq \varepsilon$ for some fixed $\varepsilon > 0$, where H is one of the H_i. Then the previously mentioned result implies that $H' \supseteq (h^H)^{*k} h^{-k} = H$ for a fixed integer k only depending on ε. Hence, if $h = (h_i)_{i \in I} \in N_0$, applying this in every component i for which $h_i \neq 1_{H_i}$ yields that $N_0 \leq N_0'$, as wished. $\square$

Remark 1.26. The argument at the end of the proof also shows that G itself is perfect.

Remark 1.27. In particular, this implies that for $N_0 \leq N \leq N_{\mathrm{pr}}$ normal in G it holds that $N \cap N_{\mathrm{rk}} = [G, N] = N'$ by the end of Lemma 1.17 and Corollary 1.25. Hence $N_{\mathrm{rk}} = (N_{\mathrm{pr}})'$ and so N_{rk}/N_0 is characteristic in G/N_0 as well. Also $N_{\mathrm{pr}}/N_{\mathrm{rk}} \cong \prod_{\mathcal{U}} S(H_i)$ from Lemma 1.17.

Using the correspondence of Lemma 1.22, we introduce the normal subgroup N_1 as the normal subgroup of N_{rk} such that $\mathrm{ct}(N_1) = J_1$, where $J_1 \coloneqq \{\overline{(r_i)}_{i \in I} \in L \mid n_i r_i < C \text{ for some } C > 0 \text{ along } \mathcal{U}\}$. It is obvious that J_1 covers the trivial order ideal $J_0 = \{\overline{(0)}_{i \in I}\}$ corresponding to N_0, i.e., there is no order ideal properly between them. Hence N_1 covers N_0. Let us mention here that for M, N normal subgroups of G between N_0 and N_{rk} it holds that M is covered by N if and only if the corresponding order ideals $\alpha(M)$ and $\alpha(N)$ of $(L, \leq)$ are of the form $\alpha(M) = \{r \in L \mid r < s\}$ and $\alpha(N) = \{r \in L \mid r \leq s\}$ for some $s \neq \overline{(0)}_{i \in I}$.

Because we will need them later, we introduce the normal subgroups $A_0 \coloneqq \{h \in N_{\mathrm{pr}} \mid \mathrm{ct}(h) \in J_0\} = \prod_{\mathcal{U}} \mathbf{Z}(H_i)$ and $A_1 \coloneqq \{h \in N_{\mathrm{pr}} \mid \mathrm{ct}(h) \in J_1\}$.

Now we complete the picture of the lattice of normal subgroups of G/N_0 by looking at an arbitrary normal subgroup $N \leq N_{\mathrm{pr}}$ of G with $N_0 \leq N$.

Let us at first assume that $N_1 \leq N$. Then the following lemma applies:

Lemma 1.28. *Define maps between the following two sets*

$$\begin{array}{c}\{N \trianglelefteq G \mid N_1 \leq N \leq N_{\mathrm{pr}}\}\\ {\scriptstyle\Psi}\big\uparrow\big\downarrow{\scriptstyle\Phi}\\ \{M \mid M \trianglelefteq G,\ N_1 \leq M \leq N_{\mathrm{rk}}\} \times \{A \mid A \trianglelefteq G,\ N_1 \leq A \leq A_1\}\end{array}$$

by $\Phi\colon N \mapsto (N \cap N_{\mathrm{rk}}, N \cap A_1)$ *and* $\Psi\colon (M, A) \mapsto MA$*. Then* Φ *and* Ψ *are isomorphisms of posets and mutually inverse to each other.*

Proof. Obviously, both maps are order preserving. Moreover, $\Phi \circ \Psi(M, A) = \Phi(MA) = (MA \cap N_{\mathrm{rk}}, MA \cap A_1) = (M(A \cap N_{\mathrm{rk}}), A(M \cap A_1))$ by Dedekind's modular law. However, it is easy to see from the definitions that $A \cap N_{\mathrm{rk}} = M \cap A_1 = N_1$, so $\Phi \circ \Psi = \mathrm{id}$ as $N_1 \leq M, A$.

Similarly, we compute $\Psi \circ \Phi(N) = \Psi(N \cap N_{\mathrm{rk}}, N \cap A_1) = (N \cap N_{\mathrm{rk}})(N \cap A_1) \leq N$. So it is enough to show that every element $h = (h_i)_{i \in I} \in N$ can be written as a product of elements from $N \cap N_{\mathrm{rk}}$ and $N \cap A_1$. We may assume that $\mathrm{ct}(h) > \overline{(0)_{i \in I}}$, i.e., $h \notin A_0$, since otherwise we can write $h = fg$ with $f = 1_G \in N \cap N_{\mathrm{rk}}$ and $g = h \in N \cap A_0 \subseteq N \cap A_1$. In the opposite case, choose λ_i such that $\ell_{\mathrm{rk}}(\lambda_i^{-1} h_i) = \ell_{\mathrm{pr}}(\overline{h}_i)$ $(i \in I)$. Let $g_i \in H_i$ be an element which has all but at most two diagonal entries equal to λ_i (which exists by Lemma 1.7). Then, setting $f := (h_i g_i^{-1})_{i \in I}$ and $g := (g_i)_{i \in I}$, $h = fg$ is the desired product decomposition. Indeed, $f \in N_{\mathrm{rk}}$ and $g \in A_1$ by construction. To show that $f, g \in N$ goes as follows: As in the proof of Lemma 1.20, find $c \in G$ such that $\mathrm{ct}(h) = \mathrm{ct}([c, h])$ and $[c, h] \in N_{\mathrm{rk}}$. Then we have that $f, gh^{-1} \in N_{\mathrm{rk}}$ and $\mathrm{ct}(f), \mathrm{ct}(gh^{-1}) \leq \mathrm{ct}(h) = \mathrm{ct}([c, h])$, so that $f, gh^{-1} \in \langle\!\langle [c, h] \rangle\!\rangle_G N_0 \leq \langle\!\langle h \rangle\!\rangle_G N_0 \leq N$ by Remark 1.23, yielding the claim that $g, f \in N$. The proof is complete. $\square$

Remark 1.29. Essentially, Lemma 1.28 (in combination with Lemma 1.22) says that the lattice of normal subgroups of G/N_1 is isomorphic to the product of the linear order of order ideals of $(L, \leq)$ different from J_0 and the subgroup lattice of the abelian group $Z \cong N_{\mathrm{pr}}/N_{\mathrm{rk}} = N_{\mathrm{rk}} A_1 / N_{\mathrm{rk}} \cong A_1/(A_1 \cap N_{\mathrm{rk}}) = A_1/N_1$ from Lemma 1.17. If $\Phi\colon N \mapsto (M, A)$, then $N \mapsto (\mathrm{ct}(M), \varphi(A))$ is the described isomorphism of lattices, where φ is the map from Lemma 1.17.

In this situation $M = N'$ is the commutator subgroup (by Remark 1.27 above) and $\varphi(A) = \varphi(MA) = \varphi(N) \cong N N_{\mathrm{rk}}/N_{\mathrm{rk}} \cong N/(N \cap N_{\mathrm{rk}}) = N/N'$ is the universal abelian quotient.

Now assume that $N_1 \not\leq N$. Then it follows from Lemma 1.20 that $\mathrm{ct}(N) = \mathrm{ct}(N \cap N_{\mathrm{rk}}) = \mathrm{ct}(N_0) = \{\overline{(0)_{i \in I}}\} = J_0$ (as by assumption $N_0 \leq N$). Hence we have established the following result.

Lemma 1.30. *If* N *is normal in* G *containing* N_0 *and* $N_1 \not\leq N$*, then* $N \leq A_0$.

Finally, we describe when a normal subgroup as in Lemma 1.30 is contained in a normal subgroup as in Lemma 1.28:

Lemma 1.31. *A normal subgroup N of G containing N_0 but not N_1 is contained in a normal subgroup $K \leq N_{\mathrm{pr}}$ containing N_1 if and only if $N \leq A$ (or equivalently $\varphi(N) \leq \varphi(A)$, with φ the map from Lemma 1.17), where $\Phi(K) = (M, A)$ is from Lemma 1.28.*

Proof. Assume $N \leq K$. Then by Lemma 1.30 it holds that $N \leq K \cap A_0 \leq K \cap A_1 = A$ implying $\varphi(N) \leq \varphi(A)$. Conversely, by Lemma 1.17 the assumption $\varphi(N) \leq \varphi(A)$ implies $NN_{\mathrm{rk}} \leq AN_{\mathrm{rk}}$. Thus $NN_{\mathrm{rk}} \cap A_1 = N(N_{\mathrm{rk}} \cap A_1) = NN_1 \leq AN_{\mathrm{rk}} \cap A_1 = A(N_{\mathrm{rk}} \cap A_1) = AN_1 = A$ by Dedekind's modular law. So $N \leq A \leq K$, completing the proof. □

We end up with a few remarks.

Remark 1.32. An ultraproduct of universal quasisimple groups of bounded rank (along the ultrafilter) will just result in a quasisimple group $X(k)$ over a pseudofinite field k, where X is the Lie type selected by $\mathcal{U}$. In this case, its lattice of normal subgroups is 'understood'. We have $N_0 = N_{\mathrm{rk}}$ and $N_{\mathrm{pr}} = \mathbf{Z}(X(k))$.

Remark 1.33. Again we have not yet covered the case that the H_i ($i \in I$) are double covers of $\Omega^{\pm}_{2m}(q)$ or $\Omega_{2m+1}(q)$ ($m \in \mathbb{Z}_+$ suitable, q odd). In this case, define N_0 as above and let A_0 be the elements of $G = \prod_{i \in I} H_i$ which are central along the ultrafilter $\mathcal{U}$. Then G/N_0 is a twofold cover of an ultraproduct G/M_0 of orthogonal groups with kernel $M_0/N_0 \cong \mathrm{C}_2$.

It follows from Remark 1.15 that, if $N_0 \leq N \trianglelefteq G$ contains an element non-central modulo N_0, i.e., it is not contained in A_0, then $M_0 \leq N$, so it corresponds to a normal subgroup of G/M_0 (which we understood).

Finally, when $N_0 \leq N \leq A_0$ with $M_0 \not\leq N$ and $M_0 \leq K \leq G$, then $N \leq K$ iff $NM_0 \leq K$ and this again can be decided by considering the lattice of normal subgroups of G/M_0, which we 'know'.

Using the same argument, one can show that the normal subgroups of an ultraproduct of the double covers of simple alternating groups are still linearly ordered by inclusion (since here $A_0 = M_0$).

Chapter 2

Metric approximation of groups by finite groups

2.1 Introduction

Eversince the work of Gromov on Gottschalk's surjunctivity conjecture [29], the class of sofic groups has attracted much interest in various areas of mathematics. Major applications of this notion arose in the work of Elek and Szabó on Kaplansky's direct finiteness conjecture [13], Lück's determinant conjecture [14], and more recently in joint work of Thom and Klyachko on generalizations of the Kervaire–Laudenbach conjecture and Howie's conjecture [41].

Despite considerable effort, no non-sofic group has been found so far. In view of this situation, attempts have been made to provide variations of the problem that might be more approachable. In the terminology of Holt and Rees, sofic groups are precisely those groups which can be approximated by finite symmetric groups equipped with the normalized Hamming length function (in the sense of Definition 1.6 of [71]). It is natural to vary the class of finite groups and also the metrics that are allowed. Note that our terminology (see Definition 2.1) differs from the one used in [57], where similar concepts were studied.

The strongest form of approximation is satisfied by LEF (resp. LEA) groups. In this case, it is well-known that a finitely presented group is not approximable by finite (resp. amenable) groups with discrete length function, i.e., it is not LEF (resp. LEA), if and only if it fails to be residually finite (resp. residually amenable). Examples of sofic groups which fail to be LEA (and thus also fail to be LEF) are given in [6] and [40] (see also [70]), answering a question of Gromov [29].

In [71] Thom proved that the so-called Higman group cannot be approximated by finite groups with commutator-contractive norm. In [35] Howie presented a group which by a result of Glebsky [**?**] turned out not to be approximable by finite nilpotent groups with arbitrary norm.

This chapter provides four more results of this type (see Sections 2.3 and 2.5). However, in our setting we restrict only to classes of finite groups and do not impose restrictions on the length functions of the approximating groups other than being invariant, i.e., being norms (see Definitions 2.1 and 2.8).

Recently, Glebsky [24] asked whether all groups can be approximated by finite solvable groups (in the sense of Definition 2.1). In Section 2.3, we answer this question by establishing that each non-trivial finitely generated perfect group is a counterexample (see Theorem 2.17). The key to this result is a theorem of Segal [63] on generators and commutators in finite solvable groups.

In Section 2.4, using results of Nikolov from [**?**] and of Liebeck and Shalev from [49], we prove that any non-trivial group which is approximable by finite groups has a non-trivial homomorphism into a metric ultraproduct of finite simple groups of type $\mathrm{PSL}_n(q)$ equipped with the normalized conjugacy length function (see Theorem 2.25).

In Section 2.5, we discuss the approximability of Lie groups by finite groups. It is easy to see that $\mathbb{R}$ as a topological group is not approximable by symmetric groups, i.e., it is not continuously embeddable into a metric ultraproduct of symmetric groups with arbitrary norms (see Remark 2.36). Using a much deeper analysis, we show that a connected Lie group is approximable by finite groups (in the sense of Definition 2.8) precisely when it is abelian (see Theorem 2.33). In [10, Question 2.11] Doucha asked for groups which can be equipped with a norm such that they do not embed into a metric ultraproduct of normed finite groups. Our result implies that any compact, connected, and non-abelian Lie group is an example of such a group. Thus the simplest example of a topological group which is not weakly sofic, i.e., not continuously embeddable into a metric ultraproduct of normed finite groups, is $\mathrm{SO}_3(\mathbb{R})$. However, we remark that every linear Lie group is an abstract subgroup of the algebraic ultraproduct of finite groups indexed over $\mathbb{N}$ (see Remark 2.35).

Furthermore, in the same section, we answer the question of Zilber [79] if there exists a compact simple Lie group which is not a quotient of an algebraic ultraproduct of finite groups. Indeed, we show that a Lie group which can be equipped with a norm generating its topology and which is an abstract quotient of a product of finite groups has abelian identity component (see Theorem 2.37). Hence any compact simple Lie group fails to be approximable by finite groups in the sense of Zilber.

A slight variation of Theorem 2.37 also answers Question 1.2 of Pillay [59]. Moreover, we point out that Theorems 2.33 and 2.37 provide an alternative proof of the main result of Turing [74].

Finally, using the same approach as for the previous two results, we solve the conjecture of Pillay [59] that the identity component of the Bohr compactification of any pseudofinite group is abelian (see Theorem 2.38).

All results of Section 2.5 follow from a theorem on generators and commutators in finite groups of Nikolov and Segal [**?**].

2.2 Preliminaries

In this section, we introduce the notion of *$\mathcal{C}$-approximable abstract* and *topological groups*, and present examples. Recall the definition of the invariant length functions ℓ_{d}, ℓ_{c}, ℓ_{rk}, ℓ_{pr}, and $\ell_{\mathrm{Cay},S}$ (see Definition 0.12). Also recall that $\mathbf{F}$ denotes a free group of rank $\mathrm{rk}(\mathbf{F})$ (see the beginning of Section 0.1).

2.2.1 On $\mathcal{C}$-approximable abstract groups

We define metric approximation of an abstract group by a class of finite groups. Throughout this chapter, let $\mathcal{C}$ be such a class. Subsequently, a group is called a *$\mathcal{C}$-group* if it belongs to the class $\mathcal{C}$.

Definition 2.1. *An abstract group G is called* $\mathcal{C}$-approximable *if there is a function $\delta_\bullet\colon G\setminus\{1_G\} \to (0,\infty]$ such that for any finite subset $S \subseteq G$ and $\varepsilon > 0$ there exist a group $H \in \mathcal{C}$, a norm ℓ_H on H, and a map $\varphi\colon S \to H$ such that*

(i) if $1_G \in S$, then $\varphi(1_G) = 1_H$;

(ii) if $g, h, gh \in S$, then $d_H(\varphi(g)\varphi(h), \varphi(gh)) < \varepsilon$;

(iii) for $g \in S \setminus \{1_G\}$ we have $\ell_H(\varphi(g)) \geq \delta_g$.

A map φ with the above properties is called an $(S, \varepsilon, \delta_\bullet)$-homomorphism.

Remark 2.2. Note that the above definition differs slightly from Definition 1.6 in [71], as we impose no restrictions on the norms. However, it is equivalent to Definition 6 from [24]. Indeed, we may even require that $\ell_H \leq 1$ and $\delta_\bullet \equiv 1$ in the above definition, without changing its essence. Namely, choosing $\varepsilon > 0$ small enough, setting $c \coloneqq \min_{g\in S}\delta_g$, $\ell'_H \coloneqq \min\{\ell_H/c, 1\}$, and defining $\delta'\colon G\setminus\{1_G\} \to (0,\infty]$ by $\delta'_\bullet \coloneqq 1$, we can replace δ by δ' and ℓ_H by ℓ'_H. So, in the sense of [34, page 3], if we do not impose restrictions on the norms on the groups from $\mathcal{C}$, the terms '*$\mathcal{C}$-approximation property*', '*discrete $\mathcal{C}$-approximation property*', and '*strong $\mathcal{C}$-approximation property*' coincide. Hence, for simplicity, subsequently, we assume the function $\delta_\bullet$ to be constant and write just δ for δ_g ($g \in G$).

Moreover, similar to soficity, being $\mathcal{C}$-approximable is a local property. This is expressed in the following remark.

Remark 2.3. An abstract group is $\mathcal{C}$-approximable if and only if every finitely generated subgroup has this property.

Let us now present some examples of $\mathcal{C}$-approximable abstract groups. Subsequently, denote by **Alt** (resp. **Fin**) the class of finite alternating groups (resp. the class of all finite groups). Indeed, $\mathcal{C}$-approximable abstract groups (in the above sense) can be seen as a generalization of sofic (resp. weakly sofic) groups as it is shown in Section 2 of [24]:

Example 2.4. A group is sofic (resp. weakly sofic) if and only if it is **Alt**-approximable (resp. **Fin**-approximable) as an abstract group.

Groups approximable by certain classes of finite simple groups of Lie type have been studied in [1] and [73, **?**].

Every $\mathcal{C}$-group G is certainly $\mathcal{C}$-approximable, since we can take $H := G$, φ to be the restriction of the identity on G to S, and $\ell_H := \ell_{\mathrm{d},H}$ to be the discrete length function on H in Definition 2.1. Hence Remark 2.3 implies:

Example 2.5. Every locally $\mathcal{C}$-group is $\mathcal{C}$-approximable as an abstract group.

Henceforth, let $\mathbf{Ab}_d$ be the class of finite abelian groups which are a direct sum of at most d cyclic groups. The last example we mention here is the following:

Example 2.6. A finitely generated abelian group which is a direct sum of at most d cyclic groups is $\mathbf{Ab}_d$-approximable as an abstract group.

This example will follow from the fact that such a group embeds in a connected abelian Lie group, which is $\mathbf{Ab}_d$-approximable as a topological group by Lemma 2.16 below.

There is another common equivalent characterization of $\mathcal{C}$-approximable groups via metric ultraproducts of normed $\mathcal{C}$-groups with norm. Recall Definition 0.7. Let us call the class $\mathcal{C}$ trivial if either $\mathcal{C} = \emptyset$ or $\mathcal{C} = \{\mathbf{1}\}$. Here is the promised characterization.

Lemma 2.7. *If $\mathcal{C}$ is a non-trivial class, every abstract $\mathcal{C}$-approximable group G is isomorphic to a discrete subgroup of a metric ultraproduct $(H_{\mathcal{U}}, \ell_{\mathcal{U}}) = \prod_{\mathcal{U}} (H_i, \ell_i)$ of $\mathcal{C}$-groups H_i with norms ℓ_i $(i \in I)$ such that $\mathrm{diam}(H_i, \ell_i) = 1$ and the distance between the images of any two different elements of G is one. If G is countable, I can be chosen to be $\mathbb{N}$ with the natural order and $\mathcal{U}$ to be some non-principal ultrafilter on it.*

Conversely, any subgroup of a metric ultraproduct of normed $\mathcal{C}$-groups is $\mathcal{C}$-approximable as an abstract group.

The proof of this result is identical to the corresponding proof in the sofic case, which is well-known. Hence we omit it here.

2.2.2 On $\mathcal{C}$-approximable topological groups

In view of Lemma 2.7 it is natural to generalize the notion of a $\mathcal{C}$-approximable (abstract) group to topological groups using ultraproducts:

Definition 2.8. *A topological group is called* $\mathcal{C}$-approximable *if it embeds continuously into a metric ultraproduct of normed $\mathcal{C}$-groups.*

Lemma 2.7 indicates the following class of examples of $\mathcal{C}$-approximable topological groups:

Example 2.9. Every $\mathcal{C}$-approximable abstract group equipped with the discrete topology is $\mathcal{C}$-approximable as a topological group.

Conversely, a $\mathcal{C}$-approximable topological group is $\mathcal{C}$-approximable as an abstract group when we 'forget' its topology.

To present more classes of examples, we need an auxiliary result. The following lemma gives a sufficient condition for a normed group to be isomorphic to an ultraproduct of normed finite groups. Its proof is trivial.

Lemma 2.10. *Let (G, ℓ_G) be a normed group, I an index set, $\mathcal{U}$ an ultrafilter on I, $(K_i, \ell_i)_{i \in I}$ a sequence of normed finite groups, and $(K_\mathcal{U}, \ell_\mathcal{U}) \coloneqq \prod_\mathcal{U} (K_i, \ell_i)$ its metric ultraproduct.*

(i) Assume there are maps $\varphi_i \colon G \to K_i$, which are isometric and a homomorphism in the $\mathcal{U}$-limit, i.e.,

$$\lim_{\mathcal{U}} d_i(\varphi_i(g), \varphi_i(h)) = d_G(g, h) \text{ and } \lim_{\mathcal{U}} d_i(\varphi_i(g)\varphi_i(h), \varphi_i(gh)) = 0$$

for all $g, h \in G$. Then there is an isometric embedding $\varphi \colon (G, \ell_G) \hookrightarrow (K_\mathcal{U}, \ell_\mathcal{U})$ in the ultraproduct defined by $\varphi(g) \coloneqq \overline{(\varphi_i(g))_{i\in I}}$.

(ii) The embedding φ is surjective if and only if for every $(k_i)_{i \in I} \in K \coloneqq \prod_{i\in I} K_i$ there exists $g \in G$ such that $\lim_\mathcal{U} d_i(\varphi_i(g), k_i) = 0$.

(iii) It surjects onto the subgroup of elements of finite length of $(K_\mathcal{U}, \ell_\mathcal{U})$ if the previous assertion holds for all $(k_i)_{i\in I} \in K$ with $\sup_{i\in I} \ell_i(k_i) < \infty$.

Let $\mathcal{C}^{\mathrm{P}}$ and $\mathcal{C}^{\mathrm{SP}}$ be the class of finite products of $\mathcal{C}$-groups and the class subgroups of finite products of $\mathcal{C}$-groups, respectively. Now we investigate which profinite groups are $\mathcal{C}$-approximable as topological groups. The standard example is given by the following lemma:

Lemma 2.11. *Let H_i $(i \in \mathbb{Z}_+)$ be $\mathcal{C}$-groups. Then the profinite group $P \coloneqq \prod_{i\in\mathbb{Z}_+} H_i$ is isomorphic to a metric ultraproduct of $\mathcal{C}^{\mathrm{P}}$-groups and so $\mathcal{C}^{\mathrm{P}}$-approximable as a topological group.*

Proof. We want to apply (i) and (ii) of Lemma 2.10 to $G \coloneqq P$. Equip G with the norm $\ell_G(h) \coloneqq \max\{1/i \,|\, i \in \mathbb{Z}_+,\, h_i \neq 1_{H_i}\} \cup \{0\}$, where $h = (h_i)_{i\in I} \in G$. Let $I \coloneqq \mathbb{Z}_+$ and $\mathcal{U}$ be some non-principal ultrafilter on I. Set $K_i \coloneqq H_1 \times \cdots \times H_i \leq G$ and let ℓ_i be the restriction of ℓ_G to K_i. Define $\varphi_i \colon G \to K_i$ in such a way that for every $g \in G$ the distance $d_G(\varphi_i(g), g)$ is minimal. By definition of ℓ_G, we have that $d_G(\varphi_i(g), g) < 1/i$. Hence it is easy to verify that condition (i) of Lemma 2.10 is fulfilled. For (ii) define g as the $\mathcal{U}$-limit in G of the sequence $(k_i)_{i\in I} \in K = \prod_{i\in I} K_i \leq G^I$ (which exists by compactness of G). Then $\lim_\mathcal{U} d_i(\varphi_i(g), k_i) \leq \lim_\mathcal{U} d_G(\varphi_i(g), g) + \lim_\mathcal{U} d_G(g, k_i) = 0$. This ends the proof. □

From the previous example we derive the following result.

Lemma 2.12. *For a pro-$\mathcal{C}^{\mathrm{SP}}$ group P the following are equivalent:*

(i) P is $\mathcal{C}^{\mathrm{P}}$-approximable as a topological group.

(ii) P is metrizable.

(iii) P is first-countable.

(iv) P is the inverse limit of a countable inverse system of $\mathcal{C}^{\mathrm{SP}}$-groups with all maps surjective.

(v) P is a closed topological subgroup of a countable product of $\mathcal{C}$-groups.

Proof. The implications (i)$\Rightarrow$(ii)$\Rightarrow$(iii) are trivial. (iii)$\Rightarrow$(iv): Let $\mathcal{B}$ be a countable system of open neighborhoods at 1_G. For each $B \in \mathcal{B}$ we can find an open normal subgroup $N \subseteq B$ such that P/N is a subgroup of a $\mathcal{C}^{\mathrm{SP}}$-group, so itself a $\mathcal{C}^{\mathrm{SP}}$-group (by Proposition 0.3.3(a) and Proposition 1.2.1 of [77]). Let $\mathcal{N}$ be the collection of these subgroups. Since $\bigcap \mathcal{B} = \{1_P\}$, as P is Hausdorff, the same holds for $\mathcal{N}$. Moreover, for $M, N \in \mathcal{N}$ we have $P/(M \cap N) \leq P/M \times P/N$, so $P/(M \cap N)$ is a $\mathcal{C}^{\mathrm{SP}}$-group, too. Hence we may assume that $\mathcal{N}$ is closed for finite intersections and apply Proposition 1.2.2 of [77] to obtain that P is the inverse limit of the $\mathcal{C}^{\mathrm{SP}}$-groups P/N $(N \in \mathcal{N})$ with respect to the natural maps $P/M \to P/N$ for $M \leq N$ $(M, N \in \mathcal{N})$.

(iv)$\Rightarrow$(v): By the standard construction of the inverse limit, it embeds into a countable product of $\mathcal{C}^{\mathrm{SP}}$-groups, which (by definition) embeds into a countable product of $\mathcal{C}$-groups. For (v)$\Rightarrow$(i) we only need to show that a countable product of $\mathcal{C}$-groups is $\mathcal{C}^{\mathrm{P}}$-approximable. This is Lemma 2.11. The proof is complete. $\square$

Remark 2.13. Lemma 2.12 implies that, if a pro-$\mathcal{C}^{\mathrm{SP}}$ group embeds continuously into a metric ultraproduct of normed $\mathcal{C}^{\mathrm{P}}$-groups, then it already embeds into such an ultraproduct of countably many groups.

We are now able to present the following important example:

Example 2.14. If $P = \overline{\langle x_1, \ldots, x_r \rangle}$ is a topologically finitely generated pro-$\mathcal{C}^{\mathrm{SP}}$ group, then P is $\mathcal{C}^{\mathrm{P}}$-approximable as a topological group.

Proof. Indeed, P embeds continuously into the product of all its continuous finite quotients $\prod_N P/N$ and finite generation implies that there are only countably many of these. By Proposition 1.2.1 of [77], we can restrict this map to a product of subgroups of $\mathcal{C}^{\mathrm{SP}}$-groups (which are itself $\mathcal{C}^{\mathrm{SP}}$-groups) such that it still is an embedding. But the latter embeds into a countable product of $\mathcal{C}$-groups. Hence P is $\mathcal{C}^{\mathrm{P}}$-approximable (by Lemma 2.12). $\square$

However, it is also simple to find examples of profinite groups that are not approximable by finite groups:

Example 2.15. Uncountable products of (non-trivial) finite groups are not metrizable and hence not approximable by finite groups.

Now we turn to Lie groups. The following example demonstrates that connected abelian Lie groups can always be approximated by finite abelian groups in the sense of Definition 2.8. Recall that $\mathbf{Ab}_d$ denotes the class of finite abelian groups which are a direct sum of at most d cyclic groups.

Lemma 2.16. *Every d-dimensional connected abelian Lie group $L = \mathbb{R}^m \times (\mathbb{R}/\mathbb{Z})^n$ $(m + n = d)$ equipped with the 'euclidean' length function ℓ_L is isometrically isomorphic to the subgroup of elements of finite length of a metric ultraproduct of normed $\mathbf{Ab}_d$-groups and hence $\mathbf{Ab}_d$-approximable.*

Proof. We wish to apply (i) and (iii) of Lemma 2.10 to $G := L$ with euclidean length function $\ell_G := \ell_L$. Let $I := \mathbb{Z}_+$ and $\mathcal{U}$ be some non-principal ultrafilter on I. For $i \in \mathbb{Z}_+$ set

$$S_i := \left\{ \frac{-i^2}{i}, \frac{-i^2+1}{i}, \dots, \frac{i^2}{i} \right\}^m \times \left(\frac{1}{i}\mathbb{Z}/\mathbb{Z} \right)^n \subseteq L,$$

and $K_i := (\mathbb{Z}/(4i^2))^m \times (\mathbb{Z}/(i))^n$. Define $\alpha_i \colon \{-i^2/i, (-i^2+1)/i, \dots, i^2/i\}^m \to (\mathbb{Z}/(4i^2))^m$ by $x \mapsto \overline{ix}$, let $\beta_i \colon (\frac{1}{i}\mathbb{Z}/\mathbb{Z})^n \to (\mathbb{Z}/(i))^n$ be the canonical isomorphism, and set $\gamma_i \colon S_i \to K_i$ to be the map $(x, y) \mapsto (\alpha_i(x), \beta_i(y))$. Moreover, equip K_i with the unique length function that turns γ_i into an isometry. Let $\delta_i \colon G \to S_i$ be a map such that $d_G(\delta_i(g), g)$ is minimal for all $g \in G$. Now define $\varphi_i := \gamma_i \circ \delta_i$. Clearly, condition (i) of Lemma 2.10 is now fulfilled. Condition (iii) follows from compactness of closed balls of finite radius in G by the same argument as at the end of the proof of Lemma 2.11. The proof is complete. □

We will see in Theorem 2.33 of Section 2.5 that connected abelian Lie groups are the only **Fin**-approximable connected Lie groups.

2.3 On Sol-approximable groups

Subsequently, let **Sol** (resp. **Nil**) be the class of finite solvable (resp. nilpotent) groups. The goal of this section is to establish the following theorem.

Theorem 2.17. *Any non-trivial finitely generated and perfect group is not* **Sol***-approximable.*

As a consequence, a finite group is **Sol**-approximable if and only if it is solvable: Indeed, any finite solvable group is **Sol**-approximable. On the other hand, a non-solvable finite group contains a non-trivial perfect subgroup and hence cannot be **Sol**-approximable by Remark 2.3 and Theorem 2.17.

Initially, Howie proved in [35] that the group $\langle x, y \,|\, x^{-2}y^{-3}, x^{-2}(xy)^5 \rangle$ is not **Nil**-approximable. We mimic his proof for any non-trivial finitely generated perfect group and then extend it by establishing that these groups are not even **Sol**-approximable using techniques of Segal [63, 65].

In preparation of the proof of Theorem 2.17 we need an auxiliary result, which is a generalization of [**?**, Theorem 4.3] – Theorem 2.18 below. Recall that the *pro-$\mathcal{C}$ topology* on a group K is the initial topology induced by all homomorphisms to $\mathcal{C}$-groups equipped with the discrete topology. Hence the closure $\overline{S}$ of a subset $S \subseteq K$ in this topology is characterized as follows: An element $k \in K$ lies in this closure if and only if $\varphi(k) \in \varphi(S)$ for all homomorphism $\varphi\colon K \to H$, where H is a $\mathcal{C}$-group.

Now we can state the auxiliary result relating $\mathcal{C}$-approximable groups to the pro-$\mathcal{C}$ topology on a free group of finite rank.

Theorem 2.18. *Let $\mathbf{F}/N$ be a presentation of a group G where $\operatorname{rk}(\mathbf{F}) < \infty$. Then, if G is $\mathcal{C}$-approximable, for each finite sequence $n_1, \ldots, n_k \in N$ it holds that $\overline{n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}} \subseteq N$ (in the pro-$\mathcal{C}$ topology on $\mathbf{F}$). The converse holds if $\mathcal{C}$ is closed with respect to finite products and subgroups.*

The proof of Theorem 2.18 will follow from the two subsequent lemmas (Lemmas 2.20 and 2.21). In the following, set $B_\varrho(\mathbf{F}) \coloneqq \{w \in \mathbf{F} \mid \ell_{\mathbf{F}}(w) \leq \varrho\}$ (the *ϱ-ball* around $1_{\mathbf{F}}$ with respect to the Cayley length function $\ell_{\mathbf{F}} \coloneqq \ell_{\mathrm{Cay},S}$ on the free group, where S are the standard generators of $\mathbf{F}$). In preparation of Lemma 2.20 below, we need to introduce the notion of a *$\mathcal{C}$-separable normal subgroup* of the free group.

Definition 2.19 ($\mathcal{C}$-separable normal subgroups of free groups)**.** *Let $N \trianglelefteq \mathbf{F}$, $\varrho \in \mathbb{N}$, $\varepsilon > 0$, $\delta > 0$. Assume there exists a group $H \in \mathcal{C}$, a norm ℓ_H on H, and a homomorphism $\varphi\colon \mathbf{F} \to H$ such that for any $w \in B_\varrho(\mathbf{F})$ we have*

$$\ell_H(\varphi(w)) < \varepsilon \text{ for } w \in N$$

and

$$\ell_H(\varphi(w)) > \delta \text{ for } w \in \mathbf{F} \setminus N.$$

*Then we call N $(\varrho, \varepsilon, \delta)$-*separated *by φ.*

*The normal subgroup N is called $\mathcal{C}$-*separable *if there exists $\delta > 0$ such that for any $\varrho \in \mathbb{N}$ and $\varepsilon > 0$ it is $(\varrho, \varepsilon, \delta)$-separated by some suitable homomorphism to a $\mathcal{C}$-group (for some chosen norm ℓ_H).*

The subsequent lemma is a generalization of [**?**, Lemma 6.2]. The proof given here follows the second proof given in the article [**?**].

Lemma 2.20 (Characterization of $\mathcal{C}$-approximable groups by $\mathcal{C}$-separability)**.** *If $N \trianglelefteq \mathbf{F}$ is $\mathcal{C}$-separable, then the group $G = \mathbf{F}/N$ is $\mathcal{C}$-approximable. When* $\operatorname{rk}(\mathbf{F}) < \infty$*, the reverse implication implication '$\Leftarrow$' also holds.*

Proof. We first prove the direction '$\Rightarrow$', i.e., we assume $N \trianglelefteq \mathbf{F}$ is $\mathcal{C}$-separable and derive that $G = \mathbf{F}/N$ is $\mathcal{C}$-approximable.

Let $\delta > 0$ be such that for any $\varrho \in \mathbb{N}$, $\varepsilon > 0$ the normal subgroup N is $(\varrho, \varepsilon, \delta)$-separated by a homomorphism to a $\mathcal{C}$-group. We will construct an (S, ε, δ)-homomorphism (as in Definition 2.1) from an arbitrary finite subset $S \subseteq G$ to a group $H \in \mathcal{C}$ (for a suitable norm ℓ_H). Assume $S = \{\overline{s}_1, \ldots, \overline{s}_m\}$ for some $s_1, \ldots, s_m \in \mathbf{F}$ (all $\overline{s}_i$ distinct; if $1_G \in S$, then it is represented by $1_{\mathbf{F}}$). Set $\varrho := \max\{\ell_{\mathbf{F}}(s_i) \mid i = 1, \dots, m\}$ (it is well-defined since S is finite).

Choose $\varphi\colon \mathbf{F} \to H$ to $(3\varrho, \varepsilon, \delta)$-separate N (where $H \in \mathcal{C}$ and ℓ_H is a suitable norm on H). Then define $\overline{\varphi}\colon S \to H$ by $\overline{\varphi}(\overline{s}_i) := \varphi(s_i)$ (it is well-defined, since the $\overline{s}_i$ are pairwise distinct). Clearly, $\overline{\varphi}$ is an (S, ε, δ)-homomorphism:

If $1_G = \overline{s}_i \in S$, then by assumption $s_i = 1_{\mathbf{F}}$, so that $\overline{\varphi}(1_G) = \varphi(1_{\mathbf{F}}) = 1_H$ (since φ is a homomorphism). Moreover, if $\overline{s_i s_j} = \overline{s}_k$, then $s_j^{-1} s_i^{-1} s_k \in N$ and so

$$d_H(\overline{\varphi}(\overline{s}_i)\overline{\varphi}(\overline{s}_j), \overline{\varphi}(\overline{s}_k)) = \ell_H(\varphi(s_j^{-1} s_i^{-1} s_k)) < \varepsilon,$$

since $\ell_{\mathbf{F}}(s_j^{-1} s_i^{-1} s_k) \leq 3\varrho$ by the choice of $\varrho \in \mathbb{N}$ and the triangle inequality. Lastly, if $\overline{s}_i \neq 1_G$, then $s_i \notin N$ and $\ell_{\mathbf{F}}(s_i) \leq 3\varrho$, whence

$$\ell_H(\overline{\varphi}(\overline{s}_i))) = \ell_H(\varphi(s_i)) > \delta$$

again since φ is $(3\varrho, \varepsilon, \delta)$-separating.

Let us now prove the reverse direction '$\Leftarrow$'. Let $\mathbf{F}$ have a basis $x_1, \dots, x_r$. Assume there is $\delta > 0$ such that there exists an (S, ε, δ)-homomorphism $\overline{\varphi}\colon S \to H$ for all S and $\varepsilon > 0$ (where $H \in \mathcal{C}$ is equipped with norm ℓ_H). Then choose

$$S := \overline{B_\varrho(\mathbf{F})} = \{\overline{w} \in G \mid w \in B_\varrho(\mathbf{F})\}.$$

Observe that S is finite, since $\mathbf{F}$ is of finite rank $r \in \mathbb{N}$. W.l.o.g., assume $\varrho \geq 1$, so that $\overline{x_i}^{\pm 1} \in S$ (for $i = 1, \dots, r$). Define $\varphi\colon \mathbf{F} \to H$ by $\varphi(x_i) := \overline{\varphi}(\overline{x_i})$ (and extend it by freeness of $\mathbf{F}$).

We claim that N is $(\varrho, 2\varrho\varepsilon, \delta - 2\varrho\varepsilon)$-separated by φ. At first note that for any generator of $\mathbf{F}$ we have by left-invariance of ℓ_H that

$$d_H(\overline{\varphi}(\overline{x_i})^{-1}, \overline{\varphi}(\overline{x_i}^{-1})) = d_H(\overline{\varphi}(1_G), \overline{\varphi}(\overline{x_i})\overline{\varphi}(\overline{x_i}^{-1})) < \varepsilon, \tag{2.1}$$

as φ is multiplicative up to ε and $1_G, \overline{x_i}^{\pm 1} \in S$ (by its definition). Now we can prove by induction on $n = \ell_{\mathbf{F}}(w) > 0$, $n \leq \varrho$ that

$$d_H(\varphi(w), \overline{\varphi}(\overline{w})) < (2\ell_{\mathbf{F}}(w) - 1)\varepsilon. \tag{2.2}$$

For $n = 1$ we have two cases. In the first case, $w = x_i$ is a generator, so

$$d_H(\varphi(w), \overline{\varphi}(\overline{w})) = d_H(\overline{\varphi}(\overline{x_i}), \overline{\varphi}(\overline{x_i})) = 0.$$

In the second case, $w = x_i^{-1}$ is the inverse of a generator, whence

$$d_H(\varphi(w), \overline{\varphi}(\overline{w})) = d_H(\overline{\varphi}(\overline{x_i})^{-1}, \overline{\varphi}(\overline{x_i}^{-1})) < \varepsilon$$

by Inequality (2.1). Now assume that Inequality (2.2) is true for $n < \varrho$. Then we get for $\ell_{\mathbf{F}}(w) = n+1$ that

$$\begin{aligned} d_H(\varphi(w), \overline{\varphi}(\overline{w})) &= d_H(\overline{\varphi}(\overline{x_{i_1}})^{\varepsilon_1} \cdots \overline{\varphi}(\overline{x_{i_{n+1}}})^{\varepsilon_{n+1}}, \overline{\varphi}(\overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_{n+1}}^{\varepsilon_{n+1}}})) \\ &\leq d_H(\overline{\varphi}(\overline{x_{i_1}})^{\varepsilon_1} \cdots \overline{\varphi}(\overline{x_{i_{n+1}}})^{\varepsilon_{n+1}}, \overline{\varphi}(\overline{x_{i_1}})^{\varepsilon_1} \cdots \overline{\varphi}(\overline{x_{i_n}})^{\varepsilon_n} \overline{\varphi}(\overline{x_{i_{n+1}}^{\varepsilon_{n+1}}})) \\ &\quad + d_H(\overline{\varphi}(\overline{x_{i_1}})^{\varepsilon_1} \cdots \overline{\varphi}(\overline{x_{i_n}})^{\varepsilon_n} \overline{\varphi}(\overline{x_{i_{n+1}}^{\varepsilon_{n+1}}}), \overline{\varphi}(\overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}}) \overline{\varphi}(\overline{x_{i_{n+1}}^{\varepsilon_{n+1}}})) \\ &\quad + d_H(\overline{\varphi}(\overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}}) \overline{\varphi}(\overline{x_{i_{n+1}}^{\varepsilon_{n+1}}}), \overline{\varphi}(\overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_{n+1}}^{\varepsilon_{n+1}}})) \\ &< \varepsilon + (2n-1)\varepsilon + \varepsilon = (2n+1)\varepsilon, \end{aligned}$$

as desired. Here we use Inequality (2.1) and left-invariance to estimate the first summand in the second line, the induction hypothesis and right-invariance in the third line, and the property of $\overline{\varphi}$ to be multiplicative up to ε in the fourth line. Moreover, the estimate of the last line is valid, since all the elements of G in the formula lie in S (by definition).

Now if $w \in N \cap B_\varrho(\mathbf{F})$, $w \neq 1_{\mathbf{F}}$, then

$$\ell_H(\varphi(w)) = d_H(\varphi(w), 1_H) = d_H(\varphi(w), \overline{\varphi}(\overline{w})) < (2\ell_{\mathbf{F}}(w) - 1)\varepsilon < 2\varrho\varepsilon,$$

since $\overline{w} = 1_G$ as $w \in N$ and $\overline{\varphi}(1_G) = 1_H$ by definition of an (S, ε, δ)-homomorphism. The case $w = 1_{\mathbf{F}}$ is trivial.

If $w \in B_\varrho(\mathbf{F}) \setminus N$ we get

$$\ell_H(\varphi(w)) = d_H(1_H, \varphi(w)) \geq d_H(1_H, \overline{\varphi}(\overline{w})) - d_H(\varphi(w), \overline{\varphi}(\overline{w})) > \delta - 2\varrho\varepsilon$$

by definition and the triangle inequality. Thus we have shown that N is $(\varrho, 2\varrho\varepsilon, \delta - 2\varrho\varepsilon)$-separated by the homomorphism $\varphi\colon \mathbf{F} \to H$ (with respect to the norm ℓ_H), so by choosing $\varepsilon > 0$ small enough, we are done. $\square$

The next lemma and its proof is a generalization of [**?**, Lemma 6.4].

Lemma 2.21. *If $N \trianglelefteq \mathbf{F}$ is $\mathcal{C}$-separable, then for all finite sequences $n_1, \ldots, n_k \in N$ we have $\overline{n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}} \subseteq N$ (in the pro-$\mathcal{C}$-topology on $\mathbf{F}$). When $\mathrm{rk}(\mathbf{F}) < \infty$ and $\mathcal{C}$ is closed for finite products and subgroups, the reverse implication '$\Leftarrow$' is also valid.*

Proof. We begin with the direction '$\Rightarrow$'. So assume N is $\mathcal{C}$-separable and choose a finite sequence $n_1, \ldots, n_k \in N$. In order to prove that $\overline{n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}} \subseteq N$, it is enough to prove that for $w \in \mathbf{F} \setminus N$ there is a homomorphism $\varphi\colon \mathbf{F} \to H$ to a $\mathcal{C}$-group H such that

$$\varphi(w) \notin \varphi(n_1)^H \cdots \varphi(n_k)^H \supseteq \varphi(n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}).$$

So find $\delta > 0$ such that N is $(\varrho, \varepsilon, \delta)$-separable for all $\varrho \in \mathbb{N}$, $\varepsilon > 0$ (w.r.t. $\mathcal{C}$). Take

$$\varrho \coloneqq \max\{\ell_{\mathbf{F}}(w), \ell_{\mathbf{F}}(n_1), \ldots, \ell_{\mathbf{F}}(n_k)\}$$

and let $\varphi\colon \mathbf{F} \to H$ be a homomorphism to the $\mathcal{C}$-group (equipped with norm ℓ_H) which $(\varrho, \delta/k, \delta)$-separates N. Then by definition $\ell_H(\varphi(w)) > \delta$, since $w \in B_\varrho(\mathbf{F}) \setminus N$ and $\ell_H(\varphi(n_j)) < \delta/k$, since $n_j \in N \cap B_\varrho(\mathbf{F})$ for $j = 1, \ldots, k$. Thus clearly

$$\ell_H(\varphi(w)) > \delta = k\delta/k > \sum_{j=1}^{k} \ell_H(\varphi(n_j)^{h_j}) \geq \ell_H(\varphi(n_1)^{h_1} \cdots \varphi(n_k)^{h_k})$$

for any choice of h_j $(j = 1, \ldots, k)$ as ℓ_H is a norm, so that $\varphi(w) \notin \varphi(n_1)^H \cdots \varphi(n_k)^H$, as claimed.

Now we prove the reverse implication '$\Leftarrow$' under the additional assumptions. We need to construct an $(\varrho, \varepsilon, \delta)$-separating homomorphism to a normed $\mathcal{C}$-group for some fixed $\delta > 0$ and all $\varrho \in \mathbb{N}$, $\varepsilon > 0$ arbitrary. Choose an integer $l \geq \delta/\varepsilon$. For each $w \in B_\varrho(\mathbf{F}) \setminus N$ and every finite sequence $n_1, \ldots, n_k \in N \cap B_\varrho(\mathbf{F})$ with $k \leq l$ we may choose a homomorphism $\varphi_{w,n_1,\ldots,n_k} : \mathbf{F} \to H$ $(H \in \mathcal{C})$ such that

$$\varphi_{w,n_1,\ldots,n_k}(w) \notin \varphi_{w,n_1,\ldots,n_k}(n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}})$$

since $\overline{n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}} \subseteq N$ (in the pro-$\mathcal{C}$ topology on $\mathbf{F}$). As $\mathbf{F}$ has finite rank, the set $B_\varrho(\mathbf{F})$ is finite and thus there are only finitely many possibilities for the choice of w and $n_1, \ldots, n_k$. Taking the tupling of all the corresponding homomorphisms, we get a homomorphism

$$\varphi\colon \mathbf{F} \to \prod H \text{ with } \varphi(w) \notin \varphi(n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}})$$

for any $w \in B_\varrho(\mathbf{F}) \setminus N$, $n_1, \ldots, n_k \in N \cap B_\varrho(\mathbf{F})$, $k \leq l$ and $\prod H \in \mathcal{C}$ (since $\mathcal{C}$ is closed for finite products). Since $\mathcal{C}$ is closed for subgroups, we can restrict φ to its image $I \coloneqq \operatorname{im}(\varphi)$ (which is then a $\mathcal{C}$-group), so that we have

$$\varphi(w) \notin \varphi(n_1)^I \cdots \varphi(n_k)^I.$$

Now set $\mathcal{E} \coloneqq \{\varphi(n)^I \mid n \in N \cap B_\varrho(\mathbf{F})\}$, $E \coloneqq \bigcup \mathcal{E}$ and equip I with the norm $\ell_I \coloneqq \varepsilon \ell_{\mathrm{Cay},E}$ (see Definition 0.12).

We claim that $\varphi\colon \mathbf{F} \to I$ $(\varrho, \varepsilon', \delta)$-separates N for $\varepsilon' > \varepsilon$ (I being equipped with norm ℓ_I):

At first choose $w \in B_\varrho(\mathbf{F}) \setminus N$. Then by construction

$$\varphi(w) \notin \{1_I\}\varphi(n_1)^I \cdots \varphi(n_k)^I$$

for any $k \leq l$ and $n_1, \ldots, n_k \in N \cap B_\varrho(\mathbf{F})$, whence $\ell_I(\varphi(w)) > \varepsilon l \geq \delta$. But if we pick $n \in N \cap B_\varrho(\mathbf{F})$, we have

$$\varphi(n) \in \{1_I\}\varphi(n)^I,$$

so $\ell_I(\varphi(n)) \leq \varepsilon < \varepsilon'$. This completes the proof. □

Proof of Theorem 2.18. The result immediately follow from Lemmas 2.20 and 2.21. □

Remark 2.22. If $\mathcal{C}$ is closed with respect to finite products and subgroups, Theorem 2.18 implies that residually $\mathcal{C}$-groups are $\mathcal{C}$-approximable as abstract groups, since, if $G = \mathbf{F}/N$ is a finitely generated such group, for each finite sequence $n_1, \ldots, n_k \in N$ we obtain $\overline{n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}} \subseteq \overline{N} = N$ (in the pro-$\mathcal{C}$ topology on $\mathbf{F}$).

In view of Theorem 2.18, when $\mathcal{C}$ is closed for subgroups, to prove the existence of a non-$\mathcal{C}$-approximable group, it suffices to find a normal subgroup $N \trianglelefteq \mathbf{F}$ of a free group of finite rank, an element $x \in \mathbf{F} \setminus N$, and a sequence $n_1, \ldots, n_k \in N$ such that $\varphi(x) \in \varphi(n_1)^H \cdots \varphi(n_k)^H$ for any surjective homomorphism $\varphi\colon \mathbf{F} \to H$ to a $\mathcal{C}$-group.

As both classes **Nil** and **Sol** are closed with respect to subgroups, we shall construct a situation as described before.

Subsequently, let $\mathbf{F}$ be freely generated by $x_1, \ldots, x_r$. Fix a presentation $\mathbf{F}/N$ of some non-trivial perfect group P and an element $x \in \mathbf{F} \setminus N$. The assumption that P is perfect is equivalent to the fact that $\mathbf{F} = \mathbf{F}'N$. Hence we can find $n_1, \ldots, n_r, n \in N$ such that $x_i \equiv n_i$ $(i = 1, \ldots, r)$ and $x \equiv n$ modulo $\mathbf{F}'$. Consider a surjective homomorphism $\varphi\colon \mathbf{F} \twoheadrightarrow H$ to some finite group H (later H will be assumed to be nilpotent resp. solvable). Writing $y_i := \varphi(x_i)$, $h_i := \varphi(n_i)$ $(i = 1, \ldots, r)$, $y := \varphi(x)$, and $h := \varphi(n)$, the above translates to $y_i \equiv h_i$ $(i = 1 \ldots, r)$ and $y \equiv h$ modulo H'. Clearly, $h_1, \ldots, h_r$ generate H modulo H' (as φ is surjective). Now we need a lemma. To state it, we need some notation from the Section 0.1(b). Recall that in a group G we write $[g, h] = g^{-1}h^{-1}gh = g^{-1}g^h$ for the commutator of the elements $g, h \in G$. For $S \subseteq G$ and $g \in G$ we write $[S, g]$ for the set $\{[s, g] \mid s \in S\}$, and for subgroups $K, L \leq G$ write $[K, L]$ for the subgroup generated by $\{[k, l] \mid k \in K,\, l \in L\}$.

Lemma 2.23 (Proposition 1.2.5 of [65])**.** *Let $L \trianglelefteq G$ be groups and suppose that $G = G'\langle g_1, \ldots, g_m\rangle$. Then*

$$[L, G] = [L, g_1] \cdots [L, g_m][L,_l G]$$

for all $l \geq 1$. Here $[L,_l G]$ denotes the subgroup $[[L, \underbrace{G], \ldots, G}_{l}]$ of G.

Proof of Theorem 2.17, Part 1. We apply Lemma 2.23 to $G := L := H$, $m := r$, and $g_i := h_i$ $(i = 1, \ldots, r)$. Moreover, we choose $l \geq 1$ to be an integer such that $\gamma_l(H) = \gamma_\omega(H)$ (recall from Section 0.1(c) that $\gamma_i(H)$ is the ith term in the lower central series of H and $\gamma_\omega(H) = \bigcap_{i \in \mathbb{Z}_+} \gamma_i(H)$). Hence there exist $l_{ij}, l_j \in H$ $(i, j = 1 \ldots, r)$ such that $y_i \equiv h_i[l_{i1}, h_1] \cdots [l_{ir}, h_r]$ $(i = 1, \ldots, r)$ and $y \equiv h[l_1, h_1] \cdots [l_r, h_r]$ modulo $\gamma_\omega(H)$. Assuming H is nilpotent (so $\gamma_\omega(H) = \mathbf{1}$), the last congruence shows that

$$y = \varphi(x) \in \varphi(n_1')^H \cdots \varphi(n_{k_{\mathbf{Nil}}}')^H,$$

where $k_{\mathbf{Nil}} = 2r + 1$ and $(n_j')_{j=1}^{k_{\mathbf{Nil}}}$ is the sequence $(n, n_1^{-1}, n_1, \ldots, n_r^{-1}, n_r)$. Thus P cannot be **Nil**-approximable. □

To prove that P is not **Sol**-approximable, we need the following deeper result of Segal on finite solvable groups:

Theorem 2.24 (Theorem 2.1 of [63]). *Assume G is a finite solvable group and*

$$\gamma_\omega(G)\langle g_1, \dots, g_m\rangle = G$$

for some $m \in \mathbb{N}$. Moreover, assume that G is generated by d elements. Then there is a fixed sequence $(i_j)_{j=1}^{m'}$ of indices in $\{1, \dots, m\}$, whose entries and length m' only depend on d and m, such that

$$\gamma_\omega(G) = \prod_{j=1}^{m'} [\gamma_\omega(G), g_{i_j}].$$

Proof of Theorem 2.17, Part 2. Assume that H is solvable. We want to apply Theorem 2.24 to $G \coloneqq H$. Since φ is surjective, the elements $y_1 = \varphi(x_1), \dots, y_r = \varphi(x_r)$ generate H, so we may set $d \coloneqq r$. We still have to define the elements $g_1, \dots, g_m \in G$. From the above congruences we conclude that the sequence

$$h_1, \dots, h_r, (h_1^{-1})^{l_{11}}, \cdots, (h_r^{-1})^{l_{1r}}, \cdots, (h_1^{-1})^{l_{r1}}, \cdots, (h_r^{-1})^{l_{rr}}$$

is a good choice for $g_1, \dots, g_m$. Thus $m \coloneqq r(r+1)$ is bounded in terms of r.

The theorem gives us, similarly as in the nilpotent case, a fixed sequence $(n''_j)_{j=1}^{k_{\mathbf{Sol}}}$ with entries in $\{n, n_1^{\pm 1}, \cdots, n_r^{\pm 1}\}$, whose length $k_{\mathbf{Sol}} = k_{\mathbf{Nil}} + 2m'$ is bounded in terms of r, such that

$$y = \varphi(x) \in \varphi(n''_1)^H \cdots \varphi(n''_{k_{\mathbf{Sol}}})^H.$$

Thus P cannot be **Sol**-approximable. □

Note that finite generation is crucial here. Indeed, there exist countably infinite locally finite-p groups which are perfect and even characteristically simple [53]. By Example 2.5, these groups are **Nil**-approximable (since finite p-groups are nilpotent), but by definition they are not finitely generated. It is known that locally finite-solvable groups cannot be non-abelian simple [60, page 154], but it seems to be an open problem if there exist **Sol**-approximable non-abelian simple groups.

2.4 On Fin-approximable groups

Let **PSL** be the class of simple groups of type $\mathrm{PSL}_n(q)$, i.e., $n \in \mathbb{N}_{\geq 2}$ and q is a prime power and $(n, q) \neq (2, 2), (2, 3)$, and recall that **Fin** is the class of all finite groups. In this section, we prove the following result, which is motivated by a private note of Nikolov.

Theorem 2.25. *Any non-trivial finitely generated* **Fin**-*approximable group has a non-trivial* **PSL**-*approximable quotient. In particular, every simple* **Fin**-*approximable group is* **PSL**-*approximable.*

To prove Theorem 2.25 we need some preparation. At first we recall a classical lemma of Goursat [28]:

Lemma 2.26 (Goursat's lemma)**.** *Let $G \leq K \times L$ be a subdirect product, i.e., the restricted projection maps $\pi_K\colon G \to K$, $\pi_L\colon G \to L$ are surjective. Set $M := \ker(\pi_L)$ and $N := \ker(\pi_K)$. Then $M \trianglelefteq K$, $N \trianglelefteq L$, and the image of G in $K/M \times L/N$ is the graph of an isomorphism.*

We need Lemma 2.26 for the following auxiliary result (Lemma 2.27 below). Recall that a profinite group is called semisimple if it is the direct product of finite non-abelian simple groups. Moreover, a finite group G is *almost simple* if it has a unique minimal normal subgroup N which is non-abelian simple; in this case $N \trianglelefteq G \leq \mathrm{Aut}(N)$.

Lemma 2.27. *Let G be a closed subdirect product of a profinite group $A = \prod_{i\in I} A_i$, where A_i is almost simple ($i \in I$). Then G contains a closed normal semisimple subgroup H such that G/H is solvable of derived length at most three and each simple factor of H is normal in G.*

Proof. For $J \subseteq I$ let $\pi_J\colon A \to \prod_{j\in J} A_j$ be the projection maps. Then by Proposition 1.2.2 of [77] the group G is the inverse limit of the groups $\pi_J(G)$ ($J \subseteq I$ finite) together with the natural maps $\pi_J(G) \to \pi_{J'}(G)$ for $J' \subseteq J$.

Using Goursat's lemma, one can show by induction on $|J|$ that for $J \subseteq I$ finite there exist $r \in \mathbb{N}$ and finite non-abelian simple groups $S_1, \dots, S_r$ such that $S_1 \times \cdots \times S_r \trianglelefteq \pi_J(G) \leq \mathrm{Aut}(S_1) \times \cdots \times \mathrm{Aut}(S_r)$. In this situation, for $j_0 \in I \setminus J$ the projection $\pi_{J\cup\{j_0\}}(G) \to \pi_J(G)$ either is an isomorphism or there exists a finite non-abelian simple group S_{r+1} such that $S_1 \times \cdots \times S_{r+1} \trianglelefteq \pi_{J\cup\{j_0\}}(G) \leq \mathrm{Aut}(S_1) \times \cdots \times \mathrm{Aut}(S_{r+1})$ and the restriction of $\pi_{J\cup\{j_0\}}(G) \to \pi_J(G)$ to the socle $S_1 \times \cdots \times S_{r+1}$ of $\pi_{J\cup\{j_0\}}$ is the natural projection onto $S_1 \times \cdots \times S_r$ (the socle of $\pi_J(G)$).

Now it is clear that G, as the inverse limit of the groups $\pi_J(G)$ ($J \subseteq I$ finite) and the maps $\pi_J(G) \to \pi_{J'}(G)$, contains the inverse limit H of the socles of these groups together with the restricted maps. It is routine to check that H has the desired properties. The fact that G/H is solvable of derived length at most three is implied by Schreier's conjecture. □

Now we can start with the proof of Theorem 2.25: We will prove that our group has a non-trivial **PSL**-approximable quotient, where we endow the groups $\mathrm{PSL}_n(q)$ with the normalized conjugacy length function ℓ_c (see Definition 0.12).

If the group in the theorem is not perfect, it has a non-trivial cyclic quotient, which clearly has the desired property. So let $P = \mathbf{F}/N$ be perfect, where $\mathbf{F}$ is freely generated by $x_1, \dots, x_r$ and $N \trianglelefteq \mathbf{F}$. Let $\widehat{\mathbf{F}}$ be the profinite completion of $\mathbf{F}$ and $M := \langle\!\langle N \rangle\!\rangle_{\widehat{\mathbf{F}}}$ be the normal closure of N in $\widehat{\mathbf{F}}$. Identifying $\mathbf{F}$ with its image in the profinite completion, it follows from Theorem 2.18 that P is **Fin**-approximable if and only if $N = M \cap \mathbf{F}$, since for a sequence $n_1, \dots, n_k \in N$ we have

$$\overline{n_1^{\mathbf{F}} \cdots n_k^{\mathbf{F}}} = \mathbf{F} \cap n_1^{\widehat{\mathbf{F}}} \cdots n_k^{\widehat{\mathbf{F}}},$$

where the closure on the left is taken in $\mathbf{F}$. This is equivalent to saying that the map $\mathbf{F} \to \widehat{\mathbf{F}}/M$ induces an embedding of P in $\widehat{\mathbf{F}}/M$.

For a profinite group G set $G_0 \coloneqq \bigcap\{O \trianglelefteq_{\mathrm{o}} G \,|\, G/O \text{ is almost simple}\}$ (here $O \trianglelefteq_{\mathrm{o}} G$ means that O is an open normal subgroup of G).

Claim 2.28. *It holds that* $\mathbf{F} \not\leq \widehat{\mathbf{F}}_0 M$.

Proof. Assume the contrary. Then by perfectness of P there are $y_i \in N$ such that $x_i\mathbf{F}' = y_i\mathbf{F}'$ $(i = 1, \ldots, r)$. By assumption, there are also $z_i \in M$ such that $x_i\widehat{\mathbf{F}}_0 = z_i\widehat{\mathbf{F}}_0$ $(i = 1, \ldots, r)$. Set $Y \coloneqq \{y_1^{\pm 1}, \ldots, y_r^{\pm 1}, z_1^{\pm 1}, \ldots, z_r^{\pm 1}\}$. As $\widehat{\mathbf{F}}_0$ is closed, by definition we have that

$$\widehat{\mathbf{F}} = \overline{\widehat{\mathbf{F}}_0\langle Y\rangle} = \widehat{\mathbf{F}}_0\overline{\langle Y\rangle} \quad \text{and} \quad \widehat{\mathbf{F}} = \overline{\mathbf{F}'\langle Y\rangle} = \overline{\widehat{\mathbf{F}}'\langle Y\rangle},$$

where all closures are taken in $\widehat{\mathbf{F}}$. Hence by Theorem 1.7 of [?] applied to $G \coloneqq \widehat{\mathbf{F}}$ and $H \coloneqq \widehat{\mathbf{F}}_0$ we get that $M \geq \langle\!\langle Y\rangle\!\rangle_{\widehat{\mathbf{F}}} \geq [\widehat{\mathbf{F}}_0, \widehat{\mathbf{F}}] \geq \widehat{\mathbf{F}}_0'$. Since $\widehat{\mathbf{F}}_0 M/M = \widehat{\mathbf{F}}_0/(\widehat{\mathbf{F}}_0 \cap M)$ is abelian by the preceding argument, we cannot have $\mathbf{F} \leq \widehat{\mathbf{F}}_0 M$, since otherwise P would be abelian and hence trivial. Contradiction proving the claim. □

Claim 2.28 implies that P has a non-trivial homomorphism to $\widehat{\mathbf{F}}/\widehat{\mathbf{F}}_0 M$.

Apply Lemma 2.27 to $G \coloneqq \widehat{\mathbf{F}}/\widehat{\mathbf{F}}_0$ as a subdirect product of all almost simple quotients of $\widehat{\mathbf{F}}$. Let $H = K/\widehat{\mathbf{F}}_0$ be the semisimple group provided by this lemma. As $\widehat{\mathbf{F}}/K$ is solvable, we cannot have $K \leq \widehat{\mathbf{F}}_0 M$, otherwise the image of P in $\widehat{\mathbf{F}}/\widehat{\mathbf{F}}_0 M$ would be trivial, contradicting Claim 2.28.

Hence $(K \cap \widehat{\mathbf{F}}_0 M)/\widehat{\mathbf{F}}_0$ is a proper normal subgroup of the semisimple group $H = K/\widehat{\mathbf{F}}_0 = \prod_{i\in I} S_i$, where S_i $(i \in I)$ are the simple factors, so by Theorem 5.12 of [?] it is contained in a maximal normal subgroup $L/\widehat{\mathbf{F}}_0$ of the former (to fulfill the hypothesis of this theorem, we need that $\mathrm{rk}(\mathbf{F}) < \infty$). By the same result, K/L is isomorphic (as an abstract group) to a metric ultraproduct of the S_i equipped with the normalized conjugacy length function $\ell_{\mathrm{c},i}$ $(i \in I)$. Note that in this situation L is even normal in $\widehat{\mathbf{F}}$, since $\ell_{\mathrm{c},i}$ is left invariant under $\mathrm{Aut}(S_i)$ and $S_i \trianglelefteq \widehat{\mathbf{F}}/\widehat{\mathbf{F}}_0$ by Lemma 2.27 $(i \in I)$.

Claim 2.29. *In this setting we have* $\mathbf{F} \not\leq LM$.

Proof. Otherwise $[\mathbf{F}, K] \leq [LM, K] \leq L[M, K] \leq L(M \cap K) = L$. Here the first inclusion holds by assumption, whereas the second follows from the commutator identity $[lm, k] = [l, k][[l, k], m][m, k]$ for $k \in K$, $l \in L$ and $[l, k] \in L$, since $L \trianglelefteq K$, and $[[l, k], m], [m, k] \in [K, M] = [M, K]$. The last inclusion holds as $M, K \trianglelefteq \widehat{\mathbf{F}}_0$, and the final equality by the choice of L. Hence $\widehat{\mathbf{F}}_0[\mathbf{F}, K] \leq L$.

Let S be a simple factor of H. $\widehat{\mathbf{F}}/\widehat{\mathbf{F}}_0$ maps continuously on the finite discrete group $\mathrm{Aut}(S)$ via the conjugation action. The image of this map clearly contains the inner automorphisms, since these are induced by S itself. The elements $\overline{x}_1, \ldots, \overline{x}_r$ generate a dense subgroup of $\widehat{\mathbf{F}}/\widehat{\mathbf{F}}_0$, which must induce all inner automorphisms of S by the previous fact.

As S has trivial center, we have $|S/\mathbf{C}_S(\overline{x}_{i_0})| = |[S, \overline{x}_{i_0}]| \geq |S|^{1/r}$ for some $i_0 \in \{1, \ldots, r\}$. Lemma 3.5 of [?] implies that $\prod_{i=1}^r [S, \overline{x}_i][S, \overline{x}_i^{-1}]$ contains the normal subsets $[S, \overline{x}_i]^S \subseteq S$ for $i = 1, \ldots, r$. Since $|[S, \overline{x}_{i_0}^S]| \geq |S|^{1/r}$, by Theorem 1.1 of [49] there is

$e \in \mathbb{N}$ only depending on r such that

$$\left(\prod_{i=1}^{r} [S, \overline{x_i}][S, \overline{x_i}^{-1}]\right)^{*e} = S.$$

This implies that $K \leq \widehat{\mathbf{F}}_0[\mathbf{F}, K]$, since e is independent of the simple factor S. But then $K \leq L$, a contradiction. $\square$

From the previous claim we deduce that P still has a non-trivial homomorphism to $\widehat{\mathbf{F}}/LM$. Since $\widehat{\mathbf{F}}/KM$ is solvable as a quotient of $\widehat{\mathbf{F}}/K$, this homomorphism restricts to KM/LM, which is a non-trivial homomorphic image of the metric ultraproduct K/L. Since the latter is simple by Proposition 3.1 of [67], we are only left to show that K/L, which is a metric ultraproduct of the sequence $(S_i)_{i\in I}$ of finite simple groups from above equipped with normalized conjugacy length function with respect to some ultrafilter $\mathcal{U}$, embeds into a metric ultraproduct of groups $\mathrm{PSL}_{n_i}(q_i)$ $(i \in I)$ equipped with the normalized conjugacy length function, since then P would have the same property.

Let us briefly sketch the argument for this: Firstly, if the limit of the ranks of the groups S_i $(i \in I)$ is bounded along the ultrafilter $\mathcal{U}$ (where the rank of the alternating group A_n is defined to be n and the sporadic groups are also considered as groups of bounded rank) the resulting ultraproduct will be a simple group of Lie type over a pseudofinite field k or an alternating group A_n, respectively. In the first case, it clearly embeds into $\mathrm{PSL}_n(k)$ for $n \in \mathbb{N}$ appropriately chosen. However, the latter is a metric ultraproduct of groups $\mathrm{PSL}_n(q_i)$ $(i \in I)$ equipped with the normalized conjugacy length function for some sequence $(q_i)_{i\in I}$ of prime powers. The second case is similar.

Hence we may assume that our ultraproduct does not involve finite simple groups from families of bounded rank.

Furthermore, we can assume that it contains no alternating groups, as we can replace any alternating group A_n by $\mathrm{PSL}_n(q)$ for some prime power q. Namely, the natural embedding $\mathrm{A}_n \hookrightarrow \mathrm{PSL}_n(q)$, where $\mathrm{PSL}_n(q)$ is equipped with the normalized projective rank length function ℓ_{pr}, induces norm $\frac{1}{n}\ell_{\mathrm{Cay},\tau^{\mathrm{S}_n}}$ on A_n with respect to the conjugacy class of a transposition τ of the ambient symmetric group S_n. The latter is Lipschitz equivalent to the normalized conjugacy length function on A_n by Fact 0.13.

Hence we can assume that all groups S_i $(i \in I)$ are classical Chevalley or Steinberg groups equipped with the normalized conjugacy length function. Applying Fact 0.13 once again, we can replace the norms on the S_i $(i \in I)$ by normalized projective rank length functions and embed our metric ultraproduct K/L into a metric ultraproduct of groups $\mathrm{PSL}_{n_i}(q_i)$ $(i \in I)$ equipped with the normalized projective rank length function. Thus K/L is **PSL**-approximable and so P must have a non-trivial **PSL**-approximable quotient, as wished. The proof is complete.

2.5 On the approximability of Lie groups

In this section, we utilize the following theorem of Nikolov and Segal to deduce two results concerning the approximability of Lie groups by finite groups and one result on compactifications of pseudofinite groups.

Theorem 2.30 (Theorem 1.2 of [**?**]). *Let $g_1, \ldots, g_m$ be a symmetric generating set for the finite group G. If $K \trianglelefteq G$, then*

$$[K, G] = \left(\prod_{j=1}^{m} [K, g_j] \right)^{*e},$$

where e only depends on m.

Remark 2.31. It was remarked in [**?**] that it was an open problem at the time of writing to decide whether a finite product of conjugacy classes in a non-abelian free group is always closed in the profinite topology.

It is a rather straightforward consequence of Theorem 2.30 that this is not the case. Indeed, the theorem implies that in $\mathbf{F} = \langle x_1, \ldots, x_m \rangle$ the profinite closure of the product of the $2me$ conjugacy classes of $x_1^{-1}, x_1, \ldots, x_m^{-1}, x_m$ contains the entire commutator subgroup, but it is a well-known fact (see Theorem 3.1.2 of [65]) that the commutator width in this group is infinite if $m > 1$.

This implication was first observed by Segal and independently discovered by Gismatullin.

Actually, we shall use the following immediate corollary of Theorem 2.30:

Corollary 2.32. *Let G be a quotient of a product of finite groups, then for $g, h \in G$ and $N \in \mathbb{N}$ we have*

$$[g^N, h^N] \in \left([G, g][G, g^{-1}][G, h][G, h^{-1}]\right)^{*e}$$

for some fixed constant $e \in \mathbb{N}$.

Recall that **Fin** denotes the class of all finite groups. At first we prove the following theorem:

Theorem 2.33. *A connected Lie group is* **Fin***-approximable as a topological group if and only if it is abelian.*

By Lemma 2.16, we already know that connected abelian Lie groups are **Fin**-approximable. So we are only left to prove that a **Fin**-approximable connected Lie group is actually abelian. This will be a consequence of the following auxiliary result:

Lemma 2.34. *Let $\varphi, \psi \colon \mathbb{R} \to (H_{\mathcal{U}}, \ell_{\mathcal{U}}) = \prod_{\mathcal{U}} (H_i, \ell_i)$ be continuous homomorphisms into a metric ultraproduct of finite groups H_i with norm ℓ_i ($i \in I$). Then for all $s, t \in \mathbb{R}$ it holds that $[\varphi(s), \psi(t)] = 1_{\overline{H}}$.*

Let us first prove Theorem 2.33 using Lemma 2.34.

Proof of Theorem 2.33. Assume L is a connected **Fin**-approximable Lie group. Then there is an embedding $\iota\colon L \hookrightarrow (H_{\mathcal{U}}, \ell_{\mathcal{U}}) = \prod_{\mathcal{U}} (H_i, \ell_i)$ into a metric ultraproduct of normed finite groups. If $a, b \in L$ are in the image of the exponential map, Lemma 2.34 implies that $\iota(a)$ and $\iota(b)$ commute. So as ι is injective, a and b commute. Hence by connectedness $L = L^0$ is abelian. This ends the proof. $\square$

We are still left to prove Lemma 2.34:

Proof of Lemma 2.34. For $\varepsilon > 0$ by continuity we can choose $N \in \mathbb{Z}_+$ large enough such that $\ell_{\mathcal{U}}(\varphi(s/N)), \ell_{\mathcal{U}}(\psi(t/N)) < \varepsilon$. Set $G := H_{\mathcal{U}}$, $g := \varphi(s/N)$, and $h := \psi(t/N)$ and apply Corollary 2.32. This gives

$$[\varphi(s), \psi(t)] = [g^N, h^N] \in \big([H_{\mathcal{U}}, g][H_{\mathcal{U}}, g^{-1}][H_{\mathcal{U}}, h][H_{\mathcal{U}}, h^{-1}]\big)^{*e},$$

whence $\ell_{\mathcal{U}}([\varphi(s), \psi(t)]) < 8e\varepsilon$ by invariance of $\ell_{\mathcal{U}}$ and the triangle inequality. Since $\varepsilon > 0$ was arbitrary, the proof is complete. $\square$

Note that Theorem 2.33 provides an answer to Question 2.11 of Doucha [10] whether there are groups with norm that do not embed into a metric ultraproduct of normed finite groups. Since every compact Lie group can be equipped with a norm that generates its topology, every such group with non-abelian identity component is an example of such a group by Theorem 2.33. (Indeed, the theorem even provides topological types of groups which cannot occur as subgroups of such a metric ultraproduct.)

Before we continue with our next result, let us state the following two remarks.

Remark 2.35. In Theorem 2.33, the topology of the Lie group matters. Indeed, any linear Lie group is **Fin**-approximable as an abstract group by Remark 2.3, since all its finitely generated subgroups are residually finite by Malcev's theorem and hence **Fin**-approximable by Remark 2.22.

Thus any linear Lie group L is embeddable (as an abstract group) into a metric ultraproduct of normed finite groups indexed over, say, the partially ordered set of pairs consisting of a finite subset of L and a positive rational number. We will now show that we can even choose this index set to be $\mathbb{N}$. Namely, if $L \leq \mathrm{SL}_n(\mathbb{C})$, then L can be embedded into the algebraic ultraproduct $\prod_{\mathcal{U}} \mathrm{SL}_n(p^{p!})$, where $\mathcal{U}$ is a non-principal ultrafilter on the set of prime numbers. Indeed, this ultraproduct is isomorphic to $\mathrm{SL}_n(k)$, where $k = \prod_{\mathcal{U}} \mathbb{F}_{p^{p!}}$ is a pseudofinite field. Now it is straightforward to see that k contains the field $k_0 = \prod_{\mathcal{U}} \mathbb{F}_p$ together with its algebraic closure $k' := \overline{k}_0$. However, k' is an algebraically closed field of characteristic zero and cardinality $2^{\aleph_0}$ (a result due to Shelah [66]) and hence isomorphic to $\mathbb{C}$. Note that, if we view the above algebraic ultraproduct as a metric ultraproduct, the induced topology on $\mathrm{SL}_n(\mathbb{C})$ is discrete.

Since some non-linear Lie groups admit finitely presented subgroups which are not residually finite [8], it is clear that such embeddings cannot exist without the assumption of linearity.

Remark 2.36. When one approximates with symmetric groups, one cannot even embed the real line $\mathbb{R}$ in a metric ultraproduct of such groups with norm. E.g., for the symmetric group S_n it holds that all norms ℓ on it satisfy $\ell(\sigma^l) \leq 3\ell(\sigma)$ for every $l \in \mathbb{Z}$ and $\sigma \in \mathrm{S}_n$.

Let us demonstrate this briefly: First assume that σ consists of just one n-cycle. Then for any $m \in \mathbb{Z}_+$ coprime to n we have that σ^m is as well an n-cycle and hence conjugate to σ. Recall from Section 0.2 that for a subset $S \subseteq R$ of a ring R we set

$$S^{+k} := \{s_1 + \cdots + s_k \mid s_i \in S\}.$$

At first assume that $l = 2k$ is even. Then it holds that $\bar{l} = 2\bar{k} \in ((\mathbb{Z}/(n))^\times)^{+2} \subseteq \mathbb{Z}/(n)$. This follows from the Chinese remainder theorem and the easy facts that

$$((\mathbb{Z}/(p^e))^\times)^{+2} = \mathbb{Z}/(p^e) \text{ for every prime } p > 2 \text{ and } ((\mathbb{Z}/(2^e))^\times)^{+2} = (2) \subseteq \mathbb{Z}/(2^e),$$

where $e \geq 1$.

If l is odd, then $\bar{l} \in ((\mathbb{Z}/(n))^\times)^{+3} \subseteq \mathbb{Z}/(n)$, which follows as previously from

$$((\mathbb{Z}/(p^e))^\times)^{+3} = \mathbb{Z}/(p^e) \text{ for every prime } p > 2 \text{ and } ((\mathbb{Z}/(2^e))^\times)^{+3} = 1 + (2) \subseteq \mathbb{Z}/(2^e),$$

for $e \geq 1$.

In total, if l is even, we find $m_1, m_2 \in \mathbb{Z}_+$ coprime to n such that $\sigma^l = \sigma^{m_1}\sigma^{m_2}$, and if l is odd, we find $l_1, l_2, l_3 \in \mathbb{Z}_+$ coprime to n such that $\sigma^l = \sigma^{l_1}\sigma^{l_2}\sigma^{l_3}$. Since the $\sigma^{m_i}, \sigma^{l_j}$ $(i = 1, 2,\ j = 1, 2, 3)$ are conjugate to σ, applying ℓ to both sides of these two equations, using the triangle inequality and invariance of ℓ, this immediately yields the inequality $\ell(\sigma^l) \leq 3\ell(\sigma)$.

If σ is not an n-cycle, we can apply the previous construction on every cycle of σ. The proof is complete.

Using the above inequality, it is simple to deduce that the only continuous homomorphism of $\mathbb{R}$ into a metric ultraproduct of finite symmetric groups with norm is trivial.

Referring to the question of Zilber [79, page 17] (also Question 1.1 of Pillay [59]) whether a compact simple Lie group can be a quotient of the algebraic ultraproduct of finite groups, we present the following second application of Corollary 2.32:

Theorem 2.37. *A Lie group equipped with a norm generating its topology that is an abstract quotient of a product of finite groups has abelian identity component.*

The proof of this result is almost identical to the proof of Theorem 2.33.

Proof. Let (L, ℓ_L) be such a Lie group with norm and $a, b \in L$ be in the image of the exponential map. For $\varepsilon > 0$ we find $N \in \mathbb{Z}_+$, $g, h \in L$ such that $\ell_L(g), \ell_L(h) < \varepsilon$ and

$g^N = a$, $h^N = b$. Then applying Corollary 2.32 to $G \coloneqq L$ yields

$$[a, b] = [g^N, h^N] \in \big([L, g][L, g^{-1}][L, h][L, h^{-1}]\big)^{*e},$$

whence $\ell_L([a, b]) < 8e\varepsilon$ by the invariance of ℓ_L and the triangle inequality. This shows that a and b commute. Hence, as L^0 is generated by the image of the exponential map, it must be abelian. □

Theorem 2.37 implies that any compact simple Lie group, the simplest example being $\mathrm{SO}_3(\mathbb{R})$, is not a quotient of a product of finite groups, answering Zilber's question (and hence also answers Question 1.1 of Pillay [59]).

Moreover, Theorem 2.37 remains valid if we replace the product of finite groups by a *pseudofinite group*, i.e., a group which is a model of the theory of all finite groups (as Corollary 2.32 is still valid in this case, with the same proof). It then also provides a negative answer to Question 1.2 of Pillay [59] whether there is a surjective homomorphism from a pseudofinite group to a compact simple Lie group.

Before we state the last theorem of this section, we digress briefly by pointing out a further application of Theorems 2.33 and 2.37.

Referring to [74], we call a compact group G with compatible norm ℓ_G *Turing-approximable* if for all $\varepsilon > 0$ there is a finite subset $S_\varepsilon \subseteq G$, a group H_ε, and a bijection $\gamma_\varepsilon \colon S_\varepsilon \to H_\varepsilon$ such that for all $g \in G$ there is $s \in S_\varepsilon$ with $d_G(g, s) < \varepsilon$ and $d_G(gh, \gamma_\varepsilon^{-1}(\gamma_\varepsilon(g)\gamma_\varepsilon(h))) < \varepsilon$ for $g, h \in S_\varepsilon$. Define for $g \in H_\varepsilon$

$$\ell_\varepsilon(g) \coloneqq |H_\varepsilon|^{-2} \sum_{f,h \in H_\varepsilon} d_G(\gamma_\varepsilon^{-1}(fgh), \gamma_\varepsilon^{-1}(fh)).$$

It is routine to check that ℓ_ε is a norm on H_ε and that for all $g \in H_\varepsilon$ we have

$$|\ell_\varepsilon(g) - \ell_G(\gamma_\varepsilon^{-1}(g))| < 3\varepsilon.$$

Set $\delta_\varepsilon \colon G \to S_\varepsilon$ such that $d_G(\delta_\varepsilon(g), g)$ is minimal for all $g \in G$.

In this situation we can apply Lemma 2.10, setting $I \coloneqq \mathbb{Z}_+$, $\mathcal{U}$ to be a non-principal ultrafilter on I, $K_i \coloneqq H_{1/i}$, and $\varphi_i \coloneqq \gamma_{1/i} \circ \delta_{1/i}$. Again one checks easily that we may apply (i) and (ii) of this lemma. Hence a Turing-approximable group is isomorphic to a metric ultraproduct of normed finite groups. Thus Theorem 2.33 as well as Theorem 2.37 imply that a Turing-approximable Lie group has abelian identity component. This is the main result of [74]. By Lemma 3.4 of [23], the latter condition is also sufficient for a compact Lie group to be Turing-approximable.

Let us now turn to pseudofinite groups. By a *compactification* of an abstract group G we mean a compact group C together with a homomorphism $\iota \colon G \to C$ with dense image. Pillay conjectured that the Bohr compactification (i.e., the universal compactification) of a pseudofinite group has abelian identity component (Conjecture 1.7 in [59]). We answer this conjecture in the affirmative by the following result:

Theorem 2.38. *Let G be a pseudofinite group. Then the identity component of any compactification C of G is abelian.*

The proof is again just an easy application of Corollary 2.32.

Proof. As G is pseudofinite it satisfies the statement of Corollary 2.32 (and so does its image in C). An easy compactness argument shows that C has the same property. Now let $\varrho_i \colon C \to \mathrm{GL}(V_i)$ be the irreducible unitary representations of C and L_i the image of ϱ_i $(i \in I)$. By the Peter–Weyl theorem, C embeds continuously into $\prod_{i \in I} L_i$, and so C^0 embeds into $\prod_{i \in I} L_i^0$.

But as L_i is a compact quotient of C, Corollary 2.32 holds in it, and so, as in the proof of Theorem 2.37, it follows that L_i^0 is abelian $(i \in I)$. But then C^0 must be abelian as well, from the above embedding. $\square$

Chapter 3

Word maps are surjective on metric ultraproducts

3.1 Introduction

Recently, there has been increasing interest in word maps on finite, algebraic, and topological groups [2, 3, 16, 22, 26, 27, 31, 32, 36, 41, 43, 44, 45, 46, 47, 51]. Recall that for a word $w \in \mathbf{F}_r$, where $\mathbf{F}_r$ denotes the free group of rank r freely generated by $x_1, \ldots, x_r$ (see the beginning of Section 0.1), and a group G the symbol $w(g_1, \ldots, g_r)$ denotes the evaluation at w of the homomorphism $\mathbf{F}_r \to G$ which is defined by $x_i \mapsto g_i$ for $i = 1, \ldots, r$. We call the map $G^r \to G$ which sends $(g_1, \ldots, g_r) \in G^r$ to $w(g_1, \ldots, g_r) \in G$ the *word map* associated to w and write $w(G) \subseteq G$ for its image.

Subsequently, fix a non-trivial word $w \in \mathbf{F}_r$. For a fixed finite group G the word image $w(G)$ can just be $\{1_G\}$ if w is a law for G – when G is a finite simple group, the possible word images were characterized by Lubotzky in [51]. In analogy, for a fixed compact group G the word image $w(G)$ can be contained in any neighborhood of the identity, as was proved by Thom in [72, Corollary 1.2]. However, examples show that for fixed w and a family $\mathcal{G}$ of finite simple groups resp. compact connected simple Lie groups, for $G \in \mathcal{G}$ one should expect $w(G)$ to be *large* in G if the order resp. dimension or rank of G is large.

There are two intriguing conjectures regarding this observation. Letting $\mathcal{G}$ be the class of finite non-abelian simple groups, Shalev conjectured [3, Conjecture 8.3] that, if w is not a proper power, the associated word map on G is surjective if the order of G is sufficiently large. Similarly, if $\mathcal{G}$ is the class of simple connected compact groups, Larsen conjectured at the 2008 Meeting of the AMS in Bloomington that w is surjective on $G \in \mathcal{G}$ if the rank of G is sufficiently large.

Shalev's conjecture was disproved for groups of type $\mathrm{PSL}_2(q)$ in [39] using trace polynomials, however, it remains plausible that such word maps are surjective once the rank is large enough (as conjectured in [48, Conjecture 4.6]). Remarkably, Lyndon proved this for infinite symmetric groups, see [11]. A weak form of Larsen's conjecture (surjectivity on SU_n (over $\mathbb{C}$) for infinitely many $n \in \mathbb{N}$) was proved by Elkasapy and Thom in [16] for all words $w \in \mathbf{F}_2 \setminus \mathbf{F}_2^{(2)}$.

In this chapter, we prove that metric versions of Shalev's and Larsen's conjectures are true. Let us explain what we mean by this. For $\varepsilon > 0$ and a subset $Y \subseteq X$ of a metric space (X, d) say that Y is ε-dense in X if $d(x, Y) := \inf_{y \in Y} d(x, y) \leq \varepsilon$ for all $x \in X$. Recall from Chapter 0.1(d) that S_n denotes the symmetric group acting on the set $\underline{n} = \{1, \ldots, n\}$ and $\ell_{\mathrm{H}} \colon \mathrm{S}_n \to [0, 1]$ the normalized Hamming length function (see Definition 0.12; the associated metric is denoted by d_{H}). The following metric analog of Shalev's conjecture holds for symmetric groups.

Theorem 3.1. *Let $w \in \mathbf{F}_r$ be a non-trivial word and $\varepsilon > 0$. There exists an integer $N(\varepsilon, w)$ such that $w(\mathrm{S}_n)$ is ε-dense in S_n with respect to the normalized Hamming metric if $n \geq N(\varepsilon, w)$.*

For a classical group $G \leq \mathrm{GL}(V)$ of Lie type with natural module V (see Section 0.1(f)) set $n = n(G) := \dim(V)$. Recall from Definition 0.12 that $\ell_{\mathrm{rk}} \colon G \to [0, 1]$ denotes the normalized rank length function (write d_{rk} for the associated metric). In analogy to Theorem 3.1 we then have the following.

Theorem 3.2. *Let $w \in \mathbf{F}_r$ be a non-trivial word and $\varepsilon > 0$. Let G be one of the groups $\mathrm{GL}_n(q)$, $\mathrm{Sp}_{2m}(q)$, $\mathrm{GO}_{2m+1}(q)$, $\mathrm{GO}_{2m}^{\pm}(q)$ or $\mathrm{GU}_n(q)$ (q a prime power, $n \geq 2$, $m \geq 1$). There exists an integer $N(\varepsilon, w)$ such that $w(G)$ is ε-dense in G with respect to the normalized rank metric if $n = n(G) \geq N(\varepsilon, w)$.*

Recall from Section 0.1(f) that U_n resp. SU_n denote the general resp. special unitary group (over $\mathbb{C}$) of degree $n \in \mathbb{Z}_+$. Equip these groups with the normalized rank length function ℓ_{rk}. We will also prove the following metric version of Larsen's conjecture.

Theorem 3.3. *Let $w \in \mathbf{F}_r$ be a non-trivial word and $\varepsilon > 0$. There exists an integer $N(\varepsilon, w)$ such that $w(\mathrm{U}_n)$ is ε-dense in U_n with respect to the normalized rank metric for all $n \geq N(\varepsilon, w)$.*

First of all, note that in the metric context there is no notable difference between A_n and S_n, $\mathrm{GL}_n(q)$ and $\mathrm{SL}_n(q)$, $\mathrm{GO}_{2m+1}(q)$ resp. $\mathrm{GO}_{2m}^{\pm}(q)$ and $\Omega_{2m+1}(q)$ resp. $\Omega_{2m}^{\pm}(q)$, $\mathrm{GU}_n(q)$ and $\mathrm{SU}_n(q)$, and similarly between U_n and SU_n when n is large, so that we are essentially talking about families of quasisimple compact groups. We conjecture that results analogous to Theorem 3.3 hold for other families of compact Lie groups of increasing rank. Also note that density with respect to the normalized rank metric implies density with respect to the normalized Hilbert–Schmidt metric on U_n – this was also unknown to the best of our knowledge.

Let us now say some words about the proofs of Theorems 3.1, 3.2, and 3.3. First observe that it suffices to prove both results for $r = 2$. Indeed, if $w \in \mathbf{F}_r$ for $r \geq 2$, then via a suitable embedding $\mathbf{F}_r \leq \mathbf{F}_2 = \langle x, y \rangle$ we can view w as a non-trivial word in the two variables, x, y. Hence we shall restrict to the case $w \in \mathbf{F}_2 = \langle x, y \rangle$. We may also assume that w is cyclically reduced, since $w(G)$ is a characteristic subset for any group G, so that up to a change of variables $w = x^a$ or $w = x^{a_1}y^{b_1}\cdots x^{a_l}y^{b_l}$ for some $l \in \mathbb{Z}_+$. In the first case, we call w a *power word*. Note that in this case we can replace $w = x^a$ by any word $v = u^a$, where $u \in \mathbf{F}_2 \setminus (\langle x\rangle \cup \langle y\rangle)$ is arbitrary, so that we are in the second case (since then $v(G) \subseteq w(G)$ for all groups G).

Now let (G, d_G) be one of the groups from Theorem 3.1, 3.2, or 3.3 together with the corresponding metric. Write $n = n(G)$ for its permutation degree resp. the dimension of its natural module. In all three theorems, instead of proving the existence of $N(w, \varepsilon)$, we will prove the equivalent statement (including the corresponding quantitative bounds for $N(w, \varepsilon)$ mentioned below) that there exists a function $d\colon [0, 1] \to \mathbb{R}$ of type $d(x) = Cx^{1/e}$, where $e = e(w) \geq 1$ only depends on w and $C > 0$ depends on the choice of e, such that $d_G(g, w(G)) \leq d(1/n)$ for all $g \in G$. This just means that $N(w, \varepsilon) = O_w((1/\varepsilon)^{e(w)})$, where now the implied constant in the O notation may still depend on w.

Now we give a brief outline of the proofs of Theorems 3.1, 3.2, and 3.3. For the convenience of the reader we will prove Theorem 3.3 before proving Theorem 3.2, as their proofs follow the same idea, but the details of the latter are more involved.

In the proof of Theorem 3.1, at first we consider the case when $w = x^a$ is a power word ($a \notin \{0, \pm 1\}$) to determine the precise quality of the quantities $\sup_{\sigma \in \mathrm{S}_n \text{ isotypic}} d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$ and $\sup_{\sigma \in \mathrm{S}_n} d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$ for $n \to \infty$ (see Lemmas 3.4, 3.5, and 3.6). Lemmas 3.5 and 3.6 show that the above estimate with $d(x) = Cx^{1/2}$ is optimal in this case (up to the constant C) and also demonstrates that the argument using Jensen's inequality, at the end of Subsection 3.2.4, gives the optimal estimate in the power word case (see Remark 3.16). However, as remarked above, one could neglect the power word case if one is not interested in the precise quality of the function d. In the case that w is not a power word, we first settle the case where the permutation σ which we want to approximate by word values is isotypic (see Section 0.1(d)) using the cycle structure of elements of $\mathrm{PSL}_2(q)$ acting on the projective line L_q (see Subsections 3.2.2 and 3.2.3), and then deduce the general case using an application of Jensen's inequality and the fact that a permutation $\sigma \in \mathrm{S}_n$ has less than $\sqrt{2n}$ distinct cycle types (see Subsection 3.2.4). Note that the idea used in the proof of partitioning the set $\underline{n}$ into copies of projective lines L_q and letting copies of groups $\mathrm{PSL}_2(q)$ act on them already appears in [43, Proof of Proposition 8]. In this case, one needs a number theoretic result by Linnik [50] to prove the existence of the constant $e = e(w) \geq 1$. However, the qualitative statement of Theorem 3.1 remains true without this assumption. Assuming a conjecture of Chowla [5] one can show that any $e > 2(l + 1)$ works.

We proceed by giving an overview of the proof of Theorem 3.3. The proof of the results in [16] relied on the analysis of a certain algebraic condition on the abelianized Fox derivative of w and our new strategy is a generalization of this – see Subsection 3.3.3 for details. We use monomial matrices and draw a connection to the normalized dimension of the second cohomology group of finite quotients of the Cayley complex of the one-relator group $K = \mathbf{F}_2/\langle\langle w \rangle\rangle$. As an intermediate step, we consider the largest free nilpotent group H which is a quotient of K, i.e., when $c = c(w) \geq 0$ is determined by $w \in \gamma_{c+1}(\mathbf{F}_2) \setminus \gamma_{c+2}(\mathbf{F}_2)$ (recall that $(\gamma_i(L))_{i\in\mathbb{Z}_+}$ denotes the lower central series of the group L; see Section 0.1(c)), then $H = \mathbf{F}_2/\gamma_{c+1}(\mathbf{F}_2)$. Using Jennings' embedding theorem, for all sufficiently large primes p we find arbitrary large finite p-groups $H(p)$ of composition length $h = h(w)$ equal to the Hirsch length of H, which are quotients of K and where the above normalized dimension gets arbitrarily small. A quantitative analysis then reveals that one can take any number greater than $h(w)$ for the exponent $e = e(w)$. In the worst case, we get that $c \leq 2l$ by a result of Fox [19] and then $h(w) \leq \sum_{k=1}^{c} 2^k < 2^{2l+1}$.

In Subsection 3.3.2, we point out that our method of proof together with a fact on the linearized permutation representation of S_n (see Lemma 3.23) implies as well that $w_1(\mathrm{SU}_n)w_2(\mathrm{SU}_n) = \mathrm{SU}_n$ for non-trivial words $w_1, w_2 \in \mathbf{F}_r$ and large n, providing an alternative proof for Theorem 2.3 of [36]. However, it still remains unclear how to prove surjectivity of single words w in general.

Finally, in Section 3.4, we use the same cohomological method as in Section 3.3, but with coefficient groups $(k[X]/(\chi))^\times$ for $\chi \in k[X]$ a polynomial instead of U_1 and a modified version of Lemma 3.21 (namely Corollary 3.26) to settle Theorem 3.2. We remark here that our proof for $\mathrm{GL}_n(k)$ works for all fields k (not only for finite ones) and we conjecture that the same is true for the other Lie types.

After finishing a first version of this chapter, we noted that there is an alternative root to the proof of Theorem 3.1, which uses the ideas from Section 3.3 and 3.4, but with finite cyclic groups C_k instead of coefficients in U_1. We present this in Subsection 3.4.3.

The statement that w has dense word image on a suitable class of normed groups (as in Theorems 3.1, 3.2, and 3.3) is equivalent to surjectivity of w on metric ultraproducts of those normed groups, where $n(G) \to \infty$ along the ultrafilter. Forming such a metric ultraproduct of symmetric groups leads to a so-called universal sofic group, whereas for complex unitary groups we obtain a group which surjects continuously on a universal hyperlinear group.

Let us end by drawing some connections to related questions and articles. To prove that the cardinality of the word image $w(G)$ for G a quasisimple group is large (which was done in [43, Theorem 2] and [44, Theorems 1.9 and 1.11]), Shalev and Larsen approximate an element which has a logarithmically large conjugacy class (e.g., in S_n an n-cycle and in a classical group of Lie type an element admitting a cyclic vector) and exploit that the cardinality of conjugacy classes is continuous in the normalized Hamming resp. rank metric, i.e., $|g^G|/|h^G| \leq |(gh^{-1})^G|$ which is bounded by $|G|^{Ld_{\mathrm{H}}(g,h)}$ resp. $|G|^{Ld_{\mathrm{rk}}(g,h)}$ (for some constant $L > 0$; see Fact 0.13; this is used, e.g., in [49]; see also [67, Corollary 2.14 and Theorem 2.15]). Hence metric density also implies that $\log_{|G|}|w(G)| \to 1$ when $n(G) \to \infty$. In private communication with Shalev, he conjectured, because of the above connection, that the word image $w(\mathrm{S}_n)$ is actually C/n-dense for a fixed constant $C > 0$, as indicated by Theorem 1.9 of [44] stating that $|w(\mathrm{S}_n)| \geq n^{-4-\varepsilon}n!$ for n sufficiently large. To prove the latter fact, it is enough to find one conjugacy class in the image $w(\mathrm{S}_n)$ which comes from an element being C-close to an n-cycle, as such a class has a centralizer of order polynomially bounded in n. However, in [42] it is shown that for power words the even better estimate $|w(\mathrm{S}_n)| \geq Cn^{-1}n!$ holds, but Lemma 3.6 of Subsection 3.2.1 demonstrates that though the word image misses some large ε-balls, i.e., $\varepsilon = \Omega(1/\sqrt{n})$. Hence the word image can be distributed quite non-uniformly in the metric sense.

Similarly, using results of [49] or [12], if one can approximate a conjugacy class of large cardinality resp. norm, one gets bounded width of the word image. In any case, maybe both questions about cardinality and width of the word image have the same simple answer, namely that w is eventually surjective (when w is not a proper power in the case of finite simple groups).

The rest of this chapter is structured as follows. In Section 3.2, we give the proof of Theorem 3.1, Section 3.3 presents the proof of Theorem 3.3, and in Section 3.4, we prove Theorem 3.2.

3.2 Symmetric groups

This section is devoted to the proof of Theorem 3.1. Recall the notation from Section 0.1(d) and Section 0.2. At first we consider the case that $w = x^a$ is a power word separately to determine the quantity $\sup_{\sigma \in \mathrm{S}_n} d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$ as a function of $1/n$ for large n up to a factor $2\sqrt{2}$ (see Lemmas 3.5 and 3.6). We also consider the quantity $\sup_{\sigma \in \mathrm{S}_n \text{ isotypic}} d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$ (see Lemma 3.4 below). However, as explained in the introduction of this chapter, this case is somehow superfluous by the argument given there. Hence, if the reader is not interested in these details, we propose to skip this section.

3.2.1 Power words

We can certainly assume that $a \geq 2$, replacing $a < 0$ by $-a$ and ignoring the trivial case $w = x$. We need some number theoretic notation for this subsection. For $x, y \in \mathbb{Z}_+$ denote by $\mathrm{rad}(x)$ the *radical* of x, which is the product of all primes dividing x, and by the *y-part* $\pi_y(x)$ of x the biggest integer $z \in \mathbb{Z}_+$ such that $\mathrm{rad}(z) = \mathrm{rad}(\gcd\{x, y\})$ and $z \mid x$.

The isotypic case. Assume first that the permutation $\sigma \in \mathrm{S}_n$ which we want to approximate by w-values is isotypic. Then we have the following.

Lemma 3.4. *Let $\sigma \in \mathrm{S}_n$ be a k-isotypic permutation and set $c_k := c_k(\sigma)$. Then we have $d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq a/n$ with equality if and only if* $\mathrm{rad}(a) \mid k$ *and* $c_k - a\lfloor c_k/a \rfloor = a - 1$.

Proof. Let $\tau \in \mathrm{S}_n$ be such that $d_{\mathrm{H}}(\sigma, \tau^a)$ is minimal. Set $a' := \pi_k(a)$. Note that an l-cycle of τ gets transformed into $\gcd\{l, a\}$ cycles of length $l/\gcd\{l, a\}$ of τ^a, so $c_k(\tau^a)$ is a multiple of a'. Hence, subsequently, we can certainly restrict to the case $a' \nmid c_k$, since otherwise we can take τ of cycle type $((ka')^{c_k/a'})$, so that $\tau^a = \sigma$ (so in particular $k, a', \gcd\{k, a\} \neq 1$).

Assume first that $\mathrm{rad}(a) \mid k$, i.e., $a = a'$. Then $c_k(\tau^a)$ is a multiple of a. Let c'_k be the number of k-cycles in σ which are not cycles in τ^a. Call these k-cycles *bad*. Then $c'_k \geq c_k - a\lfloor c_k/a \rfloor \geq 1$ and the middle term is at most $a-1$. Hence σ and τ^a differ in at least c'_k points (at least one point for each bad k-cycle of σ). However, if $d_{\mathrm{H}}(\sigma, \tau^a) = c'_k/n = (c_k - a\lfloor c_k/a \rfloor)/n$, then each bad k-cycle of σ would have support strictly contained in the support of a cycle of τ^a (by optimality of τ the supports cannot be equal). But this is impossible, since then τ^a would have a cycle of length bk for some $b \geq 2$, so $c_{bk}(\tau^a) \geq a$ and our set $\underline{n}$ would be to small. So σ and τ^a differ in at least $c_k - a\lfloor c_k/a \rfloor + 1 \leq a$ points. This bound is also attained by choosing τ of cycle type $(1^1, (k(c_k - a\lfloor c_k/a \rfloor) - 1)^1, (ak)^{\lfloor c_k/a \rfloor})$ (so that τ^a is of cycle type $(1^1, (k(c_k - a\lfloor c_k/a \rfloor) - 1)^1, k^{a\lfloor c_k/a \rfloor})$).

Now assume that $\mathrm{rad}(a) \nmid k$, i.e., $a' < a$. Then $c_k(\tau^a)$ is a multiple of a'. If $\mathrm{rad}(a) \mid k(c_k - a'\lfloor c_k/a' \rfloor)$, similarly to the above, we can take τ of cycle type $(1^1, (k(c_k - a'\lfloor c_k/a' \rfloor) - 1)^1, (ka')^{\lfloor c_k/a' \rfloor})$, so that τ^a is of cycle type $(1^1, (k(c_k - a'\lfloor c_k/a' \rfloor) - 1)^1, k^{a'\lfloor c_k/a' \rfloor})$, and $d_{\mathrm{H}}(\sigma, \tau^a) = c_k - a'\lfloor c_k/a' \rfloor + 1 \leq a' < a$. In the opposite case, $\mathrm{rad}(a) \nmid k(c_k - a'\lfloor c_k/a' \rfloor)$ write $a = a_1 a_2$, where $a_1 < a$ is the $k(c_k - a'\lfloor c_k/a' \rfloor)$-part of a. Then in the arithmetic progression $k(c_k - a'\lfloor c_k/a' \rfloor) - 1, k(c_k - a'\lfloor c_k/a' \rfloor) - 1 - a_1, \ldots, k(c_k - a'\lfloor c_k/a' \rfloor) - 1 - a_1(a_2 - 1)$ there is a unit u modulo a_2 (which is of course also a unit modulo a_1 and so modulo a). If the last number is the only such, we must have $a_2 = 2$. In this case, we take τ' of cycle type $(1^2, (k(c_k - a'\lfloor c_k/a' \rfloor) - 2)^1, (ka')^{\lfloor c_k/a' \rfloor})$, so that τ'^a is of cycle type $(1^2, (k(c_k - a'\lfloor c_k/a' \rfloor) - 2)^1, k^{a'\lfloor c_k/a' \rfloor})$ and $d_{\mathrm{H}}(\sigma, \tau^a) \leq d_{\mathrm{H}}(\sigma, \tau'^a) = c_k - a'\lfloor c_k/a' \rfloor + 2 \leq a' + 1 < a$ (as $a' \neq 1$). In the opposite case, we can choose $u = k(c_k - a'\lfloor c_k/a' \rfloor) - 1 - ia_1$ with $i \leq a_2 - 2$. Set $t \coloneqq 1 + ia_1$. If $u < 0$, one can choose τ' of cycle type $(1^{k(c_k - a'\lfloor c_k/a' \rfloor)}, (ka')^{\lfloor c_k/a' \rfloor})$, so that τ'^a is of cycle type $(1^{k(c_k - a'\lfloor c_k/a' \rfloor)}, k^{a'\lfloor c_k/a' \rfloor})$ and

$$d_{\mathrm{H}}(\sigma, \tau^a) \leq d_{\mathrm{H}}(\sigma, \tau'^a) = \frac{k(c_k - a'\lfloor c_k/a' \rfloor)}{n} < \frac{1 + a_1(a_2 - 2)}{n} < a/n.$$

In the opposite case, we take τ' of cycle type $(1^t, u^1, (ka')^{\lfloor c_k/a' \rfloor})$, so that τ'^a is of cycle type $(1^t, u^1, k^{a'\lfloor c_k/a' \rfloor})$ and $d_{\mathrm{H}}(\sigma, \tau^a)$ is bounded by

$$d_{\mathrm{H}}(\sigma, \tau'^a) \leq \frac{c_k - a'\lfloor c_k/a' \rfloor + t}{n} \leq \frac{(a'-1) + (1 + a_1(a_2 - 2))}{n} < a/n,$$

where the second inequality is strict when $t \geq k$, and the last inequality holds since $a' \leq a_1$. To see the first inequality, first take a big cycle of length $c_k - a'\lfloor c_k/a' \rfloor$ with support on the bad cycles of σ such that on each of these there is exactly one point where this cycle does not agree with σ. Then, adding t consecutive fixed points on this cycle, we produce at most t more errors. $\square$

The general case. Now we consider the case that the permutation $\sigma \in \mathrm{S}_n$ which we want to approximate by w-values is arbitrary. At first we estimate the quantity $d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$ from above.

Lemma 3.5. *Set* $D \coloneqq \frac{1}{\mathrm{rad}(a)} \sum_{k=1}^{\mathrm{rad}(a)} (\pi_k(a) - 1) = \frac{1}{a} \sum_{k=1}^{a} (\pi_k(a) - 1)$*. For any* $\varepsilon > 0$

and $n \in \mathbb{Z}_+$ sufficiently large (in terms of ε) we have that

$$d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq (1+\varepsilon)\sqrt{\frac{2D}{n}}.$$

for all $\sigma \in \mathrm{S}_n$.

Proof. Let the permutation σ which we want to approximate by word values be of cycle type $(k^{c_k})_{k\in\mathbb{Z}_+}$. At first we justify that we may assume $c_k \leq \pi_k(a)-1$ ($k \in \mathbb{Z}_+$). Note that the function $d(x) = (1+\varepsilon)\sqrt{2Dx}$ obeys the following properties:

(i) For all $N \in \mathbb{Z}_+$ it holds that $d(x) \geq Nx$ for $x > 0$ small enough;

(ii) d is concave;

(iii) $d(0) = 0$.

For a fixed $N \in \mathbb{Z}_+$ assume that $d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq d(1/n)$ for all $n \geq N$ and $\sigma \in \mathrm{S}_n$ with $c_k(\sigma) \leq \pi_k(a)-1$ for all $k \in \mathbb{Z}_+$. For $\sigma \in \mathrm{S}_n$ arbitrary write $\sigma = \sigma_1\sigma_2$ such that the supports $\Omega_1 \coloneqq \mathrm{supp}(\sigma_1)$ and $\Omega_2 \coloneqq \mathrm{supp}(\sigma_2)$ of σ_1 and σ_2 are disjoint and $c_k(\sigma_1) = c_k - \pi_k(a)\lfloor c_k/\pi_k(a)\rfloor \leq \pi_k(a)-1$ (so $\pi_k(a) \mid c_k(\sigma_2) = \pi_k(a)\lfloor c_k/\pi_k(a)\rfloor$). Then $d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq \frac{n_1}{n} d_{\mathrm{H}}(\sigma_1, \mathrm{Sym}(\Omega_1))$, where $n_1 \coloneqq |\Omega_1|$, since $\sigma_2 \in w(\mathrm{Sym}(\Omega_2))$. Now if $n_1 < N$, for n sufficiently large we have that

$$d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq n_1/n < N/n \leq d(1/n)$$

by Property (i). In the opposite case, if $n_1 \geq N$, Properties (i) and (ii) and Jensen's inequality imply that $d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq \frac{n_1}{n} d_{\mathrm{H}}(\sigma_1, w(\mathrm{Sym}(\Omega_1))) \leq \frac{n_1}{n} d(1/n_1) \leq d(1/n)$.

So, subsequently, fix $\sigma \in \mathrm{S}_n$ with $c_k \leq \pi_k(a)-1$ ($k \in \mathbb{Z}_+$). Let u be the biggest unit modulo a with $u \leq n$. Then set $t \coloneqq n-u \leq a-1$. Choosing $\tau \in \mathrm{S}_n$ of cycle type $(1^t, u^1)$ so that τ^a has the same cycle type, we see that

$$d_{\mathrm{H}}(\sigma, \tau^a) \leq \frac{t+1+\sum_{k\in\mathbb{Z}_+} c_k}{n} \leq \frac{a+\sum_{k\in\mathbb{Z}_+} c_k}{n} =: \frac{e}{n},$$

since we make at most one mistake on each cycle of σ (if τ^a would be one big cycle) plus the modification of $t+1$ points. This estimate is worst possible if the number of cycles of σ is maximal. Hence, assuming n is large and modifying it slightly if necessary, we may assume $c_k = \pi_k(a)-1$ for all $k \leq l\,\mathrm{rad}(a)$ and $c_k = 0$ otherwise for some large $l \in \mathbb{Z}_+$. In this situation we have

$$\begin{aligned}
n &= \sum_{k=1}^{l\,\mathrm{rad}(a)} k(\pi_k(a)-1) \\
&= l\sum_{i=1}^{\mathrm{rad}(a)} i(\pi_i(a)-1) + \binom{l}{2}\mathrm{rad}(a)\sum_{j=1}^{\mathrm{rad}(a)}(\pi_j(a)-1) \\
&\sim \frac{D\,\mathrm{rad}(a)^2}{2}l^2.
\end{aligned}$$

Here we use the fact that $\pi_k(a)$ depends only on the congruence class of k modulo $\mathrm{rad}(a)$. Similarly,

$$e = a + l\sum_{i=1}^{\mathrm{rad}(a)}(\pi_i(a)-1) \sim D\,\mathrm{rad}(a)l,$$

so that

$$\frac{e}{n} \sim \frac{2}{\mathrm{rad}(a)l} \sim \sqrt{\frac{2D}{n}}.$$

The proof is now complete. □

In the next lemma, we establish that the constant D in Lemma 3.5 is optimal up to factor $2\sqrt{2}$.

Lemma 3.6. *With D defined as in Lemma 3.5, for any $\varepsilon > 0$ there are infinitely many $n \in \mathbb{Z}_+$ and $\sigma \in \mathrm{S}_n$ such that*

$$d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \geq \frac{1-\varepsilon}{2}\sqrt{\frac{D}{n}}.$$

Proof. Set $E := \frac{1}{\mathrm{rad}(a)}\sum_{k=1}^{\mathrm{rad}(a)}\lfloor \pi_k(a)/2\rfloor = \frac{1}{a}\sum_{k=1}^{a}\lfloor\pi_k(a)/2\rfloor$. Fix $l \in \mathbb{Z}_+$ and choose $\sigma \in \mathrm{S}_n$ such that $c_k(\sigma) = \lfloor \pi_k(a)/2\rfloor$ for $k \leq l\,\mathrm{rad}(a)$ and $c_k(\sigma) = 0$ otherwise. Then we have

$$\begin{aligned}
n &= \sum_{k=1}^{l\,\mathrm{rad}(a)} k\lfloor\pi_k(a)/2\rfloor \\
&= l\sum_{i=1}^{\mathrm{rad}(a)} i\lfloor\pi_i(a)/2\rfloor + \binom{l}{2}\mathrm{rad}(a)\sum_{j=1}^{\mathrm{rad}(a)}\lfloor\pi_j(a)/2\rfloor \\
&\sim \frac{E\,\mathrm{rad}(a)^2}{2}l^2,
\end{aligned}$$

using the same trick as in the proof of Lemma 3.5.

Next we establish a lower bound for $e := nd_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$. Fix $\tau \in \mathrm{S}_n$ such that $d_{\mathrm{H}}(\sigma,\tau^a)$ is minimal. Recall that a cycle of σ is called *good* if it is a cycle of τ^a and *bad* otherwise. Similarly, we call a point $x \in \underline{n}$ good if $x.\sigma = x.\tau^a$ and bad otherwise. Now fix $k \in \mathbb{Z}_+$. If there are exactly i good k-cycles of σ (so they are k-cycles of τ^a as well), then the other

$c_k(\tau^a) - i$ k-cycles of τ^a contain at least i bad points. Indeed, either $c_k(\tau^a) = 0$ and there is nothing to show, or $c_k(\tau^a) \geq \pi_k(a)$ and $i \leq c_k(\sigma) = \lfloor \pi_k(a)/2 \rfloor$ by the choice of σ, which implies $c_k(\tau^a) - i \geq \pi_k(a) - \lfloor \pi_k(a)/2 \rfloor \geq \lfloor \pi_k(a)/2 \rfloor \geq i$.

Hence, indexing the cycles of σ with some set I and writing b_i for the number of bad points in the ith cycle, we obtain from the previous argument that $|\{i \in I \mid b_i = 0\}| \leq \sum_{i\in I} b_i = e$. This implies that $e \geq |I|/2$, since $|\{i \in I \mid b_i = 0\}| \leq \sum_{i\in I} b_i < |I|/2$ leads to the contradiction $\sum_{i\in I} b_i \geq |\{i \in I \mid b_i \neq 0\}| = |I| - |\{i \in I \mid b_i = 0\}| > |I|/2$.

Therefore

$$e \geq \frac{|I|}{2} = \frac{1}{2} \sum_{k=1}^{l\,\mathrm{rad}(a)} \lfloor \pi_k(a)/2 \rfloor = \frac{l}{2} \sum_{i=1}^{\mathrm{rad}(a)} \lfloor \pi_k(a)/2 \rfloor = \frac{E\,\mathrm{rad}(a)}{2} l$$

and we obtain

$$\frac{e}{n} \gtrsim \frac{1}{l\,\mathrm{rad}(a)} \sim \sqrt{\frac{E}{2n}}.$$

The claim follows from the fact that $D/2 \leq E (\leq D)$. □

This ends this subsection. In the following, we assume that w is not a power word. To attack this case, we start by collecting some basic facts about the groups $\mathrm{PSL}_2(q)$ for q a prime power.

3.2.2 The cycle structure of elements from $\mathrm{PSL}_2(q)$

In this subsection, we recall some well-known facts about the cycle structure of the elements from $\mathrm{PSL}_2(q) \leq \mathrm{Sym}(L_q)$ acting on the projective line L_q of order q, where $q = p^e$ is a power of the prime p. The key observation, which we will exploit in Subsection 3.2.4 to prove Theorem 3.1, is here that these elements are all almost isotypic.

Consider an element $g \in \mathrm{SL}_2(q)$ and write $\overline{g} \in \mathrm{PSL}_2(q)$ for the corresponding permutation on L_q. Then g has two eigenvalues $\lambda, \lambda^{-1} \in \mathbb{F}_{q^2}$. Recall from Section 0.1 that $o := \mathrm{ord}(\lambda)$ denotes the multiplicative order of λ. We have the following complete distinction into three disjoint cases.

Case 1: If $\lambda = \pm 1$, then g has at least one eigenvector. If it has a second eigenvector not contained in the span of the first one, we have $g = \pm\,\mathrm{id}$, so $\overline{g} = \mathrm{id}_{L_q}$. If this is not the case, in a suitable basis

$$g = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

so $\overline{g}$ has precisely one fixed point $[1:0]$ and the remaining q/p cycles are all of length p.

Case 2: In the case when $\lambda \in \mathbb{F}_q \setminus \{\pm 1\}$, we see that g is diagonalizable over $\mathbb{F}_q$, whence $o = \mathrm{ord}(g)$ divides $q - 1$. So choose coordinates such that $g = \mathrm{diag}(\lambda, \lambda^{-1})$. Then $\overline{g}$ has the two fixed points $[1:0], [0:1]$ on L_q corresponding to the eigenvectors of g over $\mathbb{F}_q$. Take any other point $x = [a:b] \in L_q$ (i.e., $ab \neq 0$). Then the orbit of x under $\langle \overline{g} \rangle \leq \mathrm{PSL}_2(q)$ has length $k = o/2$ resp. $k = o$ when o is even respectively odd. Namely, $x.\overline{g}^l = x$ is equivalent to $[\lambda^l a : \lambda^{-l} b] = [\lambda^{2l} a : b] = [a : b]$, which is equivalent to $\lambda^{2l} = 1$, meaning that $o/2 \mid l$ for o even and $o \mid l$ for o odd.

Case 3: In the last case, $\lambda \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$. Since $(X - \lambda)(X - \lambda^{-1}) = \chi_g(X) \in \mathbb{F}_q[X]$, it follows that $\lambda^{-1} = \lambda^q$ is the Galois conjugate to λ in $\mathbb{F}_{q^2}$, so $o = \mathrm{ord}(g)$ divides $q + 1$. Moreover, $\overline{g}$ has no fixed points as g has no eigenvector over $\mathbb{F}_q$. However, embedding into $\mathrm{SL}_2(q^2)$, we can again assume $g = \mathrm{diag}(\lambda, \lambda^{-1})$. Then the same argument as above shows that all cycles of $\overline{g}$ have length $k = o/2$ resp. $k = o$ when o is even resp. odd.

To summarize our observations, let us state the following corollary.

Corollary 3.7. *The cycle type of $\overline{g}$ acting on L_q only depends on $o = \mathrm{ord}(\lambda)$ and q if $o > 2$. Namely, if $2 < o$ and $o \mid q - 1$ (Case 2), then it is $(1^2, (o/2)^{2(q-1)/o})$ resp. $(1^2, o^{(q-1)/o})$ when o is even resp. odd, and if $2 < o$ and $o \mid q + 1$ (Case 3) it is $((o/2)^{2(q+1)/o})$ resp. $(o^{(q+1)/o})$ when o is even resp. odd.*

3.2.3 Effective surjectivity of word maps over finite fields

In this subsection, using the facts from Subsection 3.2.2, we demonstrate that permutations of certain cycle type are attained as w-values inside groups of type $\mathrm{PSL}_2(q)$, thus providing the crucial ingredient for the proof of Theorem 3.1 in Subsection 3.2.4.

As in the previous subsection, let $q = p^e$ be a power of the prime p. The map $\mathrm{tr}_w \colon \mathrm{SL}_2(\overline{\mathbb{F}}_q)^2 \to \overline{\mathbb{F}}_q$ defined by $(g, h) \mapsto \mathrm{tr}(w(g, h))$ is surjective. Indeed, this can be seen from the existence of trace polynomials and the theorem of Borel [4] that the word map associated to w on $\mathrm{SL}_2(\overline{\mathbb{F}}_q)$ is dominant – but it follows also from direct inspection as explained below. Here we show surjectivity of tr_w for p large enough but in an effective way. Throughout this subsection, assume that w is not a power word and $p \nmid a_i, b_i$ for $i = 1, \ldots, l$, where $w = x^{a_1} y^{b_1} \cdots x^{a_l} y^{b_l}$ as in the introduction.

Lemma 3.8. *For any $t \in \mathbb{F}_q$ there exists $m \leq l$ and unipotent elements $g, h \in \mathrm{SL}_2(q^m)$ such that $\mathrm{tr}_w(g, h) = t$.*

Proof. A classical result (going back to [20]) says that

$$\mathrm{tr}_w(g, h) = f(\mathrm{tr}(g), \mathrm{tr}(h), \mathrm{tr}(gh))$$

is a polynomial in $\mathrm{tr}(g)$, $\mathrm{tr}(h)$, and $\mathrm{tr}(gh)$, where

$$f(X, Y, Z) = f_l(X, Y) Z^l + \cdots + f_0(X, Y) \in \mathbb{Z}[X, Y, Z]$$

is uniquely determined and called the *trace polynomial* of w.

Now we define $g(U, V), h(U, V) \in \mathbb{Z}[U, V]$ by

$$g(U, V) := \begin{pmatrix} 1 & 0 \\ U & 1 \end{pmatrix} \text{ and } h(U, V) := \begin{pmatrix} 1 & V \\ 0 & 1 \end{pmatrix}.$$

Then $\mathrm{tr}(g(U,V)) = \mathrm{tr}(h(U,V)) = 2$ and $\mathrm{tr}(gh) = UV + 2$ in $\mathbb{Z}[U,V]$. Computing $f_n(2,2)$ from above gives $f_n(2,2) = a_1b_1 \cdots a_lb_l$, in particular, f is non-trivial.

Now substituting 1 for V and reducing modulo p gives a polynomial

$$r(U) \coloneqq \mathrm{tr}_w(g(U,1), h(U,1)) = a_1b_1 \dots a_lb_lU^l + s(U) \in \mathbb{F}_p[U],$$

with $\deg(s) < l$, of degree l by assumption on w, as $p \nmid a_1b_1 \cdots a_lb_l$. Hence the equation $r(U) - t = 0$ is an equation over $\mathbb{F}_q$ of degree l, so has a solution in one of the fields $\mathbb{F}_{q^i}$ for $i = 1, \dots, l$. □

Remark 3.9. If l is odd, then m can also be chosen odd, since then at least one irreducible factor of $r(U)$ must be of odd degree.

As a consequence of Lemma 3.8 together with the facts mentioned in Subsection 3.2.2, for any fixed integer $k > 1$ we get a word value in some groups of the form $\mathrm{PSL}_2(q^{im})$ for all $i \in \mathbb{Z}_+$, where q depends on k, consisting only of k-cycles up to two fixed points. We conclude the following corollary.

Corollary 3.10. *Let $k > 1$ be an integer. Assume that $2k \mid q-1$ resp. $k \mid q-1$ when k is even resp. odd. Then there exists $m \leq l$ such that there is an element $\sigma \in w(\mathrm{PSL}_2(q^{im})) \subseteq \mathrm{S}_{q^{im}+1}$ of cycle type $(1^2, k^{(q^{im}-1)/k})$ for all $i \in \mathbb{Z}_+$.*

Proof. Let $i \in \mathbb{Z}_+$ be arbitrary. Choose $\lambda \in \mathbb{F}_q^\times$ of order $2k$ resp. k when k is even resp. odd. Then apply Lemma 3.8 to $t \coloneqq \lambda + \lambda^{-1} \in \mathbb{F}_q$ to get $g, h \in \mathrm{SL}_2(q^m) \leq \mathrm{SL}_2(q^{im})$ for some $m \leq l$ with $\mathrm{tr}_w(g,h) = \mathrm{tr}(w(g,h)) = t$. Note that $w(g,h) \in \mathrm{SL}_2(q^m) \leq \mathrm{SL}_2(q^{im})$ is diagonalizable with eigenvalues $\lambda, \lambda^{-1} \in \mathbb{F}_q^\times \subseteq \mathbb{F}_{q^{im}}^\times$. Setting $\sigma \coloneqq \overline{w(g,h)} \in \mathrm{PSL}_2(q^{im}) \leq \mathrm{S}_{q^{im}+1}$, Corollary 3.7 of Subsection 3.2.2 immediately implies the claim. □

Remark 3.11. If l is odd, using Remark 3.9, one can even remove the two fixed points from the above element σ. Indeed, assuming $2k \mid q+1$ resp. $k \mid q+1$ when k is even resp. odd and going through the proof of Corollary 3.10 together with the fact that m can be chosen odd, one gets $\sigma \in \mathrm{S}_{q^{im}+1}$ of cycle type $(k^{(q^{im}+1)/k})$ for all odd $i \in \mathbb{Z}_+$.

The next result shows that there is also a word value in $\mathrm{PSL}_2(q) \leq \mathrm{Sym}(L_q) \cong \mathrm{S}_{q+1}$, which is close to a $(q+1)$-cycle in S_{q+1}. It can be considered as a weak version of [44, Theorem 4.1] which permits an elementary proof.

Lemma 3.12. *Assume that $q > 4l$. Then there exists $\sigma \in w(\mathrm{PSL}_2(q)) \subseteq \mathrm{Sym}(L_q) \cong \mathrm{S}_{q+1}$, which differs in less than $2 + \sqrt{lq}$ points of L_q from a $(q+1)$-cycle in S_{q+1}.*

Proof. Using the same trick as in the proof of Lemma 3.8, one sees that the map

$$\mathrm{tr}_w \colon \mathrm{SL}_2(q) \times \mathrm{SL}_2(q) \to \mathbb{F}_q$$

meets at least q/l points. This implies that the set $\Lambda \subseteq \mathbb{F}_{q^2}^\times$ of eigenvalues of elements from $w(\mathrm{SL}_2(q))$ has cardinality at least $2(q/l - 1)$ (two eigenvalues for each trace value apart from the trace values ± 2; if $2 \mid q$ one can take $2q/l-1$). Now assume that the multiplicative order of all these eigenvalues is less than $b := 2\sqrt{q/l}$. Then we obtain the contradiction

$$|\Lambda| \leq \sum_{i=1}^{\lceil b\rceil -1} \varphi(i) \leq \frac{(\lceil b\rceil - 1)^2}{2} - 2 < 2(q/l-1),$$

where the second inequality holds since by assumption $\lceil b\rceil - 1 \geq 4$. Hence Λ contains an element λ of order $o \geq 2\sqrt{q/l} > 4$. Let $f \in w(\mathrm{SL}_2(q))$ with eigenvalues $\lambda, \lambda^{-1} \in \mathbb{F}_{q^2}$. Then by Corollary 3.7 the permutation $\sigma := \overline{f} \in \mathrm{PSL}_2(q)$ consists apart from zero or two fixed points only of cycles of type $o/2$ resp. o when o is even resp. odd. This implies that σ differs in less than $2 + \sqrt{lq}$ points from a $(q+1)$-cycle in $\mathrm{Sym}(L_q) \cong \mathrm{S}_{q+1}$. □

Remark 3.13. For even q one can improve the estimate by a factor $1/2$, since o will always be odd.

3.2.4 Proof of Theorem 3.1

In this subsection, we use the facts provided by Subsection 3.2.3 to establish Theorem 3.1. We may assume that $w = x^{a_1}y^{b_1}\cdots x^{a_l}y^{b_l}$ as in Subsection 3.2.3, as the case that w is a power word was already settled by Lemma 3.5. We start with the isotypic case and prove the general case as a consequence.

The isotypic case. At first let $\sigma \in \mathrm{S}_n$ be k-isotypic, i.e., $n = c_k k$ for $c_k := c_k(\sigma)$. We can certainly restrict to $k > 1$, since the identity is always in $w(\mathrm{S}_n)$. Subsequently, we prove two estimates for the quantity $d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n))$. The first estimate will be suitable for small k, whereas the second will be better for large k.

Estimate for small k. Let p be the smallest prime such that $p \nmid a_i, b_i$ for $i = 1,\ldots,l$ and $2k \mid p-1$ resp. $k \mid p-1$ when k is even resp. odd. Apply Corollary 3.10 to $q := p$ to get the integer $m \leq l$. Set $q := p^m$ and write

$$n = \sum_{i=1}^{s} n_i(q^i+1) + n_0$$

such that $\sum_{i=1}^{j} n_i(q^i+1) + n_0 \leq q^{j+1}$ for all $0 \leq j \leq s$ (i.e., use a greedy algorithm to produce such a representation, starting with the biggest summand q^s+1).

Then $n_i \leq q-1$ for $i \geq 1$ and $n_0 \leq q$. Moreover, using a standard estimate for the q-ary representation of positive integers, one obtains $\sum_{i=0}^{s} n_i < q(\log_q(n)+1)$.

Write $\underline{n} = \bigsqcup_{i=1}^{s} n_i L_{q^i} \sqcup \underline{n_0}$ as a disjoint union of n_i copies of the projective lines L_{q^i} for $i = 1,\ldots,s$ and n_0 singletons. Using Corollary 3.10, let $g, h \in \mathrm{S}_n$ be permutations which restrict to maps $\overline{g}, \overline{h} \in \mathrm{PSL}_2(q^i)$ acting on the copies of L_{q^i} such that $w(\overline{g}, \overline{h})$ has cycle type $(1^2, k^{(q^i-1)/k})$, and which fix the remaining n_0 points. Then, if we label the points in an optimal way, we get

$$nd_{\mathrm{H}}(\sigma, w(g,h)) = n_0 + 2\sum_{i=1}^{m} n_i \leq 2\sum_{i=0}^{m} n_i \leq 2q(\log_q(n)+1).$$

By a celebrated result of Linnik [50], one has that the least prime which is congruent to 1 modulo $2k$ resp. k is bounded by $D_1 k^{D_2}$ for some constants $D_1 > 0$, $D_2 \geq 1$. Choosing D_1 large enough, one can also ensure that $p \nmid a_i, b_i$ for $i = 1, \ldots, l$, e.g., take p congruent to 1 modulo $2ka_1b_1 \cdots a_lb_l$. Hence $q \leq D_1^l k^{D_2 l}$, so that

$$d_{\mathrm{H}}(\sigma, w(g,h)) \leq 2D_1^l k^{D_2 l}(\log_2(n)+1)/n.$$

Remark 3.14. The logarithmic term in this argument can be removed if l is odd. Namely, then we require $2k \mid p+1$ resp. $k \mid p+1$ for k even resp. odd, and can choose $m \leq l$ odd, so that we are in Case 3 of Subsection 3.2.2, where no fixed points occur. However, then we may only use the odd i and thus get a bigger constant.

It is probably also true that, when w is not a square, $\sigma \in w(\mathrm{S}_n)$ if k is fixed and c_k is even and large enough in terms of k (Lemma 3.42(ii) of Subsection 3.4.3 can be seen as a weak form of this conjecture which is true). But this also would not improve our estimate.

The result of Linnik is not necessary for the qualitative statement of Theorem 3.1. We only need it to get a nice function d, which is mentioned in the introduction. There are weaker versions of Linnik's result available with an elementary proof, e.g., see [68].

Estimate for large k. Let p be the smallest prime such that $p \nmid a_i, b_i$ for $i = 1, \ldots, l$ and $p > 4l$. Set $q := p$ and write $n = \sum_{i=1}^{s} n_i(q^i+1) + n_0$ and $\underline{n} = \bigsqcup_{i=1}^{s} n_i L_{q^i} \sqcup \underline{n_0}$ as above. Using Lemma 3.12, let $g, h \in \mathrm{S}_n = \mathrm{Sym}(\underline{n})$ be permutations which restrict to maps $\overline{g}, \overline{h} \in \mathrm{PSL}_2(q^i)$ acting on the copies of L_{q^i} such that $w(\overline{g}, \overline{h})$ differs in less than $2 + \sqrt{q^i l}$ points from an (q^i+1)-cycle for $i = 1, \ldots, s$, and which fix the remaining n_0 points.

Again, under an optimal labeling, an n-cycle differs from σ in at most c_k points. Hence, using the triangle inequality,

$$d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq \frac{1}{k} + \frac{1}{n}\left(\sum_{i=1}^{s} n_i(\sqrt{lq^i}+2) + n_0\right).$$

The second term can be estimated by $D_3/\sqrt{n}$ for suitable $D_3 > 0$ depending on q and l.

Remark 3.15. By Theorem 1.3 of [44] there exists a constant $D_4 > 0$ such that there are elements $g, h \in \mathrm{S}_n$ which restrict to permutations on the support of each k-cycle of σ such that $d_{\mathrm{H}}(\sigma, w(g,h)) \leq D_4/k$. However, the proof presented there uses some algebraic geometry and the weak Goldbach conjecture, and using it instead of the above estimate would not improve the exponent e mentioned in the introduction. Note here that we found an alternative proof of this result of Larsen and Shalev after having finished a first version of this chapter, which is presented in Lemma 3.42(iii) of Subsection 3.4.3.

Global estimate for isotypic elements. Using the first estimate for

$$k \leq \left(\frac{n}{2D_1^l(\log_2(n)+1)}\right)^{\frac{1}{D_2 l+1}}$$

and the second one in the opposite case, one obtains that

$$d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) \leq C_{\mathrm{it}} n^{-1/e_{\mathrm{it}}}$$

for any $e_{\mathrm{it}} > D_2 l + 1 \geq 2$ and C_{it} appropriately. Assuming a conjecture by Chowla [5], we can take D_2 arbitrarily close to one, so that e_{it} can be taken arbitrarily close to $l+1$.

We will now use our knowledge about the isotypic case to conclude the proof of the theorem in the general case.

The general case. Now we are ready to establish Theorem 3.1. A basic ingredient we need is the elementary fact that a permutation on n letters has less than $\sqrt{2n}$ different cycle types.

Proof of Theorem 3.1. By Lemma 3.5, we may restrict to the case when w is not a power word. Set $d_{\mathrm{it}}(x) := C_{\mathrm{it}} x^{1/e_{\mathrm{it}}}$ and note that d_{it} is monotone and concave. Recall from Section 0.1(d) that $\Omega_k := \Omega_k(\sigma)$ denotes the support of all k-cycles of $\sigma \in \mathrm{S}_n$ and $n_k := |\Omega_k|$ for $k \in \mathbb{Z}_+$. Let S be the set of numbers $k \in \mathbb{Z}_+$ such that $n_k > 0$ and note that $|S| < \sqrt{2n}$. Then for $n \geq 2$

$$\begin{aligned} d_{\mathrm{H}}(\sigma, w(\mathrm{S}_n)) &\leq \sum_{k\in\mathbb{Z}_+} \frac{n_k}{n} d_{\mathrm{H}}(\sigma|_{\Omega_k}, w(\mathrm{Sym}(\Omega_k))) \\ &\leq \sum_{k\in S} \frac{n_k}{n} d_{\mathrm{it}}\left(\frac{1}{n_k}\right) \\ &\leq d_{\mathrm{it}}\left(\sum_{k\in S}\frac{1}{n}\right) \leq d_{\mathrm{it}}\left(\frac{\sqrt{2n}}{n}\right) = d_{\mathrm{it}}\left(\sqrt{2/n}\right), \end{aligned}$$

where the second last inequality is implied by Jensen's inequality applied to the concave function d_{it}, and the last one by monotonicity of d_{it}. We can now set $d(x) := d_{\mathrm{it}}(\sqrt{2x}) = \sqrt{2} C_{\mathrm{it}} x^{1/(2e_{\mathrm{it}})}$. This finishes the proof. □

Remark 3.16. In the power word case $w = x^a$ $(a > 1)$, Lemma 3.4 shows that $d_{\mathrm{it}}(x) = ax$ is optimal. Applying the above argument, this produces $d(x) = a\sqrt{2x}$. Lemmas 3.5 and 3.6 demonstrate that the term $\sqrt{x}$ in this expression is 'correct'. However, the coefficient $\sqrt{2}a$ is far from optimal, since it is apparent that $D \leq a-1$, where D is the constant from Lemmas 3.5 and 3.6. Still, the argument from above using Jensen's inequality produces the optimal bound $N(w,\varepsilon) = O((1/\varepsilon)^2)$ in this case.

3.3 Unitary groups

In this section, we present the proof of Theorem 3.3 (Subsection 3.3.1 below) and draw some connections to the article [16] of Elkasapy and Thom (see Subsection 3.3.3).

3.3.1 Proof of Theorem 3.3

Denote by K the one-relator group

$$\langle x, y \,|\, w \rangle = \mathbf{F}_2 / \langle\!\langle w \rangle\!\rangle$$

associated to w.

The key observation is the following lemma involving the second cohomology group of a quotient of the Cayley complex of K, in which we interpret monomial matrices in U_n as 1-cochains.

Let X be the Cayley complex of the presentation $\langle x, y \,|\, w \rangle$ of K, i.e., its 1-skeleton is the directed Cayley graph $\Gamma \coloneqq \mathrm{Cay}(K, \{x, y\})$ and for each vertex $v \in V(\Gamma) = K$ we glue in a 2-cell c_v along w starting at v.

For $\pi\colon K \twoheadrightarrow G$ being a surjective homomorphism to a finite group G of order n set $g \coloneqq \pi(x)$, $h \coloneqq \pi(y)$ and let $X(\pi)$ be the quotient of the 2-complex X induced by π, whose 1-skeleton is the Cayley graph $\Gamma(\pi) \coloneqq \mathrm{Cay}(G, \{g, h\})$ of G. Consider also permutations $\sigma_g, \sigma_h \in \mathrm{S}_n = \mathrm{Sym}(G)$ arising from the action of G on itself. Set $d(\pi) \coloneqq \dim(H^2(X(\pi), \mathbb{R}))$ to be the dimension of the second cohomology group of $X(\pi)$.

Lemma 3.17. *For every diagonal unitary matrix $M \in \mathrm{U}_n$ we can find monomial matrices $M_g, M_h \in \mathrm{U}_1 \wr \mathrm{S}_n \leq \mathrm{U}_n$ such that $M_g = (\lambda_i)_{i=1}^n.\sigma_g$ and $M_h = (\mu_i)_{i=1}^n.\sigma_h$ so that $w(M_g, M_h)$ is diagonal and differs in at most $d(\pi)$ diagonal entries from M. Hence, setting $\varepsilon(\pi) \coloneqq d(\pi)/n$, the image of the word map $w(\mathrm{U}_n)$ is $\varepsilon(\pi)$-dense in U_n with respect to the normalized rank metric.*

Proof. Write $C_\bullet(\pi)$ resp. $C^\bullet(\pi)$ for the chain resp. cochain complex over $\mathbb{R}$ associated to $X(\pi)$ with differentials

$$d_i(\pi)\colon C_i(\pi) \to C_{i-1}(\pi) \text{ resp. codifferentials } d^i(\pi) \coloneqq (d_i(\pi))^*\colon C^{i-1}(\pi) \to C^i(\pi)$$

for $i \in \mathbb{N}$. A 1-cochain $\alpha\colon X_1(\pi) \to \mathbb{R}$ assigns to each edge e of $\Gamma(\pi)$ a real number α_e. Then the Cayley graph $\Gamma(\pi)$ together with this assignment encodes two elements $g_\alpha, h_\alpha \in \mathbb{R} \wr \mathrm{S}_n \subseteq \mathbb{R}^{n\times n}$, where the permutation part of g_α resp. h_α is given by the action of g resp. h on the vertices $V(\Gamma(\pi)) = G$ of $\Gamma(\pi)$, and the first part is induced by the values α_e ($e \in E(\Gamma(\pi))$). The group $\mathbb{R} \wr \mathrm{S}_n$ can be seen as the set of monomial matrices in $\mathbb{R}^{n\times n}$, where the entries marked along the corresponding permutations are added instead multiplied.

Given the 1-cochain α, its image under the codifferential $d^2(\pi)\colon C^1(\pi) \to C^2(\pi)$ is defined by

$$d^2(\pi)(\alpha)(c) = \sum_{e \in \partial(c)} \varepsilon_e \alpha(e)$$

for all $c \in X_2(\pi)$, where $\partial(c)$ is the set of edges of the boundary of the cell c and $\varepsilon_e \in \{\pm 1\}$ is the corresponding orientation. Now $C^2(\pi)/\operatorname{im}(d^2(\pi)) = H^2(X(\pi), \mathbb{R})$.

Choose $M = \operatorname{diag}(\lambda_v)_{v \in G} \in \mathrm{GU}(\ell^2 G) = \mathrm{U}_n$ arbitrary and find $\beta_v \in \mathbb{R}$ such that $\lambda_v = e^{i\beta_v}$ for $v \in G$. Then there exists a function $\alpha\colon X_1(\pi) \to \mathbb{R}$ such that $d^2(\pi)(\alpha)(c_v) = \beta_v$ for all but at most $d(\pi)$ vertices $v \in V(\Gamma(\pi)) = G$.

But, letting $\varphi\colon \mathbb{R} \wr \mathrm{S}_n \to \mathrm{U}_1 \wr \mathrm{S}_n \leq \mathrm{U}_n$ be the homomorphism induced by exponentiation, we also see that

$$w(\varphi(g_\alpha), \varphi(h_\alpha)) = (e^{i d^2(\pi)(\alpha)(c_v)})_{v \in G}.\operatorname{id}.$$

Hence $M_g := \varphi(g_\alpha)$, $M_h := \varphi(h_\alpha)$ is a suitable choice of matrices. The last statement of the lemma follows from the definition of the normalized rank metric on U_n (see Definition 0.12). This completes the proof. ☐

Remark 3.18. Subsequently, for a chain $x \in C_i(\pi)$ $(i = 0, 1, 2)$ write $x^* \in C^i(\pi)$ for the corresponding dual cochain defined by $\langle x, \cdot \rangle = x^*$, where $\langle \cdot, \cdot \rangle$ is the inner product associated to the basis $X_i(\pi)$ of $C_i(\pi)$.

If $w \in \mathbf{F}_2'$, it is clear that $d(\pi) \geq 1$, as any element

$$\sum_{v \in G} \lambda_v c_v^* \in \operatorname{im}(d^2(\pi))$$

lies in the hyperplane given by $\sum_{v \in G} \lambda_v = 0$ (here we use that G is finite). This reflects the fact that then $w(\mathrm{U}_n) \subseteq \mathrm{SU}_n$. Moreover, in this case, if $d(\pi) = 1$, the word map $w\colon \mathrm{SU}_n \times \mathrm{SU}_n \to \mathrm{SU}_n$ is surjective (by transitivity we can then achieve equality on any $n-1$ diagonal entries in the above proof). Namely, if $w(g,h) = u$ for $g, h \in \mathrm{U}_n$ and $u \in \mathrm{SU}_n$, we can find $\lambda, \mu \in \mathbb{C}$ such that $\lambda^n = \det(g)$ and $\mu^n = \det(h)$. Then $g' := \lambda^{-1} g$, $h' := \mu^{-1} h$ lie in SU_n and satisfy $w(g', h') = u$.

In the opposite case, when $w \in \mathbf{F}_2 \setminus \mathbf{F}_2'$, either $w(1,x)$ or $w(x,1)$ is of the form x^m for $m \in \mathbb{Z} \setminus \{0\}$. So the word w is always surjective on U_n and SU_n, since every element of these groups is diagonalizable and hence has an mth root (of determinant one in the case of SU_n).

Remark 3.19. In the situation of the proof of Lemma 3.17 write $C_\bullet = C_\bullet(X)$ resp. $C^\bullet = C^\bullet(X)$ for the chain resp. cochain complex over $\mathbb{R}$ associated to X. Then we have a commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & C^0 & \xrightarrow{d^1} & C^1 & \xrightarrow{d^2} & C^2 & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & C^0(\pi) & \xrightarrow{d^1(\pi)} & C^1(\pi) & \xrightarrow{d^2(\pi)} & C^2(\pi) & \longrightarrow & 0
\end{array},$$

where the top arrows are K-equivariant, the bottom arrows are G-equivariant, and the vertical arrows are induced by π. The duality defined in Remark 3.18 identifies $C^i(\pi)$ G-equivariantly with $C_i(\pi)$. Then, further identifying via the isomorphisms

$$C_1(\pi) \cong g \cdot \mathbb{R}[G] \oplus h \cdot \mathbb{R}[G] \quad \text{and} \quad C_2(\pi) \cong \mathbb{R}[G],$$

where the latter is given by $c_v \mapsto v$ $(v \in G)$, and letting ${}^*\colon \mathbb{R}[G] \to \mathbb{R}[G]$ be the natural involution given by $g \mapsto g^{-1}$ $(g \in G)$, one computes that the map $d^2(\pi)$ is then given by $d^2(\pi)(g \cdot 1) = \pi(\partial w/\partial x)^*$ and $d^2(\pi)(h \cdot 1) = \pi(\partial w/\partial y)^*$. Here $\partial w/\partial x$ resp. $\partial w/\partial y$ denote the Fox derivative [19] of w with respect to x resp. y, i.e., if $w = x^{a_1}y^{b_1} \cdots x^{a_l}y^{a_l}$ and $\varepsilon_i := \operatorname{sgn}(a_i)$, $\delta_i := \operatorname{sgn}(b_i)$, then

$$\begin{aligned} \frac{\partial w}{\partial x} &= \sum_{i=1}^{l} \varepsilon_i x^{a_1}y^{b_1} \cdots x^{a_{i-1}}y^{b_{i-1}}x^{\frac{\varepsilon_i - 1}{2}}(1 + x^{\varepsilon_i} \cdots + x^{\varepsilon_i|a_i-1|}), \\ \frac{\partial w}{\partial y} &= \sum_{i=1}^{l} \delta_i x^{a_1}y^{b_1} \cdots x^{a_i}y^{\frac{\delta_i - 1}{2}}(1 + y^{\delta_i} \cdots + x^{\delta_i|b_i-1|}). \end{aligned}$$

Later we will apply Lemma 3.17 to a family of surjective homomorphisms $\pi(p)\colon K \twoheadrightarrow H(p)$ (p a sufficiently large prime) to finite groups $H(p)$ of order $n_p = p^h$ (h is a constant defined later) such that $\varepsilon(\pi(p)) \to 0$ as $p \to \infty$. This is only possible if the corresponding map $d^2(\pi(p))$ in the above proof for $G = H(p)$ is non-trivial for sufficiently large p. Hence next we characterize when this happens for an arbitrary homomorphism $\pi\colon K \twoheadrightarrow G$.

Lemma 3.20. *Let $G = \mathbf{F}_2/N = \langle x, y\rangle/N$ be a (not necessarily finite) quotient of the one-relator group K, i.e., $w \in N$, and set $g := \overline{x}$, $h := \overline{y}$ in G. Define $\pi\colon K \twoheadrightarrow G$ as in Lemma 3.17. Then $d^2(\pi)$ in the proof of Lemma 3.17 is identically zero if and only if $w \in N' = [N, N]$.*

Proof. By assumption, we have $w \in N$. If $w \in N'$, then $w = \prod_{i=1}^k n_i$ is a product of elements $n_i \in N$ $(i = 1, \ldots, k)$, where the multiset $(n_i)_{i=1}^k$ equals $(n_i^{-1})_{i=1}^k$. Consider the w-loop $l_v(w)$ with arbitrary starting vertex $v \in V(\Gamma(\pi))$. The above shows that each subloop of $l_v(w)$ associated to n_i $(i = 1, \ldots, k)$ returns to v and hence any edge in $\Gamma(\pi)$ is traversed equally often in both directions. But then one sees immediately that $d^2(\pi) = 0$.

Conversely, the assumption $d^2(\pi) = 0$ implies that the loops $l_v(w)$ $(v \in V(\Gamma(\pi)))$ traverse all of its edges equally often in both directions. Let Δ be the undirected simple graph which is the image of the loop $l_v(w)$. Then Δ is homotopic to a bouquet of circles each of which is traversed equally often in both directions by the loop l corresponding to $l_v(w)$ under the chosen homotopy. But this means precisely that the homotopy class of l lies in $\pi_1(\Delta)'$. Pulling back the generators of the group $\pi_1(\Delta)$ to elements of N, we see that $w \in N'$. □

Now we show how to define the maps $\pi(p)\colon K \twoheadrightarrow H(p)$ and quotients $H(p)$ appropriately (for p a sufficiently large prime) such that $\varepsilon(\pi(p)) \to 0$ for $p \to \infty$.

Since $\mathbf{F}_2$ is residually nilpotent, there exists a unique integer $c = c(w) \geq 0$ such that $w \in \gamma_{c+1}(\mathbf{F}_2) \setminus \gamma_{c+2}(\mathbf{F}_2)$. Set $H := \mathbf{F}_2/\gamma_{c+1}(\mathbf{F}_2)$ to be the free 2-generated nilpotent group of class c (in which w is trivial) and let $\pi\colon K \twoheadrightarrow H$ be the corresponding quotient map.

By Jennings' embedding theorem, every finitely generated torsion-free nilpotent group N can be embedded into the group $U := \mathrm{UT}_d(\mathbb{Z})$ of upper unitriangular matrices over $\mathbb{Z}$ (for $d \in \mathbb{Z}_+$ sufficiently large; such an embedding can even be explicitly computed from a polycyclic representation of N by an algorithm due to Nickel [54]; see also [30] and [64]). Since the factors of the lower central series $\gamma_i(\mathbf{F}_2)/\gamma_{i+1}(\mathbf{F}_2)$ $(i = 1, \ldots, c)$ are free abelian, H is a poly-$\mathbb{Z}$ group and we obtain that H can be concretely realized as a subgroup of $\mathrm{UT}_d(\mathbb{Z})$ for some dimension $d = d(w)$.

Define the central series $H_i := H \cap \gamma_i(U)$ of H for $i = 1, \ldots, d$ (note that the group $\gamma_l(U)$ consists of the upper unitriangular matrices $u = (u_{ij}) \in U$ with $u_{ij} = 0$ for $1 \leq j-i \leq l-1$). Then $H_i/H_{i+1} \leq \gamma_i(U)/\gamma_{i+1}(U) \cong \mathbb{Z}^{d-i}$ (the ith off-diagonal). Let $B_i \subseteq \mathbb{Z}^{d-i}$ be a basis of H_i/H_{i+1} and set $h_i := \dim(H_i/H_{i+1}) = |B_i|$ $(i = 1, \ldots, d-1)$. Let p be a prime not dividing some $h_i \times h_i$ minor of the $(d-i) \times h_i$-matrix associated to B_i for all $i = 1, \ldots, d-1$. Then, writing $U(p) := \mathrm{UT}_d(\mathbb{Z}/(p))$ and letting $H(p), H_i(p)$ $(i = 1, \ldots, d)$ be the image of H, H_i in $U(p)$, we see that $H_i/H_{i+1} \cong \mathbb{Z}^{h_i} \twoheadrightarrow H_i(p)/H_{i+1}(p) \cong (\mathbb{Z}/(p))^{h_i}$. Define $\pi(p)\colon K \twoheadrightarrow H(p)$ to be the induced quotient map to the finite p-group $H(p)$.

Now refine the central series $(H_i)_{i=1}^{d}$ to a central series $(L_j)_{j=1}^{h+1}$ such that $L_j/L_{j+1} \cong \mathbb{Z}$ for $j = 1, \ldots, h$, where $h := \sum_{i=1}^{d-1} h_i$ is the Hirsch length of H. Then still $L_j/L_{j+1} \cong \mathbb{Z} \twoheadrightarrow L_j(p)/L_{j+1}(p) \cong \mathbb{Z}/(p)$ for $j = 1, \ldots, h$ and p as above. Let $x_j \in H$ be such that $\langle x_j \rangle L_{j+1} = L_j$ for $j = 1, \ldots, h$. Note that the map $d^2(\pi)$ associated to the surjective homomorphism $\pi\colon K \twoheadrightarrow H$ is non-trivial, since if it where trivial, then by Lemma 3.20 applied to π and $N = \ker(\pi) = \gamma_{c+1}(\mathbf{F}_2)$ we would have $w \in N' = [\gamma_{c+1}(\mathbf{F}_2), \gamma_{c+1}(\mathbf{F}_2)] \leq \gamma_{c+2}(\mathbf{F}_2)$, which is not the case by the choice of c. Hence from the local nature of the definition of $d^2(\pi)$ it follows that there is an edge $e \in E(\Gamma(\pi))$ such that $0 \neq d^2(\pi)(e^*) = \sum_v \lambda_v c_v^*$ with $\lambda_v \in \mathbb{Z} \setminus \{0\}$. This element corresponds to the element $0 \neq y = \sum_v \lambda_v v \in \mathbb{Z}[H]$ in the group ring. Subsequently, let k be a field of large enough characteristic such that the image of y in $k[H]$ is non-trivial, and for $z \in k[H]$ an element in the group algebra of H, write $z(p)$ for its image in $k[H(p)]$. It follows that for p large enough the elements v in the support $\mathrm{supp}(y) \subseteq H$ are mapped injectively to the elements $v(p) = \pi(p)(v) \in \mathrm{supp}(y(p)) \subseteq H(p)$ (e.g., take p larger than all matrix entries of elements v form $\mathrm{supp}(y)$).

Now the k-dimension of $\mathrm{im}(d^2(\pi(p)))$ can be bounded from below by the k-dimension of the right ideal $y(p)k[H(p)]$ as the action of $H(p)$ on $C^2(\pi(p))$ (here with coefficients in k) equals its right action on the group algebra $k[H(p)]$. We bound the dimension of the latter from below by the following lemma.

Lemma 3.21. *In this situation the right ideal $y(p)k[H(p)] \subseteq k[H(p)]$ generated by $y(p)$ has k-dimension at least $(p-f)^h$ for a constant $f = f(w)$ only depending on w.*

Proof. For $l = 0, \dots, h$ write

$$y = \sum_{e \in \mathbb{Z}^{h-l}} x_1^{e_1} \cdots x_{h-l}^{e_{h-l}} c_e$$

for $e = (e_1, \dots, e_{h-l})$ and $c_e \in \mathbb{Z}[L_{h+1-l}]$. We prove by induction on l that for all $e \in \mathbb{Z}^{h-l}$ the right ideal $c_e(p)k[L_{h+1-l}(p)]$ is either zero or has k-dimension at least $(p-f)^l$, obtaining the claim for $l = h$ as $c_\emptyset = y \neq 0$ and so $y(p) \neq 0$ for p large enough.

For $l = 0$ there is nothing to prove, as c_e is either zero or it spans a one-dimensional ideal in $k = k[L_{h+1}]$. Now for the induction step assume the statement is proven for $l \geq 0$. Let $e \in \mathbb{Z}^{h-1-l}$ be arbitrary and write $c_e = \sum_{i \in \mathbb{Z}} x_{h-l}^i c_{(e,i)}$. If $c_e = 0$, we are done, so assume the opposite. Then, certainly, the set $S \coloneqq \{i \in \mathbb{Z} \mid c_{(e,i)} \neq 0\} \neq \emptyset$ is an invariant of y and so of w, hence $m \coloneqq \max S - \min S \leq f(w)$ for some function f of w. Set $z \coloneqq x_{h-l}^{\min S}$. Since the right $k[L_{h-l}(p)]$-ideals generated by $c_e(p)$ and $(z^{-1}c_e)(p)$ have the same dimension, we may consider the element $u \coloneqq z^{-1}c_e$ instead of c_e. This equals $u = \sum_{i=0}^m x_{h-l}^i c_{(e,i+\min S)}$. Now it is easy to see that the set of linear combinations $\sum_{k=0}^{p-m-1} (u x_{h-l}^k d_k)(p)$ with $d_k \in k[L_{h+1-l}]$ arbitrary for $k = 0, \dots, p-m-1$ generate a k-subspace of dimension at least $(p-m)(p-f)^l$. Indeed, by choosing d_0 appropriately, one can obtain any element of $c_{(e,\min S)}k[L_{h+1-l}]$ as the left coefficient in $k[L_{h+1-l}]$ of x_{h-l}^0. Then, choosing d_1 such that $x_{h-l}d_1x_{h-l}^{-1} \in k[L_{h+1-l}]$ is appropriate, one can obtain any left coefficient in front of x_{h-l}^1 in some coset of the right ideal $c_{(e,\min S)}k[L_{h+1-l}]$, etc. Since by assumption $p - m \geq p - f$, we are done. □

Now, as a consequence of Lemma 3.21, we obtain the following immediate corollary.

Corollary 3.22. *Applying Lemma 3.17 to $\pi(p)$ as above, we obtain that $\varepsilon(\pi(p)) \leq 1 - (1 - f/p)^h \leq hf/p = hfn^{-1/h}$ for $n = p^h$, where $h = h(w)$ and $f = f(w)$ are defined as above.*

Proof. Lemma 3.21 and the comment preceding it imply that

$$\dim(H^2(X(\pi(p)), \mathbb{R})) \leq p^h - (p-f)^h.$$

Normalizing, we obtain the desired identity. □

The homomorphisms $\pi(p)\colon K \twoheadrightarrow H(p)$ for $p \geq p_0 = p_0(w)$ a sufficiently large prime suffice now to prove the quantitative version of Theorem 3.3 given in the introduction.

Namely, one proves by induction on $n \geq 1$ that for $D = D(w) > 0$ sufficiently large

$$d_{\mathrm{rk}}(g, w(\mathrm{U}_n)) \leq \varepsilon(n) \coloneqq (D\log(n) + 1)n^{-1/h}$$

for all $g \in \mathrm{U}_n$. Set $\varepsilon(0) \coloneqq 0$.

Indeed, this is true for $n < p_0^h$. Now for $n \geq p_0^h$ we pick the largest prime p such that $p^h \leq n$ and the largest integer $l \geq 1$ such that $lp^h \leq n$. Then via the embedding

$\mathrm{U}_{n-lp^h} \oplus \mathrm{U}_{p^h}^{\oplus l} \leq \mathrm{U}_n$, writing $n_1 \coloneqq n - lp^h$ and $n_2 \coloneqq lp^h$, we see that

$$d_{\mathrm{rk}}(g, w(\mathrm{U}_n)) \leq \frac{n_1}{n}\varepsilon(n_1) + \frac{n_2}{n}hf/p$$

for all $g \in \mathrm{U}_n$ by the induction hypothesis and Corollary 3.22. Since p is the largest prime such that $p \leq n^{1/h}$, Bertrand's postulate implies that $p \geq n^{1/h}/2$. Moreover, by construction $n_1 < n/2$, so that the above term can be bounded by

$$\frac{1}{2}(D\log(n/2)+1)(n/2)^{-1/h} + 2hfn^{-1/h} \leq (D\log(n)+1)n^{-1/h}$$

again if D is large enough.

3.3.2 Further implications

It is easy to see that our method of proof implies that $w(\mathrm{SU}_n)$ has width at most two in SU_n for n large enough, which was first proven in [36, Theorem 2.3] using Gotô's trick, Borel's theorem and the representation theory of SU_2. The reason for this is the following basic fact about the linearized permutation representation of the Weyl group S_n.

Lemma 3.23. *Let $V = \mathbb{R}^n$ be the permutation representation of S_n. Let V_0 be the subrepresentation of V of all vectors whose entries sum to zero. If $U_1, U_2 \leq V_0$ are subspaces and $\dim(U_1) + \dim(U_2) \geq n-1$, then $U_1 + U_2.\sigma = V_0$ for some $\sigma \in \mathrm{S}_n$.*

Proof. This is a consequence of the fact that the exterior power $\Lambda^k(V_0)$ is irreducible for $k = 0, \ldots, n-1$, see Proposition 3.12 of [21]. Note that the determinant pairing $h\colon \Lambda^k(V_0) \times \Lambda^{n-1-k}(V_0) \to \mathbb{R}$ given by $h(v_1\wedge\cdots\wedge v_k, v_{k+1}\wedge\cdots\wedge v_{n-1}) = v_1\wedge\cdots\wedge v_{n-1} \in \Lambda^{n-1}(V_0) \cong \mathbb{R}$, is non-degenerate. Using irreducibility, we can easily see that this implies the claim. Indeed, set $k \coloneqq \dim(U_1)$, choose a basis $u_1, \ldots, u_k$ for U_1 and a linearly independent set $u'_{k+1}, \ldots, u'_{n-1}$ in U_2. Now we have that $h(u_1 \wedge \ldots \wedge u_k, (u'_{k+1} \wedge \ldots \wedge u'_{n-1}).\sigma) \neq 0$ if and only if $U_1 + U_2.\sigma = V_0$. By irreducibility, the set $\{(u'_{k+1} \wedge \ldots \wedge u'_{n-1}).\sigma \mid \sigma \in \mathrm{S}_n\}$ spans $\Lambda^{n-1-k}(V_0)$. The fact that h is non-degenerate implies the claim. □

We immediately obtain the following corollary.

Corollary 3.24. *Let $w_1, w_2 \in \mathbf{F}_2$ be non-trivial. Then*

$$w_1(\mathrm{SU}_n)w_2(\mathrm{SU}_n) = \mathrm{SU}_n$$

for n sufficiently large.

Proof. Set $U_i \leq w_i(\mathbb{R} \wr \mathrm{S}_n) \cap \mathbb{R}^n \leq V \coloneqq \mathbb{R}^n$ to be a vector subspace of the diagonal matrices $\mathbb{R}^n$ which lies in the above w_i-image and has maximal dimension with respect to this property ($i = 1, 2$). In Lemma 3.17, Corollary 3.22, and the remarks thereafter, we have shown that $\dim(U_i) \geq \frac{n-1}{2}$ for n large enough. Applying Lemma 3.23 to U_1, U_2 and V, and exponentiating, we see that for every diagonal matrix $g \in \mathrm{SU}_n$ there are $h_i \in w_i(\mathrm{SU}_n)$ ($i = 1, 2$) such that $g = h_1 h_2^\sigma$, so that $w_1(\mathrm{SU}_n)w_2(\mathrm{SU}_n) = \mathrm{SU}_n$. □

3.3.3 Concluding remarks

Lemma 3.17 and the above proof of Theorem 3.3 can be seen as a generalization of the methods used in [16] and clarify various aspects of it. Let us demonstrate this briefly. For a word $w \in \mathbf{F}_2' \setminus \mathbf{F}_2''$ set $\pi\colon K \twoheadrightarrow H \coloneqq \mathbf{F}_2/\mathbf{F}_2' = \mathbb{Z}^2 = \langle g, h\rangle$ to be the natural homomorphism. Applying Lemma 3.20 to π, we see that $d^2(\pi)$ is non-trivial, so again we get an edge $e \in \Gamma(\pi)$ such that $d^2(\pi)(e^*) \neq 0$, which corresponds to an element $z = z(g,h) \in \mathbb{Z}[H] = \mathbb{Z}[\mathbb{Z}^2] = \mathbb{Z}[g^{\pm 1}, h^{\pm 1}]$ in the integral group ring. By symmetry, it is no loss to assume that e is rooted at 1_H and labeled by h.

The Laurent polynomial $p_w(X)$ defined in Section 3 of [16] is now precisely equal to $z^*(X,1) = z(X^{-1},1)$, where $z(g,h) = d^2(\pi)(e^*) = \pi(\partial w/\partial y)^*$ (see Remark 3.19). Here $z^*(X,1)$ is just the image of z under the homomorphism $\mathbb{Z}[\mathbb{Z}^2] = \mathbb{Z}[g^{\pm 1}, h^{\pm 1}] \twoheadrightarrow \mathbb{Z}[\mathbb{Z}] = \mathbb{Z}[X^{\pm 1}]$ induced by $g^a h^b \mapsto X^{-a}$, as, e.g., for $w = [x^a, y^b] = x^{-a}y^{-b}x^a y^b$, $a, b > 0$ we have $z(g,h) = (h + \cdots + h^b)(1 - g^a)$ and $p_w(X) = -b(X^{-a} - 1)$.

Now we can find a suitable homomorphism $\varphi\colon \mathbb{Z}^2 = \langle g, h\rangle \twoheadrightarrow \mathbb{Z} = \langle X\rangle$ such that the induced ring homomorphism $\mathbb{Z}[\mathbb{Z}^2] = \mathbb{Z}[g^{\pm 1}, h^{\pm 1}] \twoheadrightarrow \mathbb{Z}[\mathbb{Z}] = \mathbb{Z}[X^{\pm 1}]$ maps z to a non-zero element $\varphi(z) = p(X)$ (e.g., take as the kernel of the homomorphism φ a saturated copy of $\mathbb{Z}$ in $\mathbb{Z}^2$ which does not hit any element in the support of z). For $n \in \mathbb{Z}_+$ we define the homomorphism $\pi(n)\colon K \twoheadrightarrow H(n) = \mathbb{Z}/(n)$ just by composing $\varphi \circ \pi$ with the natural projection $\mathbb{Z} \twoheadrightarrow \mathbb{Z}/(n)$. One now quickly derives the conclusions of Lemma 3.1, Corollary 3.2, and Proposition 3.8 of [16] from the following lemma.

Lemma 3.25. *Let $p(X)$ be as above. Write $z(n)$ for the image of z in $\mathbb{Z}[H(n)]$. Define $W_n \coloneqq \{\omega \in \mathbb{C} \mid p(\omega) = 0 \text{ and } \omega^n = 1\}$. The (right) ideal $z(n)\mathbb{R}[H(n)]$ has codimension $|W_n|$, so in particular, if the least prime dividing n is large enough, then it has codimension one and the word map w on SU_n is surjective by Lemma 3.17 and Remark 3.18.*

Proof. By the Chinese remainder theorem, we have the isomorphism

$$\mathbb{R}[H(n)] = \mathbb{R}[\mathbb{Z}/(n)] \cong \mathbb{R}[X]/(X^n - 1) \cong \bigoplus_{\substack{\chi \mid X^n - 1 \\ \chi \text{ irreducible}}} \mathbb{R}[X]/(\chi) \cong \mathbb{R}^{e_n} \oplus \mathbb{C}^{\lceil n/2\rceil - 1},$$

where $e_n = 2$ if n is even and $e_n = 1$ if n is odd. This holds, since the (monic) irreducible polynomials $\chi \mid X^n - 1$ are either of the form $X \pm 1$ (so that $\mathbb{R}[X]/(\chi) \cong \mathbb{R}$) or of the form $(X - \omega)(X - \overline{\omega})$ for $\omega \in \mathbb{C} \setminus \mathbb{R}$ an nth root of unity (so that $\mathbb{R}[X]/(\chi) \cong \mathbb{C}$). The last isomorphism in the above equation is given by $\overline{X} \mapsto (\omega_\chi)_\chi$, where ω_χ is a root of χ. Hence the ideal generated by $z(n)$ has as codimension precisely the number of nth roots ω for which $p(\omega) = 0$, as claimed. The second claim follows from the fact that $p(X) \in \mathbb{Z}[X^{\pm 1}]$ and the minimal polynomial of a primitive mth root of unity, $m > 1$ dividing n, over $\mathbb{Q}$ is the cyclotomic polynomial $\Phi_m(X)$ of degree $\varphi(m) \geq p - 1$, where p is the least prime divisor of n. Hence, if $p - 1 > \deg(p(X))$, we have $W_n = \{1\}$. This completes the proof. ∎

The above shows that the result from [16] is precisely the simplest application of Lemma 3.17, namely when $G = H(n)$ is taken to be cyclic. We can now also understand that Question 4.4 from [16] has a negative answer. Indeed, assume that for every choice of $\varphi\colon \mathbb{Z}^2 = \langle g, h\rangle \twoheadrightarrow \mathbb{Z} = \langle X\rangle$ in the above construction, the image $\mathrm{im}(d^2(\pi(n))) \subseteq \mathbb{R}[H(n)] \cong \mathbb{R}[X]/(X^n - 1)$ has codimension greater than one, i.e., $d^2(\varphi\circ\pi)(\varphi(g)\cdot 1)$ and $d^2(\varphi\circ\pi)(\varphi(h)\cdot 1)$ as Laurent polynomials in $\mathbb{Z}[X^{\pm 1}]$ have a non-trivial nth root of unity as a common root. Now choose $\alpha \in \mathrm{Aut}(\mathbf{F}_2)$ arbitrary. Replacing w by $\alpha(w)$ will not improve this situation. To see this, set $x' \coloneqq \alpha(x)$, $y' \coloneqq \alpha(y)$ and obtain by the chain rule

$$
\begin{aligned}
d^2_{\alpha(w)}(\varphi\circ\pi)(\varphi(g)\cdot 1) &= \varphi\circ\pi\left(\frac{\partial w}{\partial x}(x',y')\frac{\partial x'}{\partial x} + \frac{\partial w}{\partial y}(x',y')\frac{\partial y'}{\partial x}\right)^*,\\
d^2_{\alpha(w)}(\varphi\circ\pi)(\varphi(h)\cdot 1) &= \varphi\circ\pi\left(\frac{\partial w}{\partial x}(x',y')\frac{\partial x'}{\partial y} + \frac{\partial w}{\partial y}(x',y')\frac{\partial y'}{\partial y}\right)^*,
\end{aligned}
$$

where $d^2_{\alpha(w)}(\varphi\circ\pi)$ denotes the map corresponding to $d^2(\varphi\circ\pi)$ but with $\alpha(w)$ in the role of w. But then, as α is an automorphism, we see that $\varphi\circ\pi(x') = X^a$ and $\varphi\circ\pi(y') = X^b$ with $a, b \in \mathbb{Z}$ coprime (as they must generate $\mathbb{Z} = \langle X\rangle$). Hence

$$
\varphi\circ\pi\left(\frac{\partial w}{\partial x}(x',y')\right) = \frac{\partial w}{\partial x}(X^a, X^b) \quad \text{resp.} \quad \varphi\circ\pi\left(\frac{\partial w}{\partial y}(x',y')\right) = \frac{\partial w}{\partial y}(X^a, X^b),
$$

which is equal to $d^2(\varphi'\circ\pi)(\varphi'(g)\cdot 1)^*$ resp. $d^2(\varphi'\circ\pi)(\varphi'(h)\cdot 1)^*$, where $\varphi'\colon \mathbb{Z}^2 = \langle g,h\rangle \twoheadrightarrow \mathbb{Z} = \langle X\rangle$ is given by $g \mapsto X^a$, $h \mapsto X^b$ (see Remark 3.19). But these two expressions seen as Laurent polynomials in $\mathbb{Z}[X^{\pm 1}]$ by our assumption have a non-trivial nth root of unity ω as a common root. But then, by the above equations, $d^2_{\alpha(w)}(\varphi\circ\pi)(\varphi(g)\cdot 1)$ and $d^2_{\alpha(w)}(\varphi\circ\pi)(\varphi(h)\cdot 1)$ also must have ω as a root, so that the image $\mathrm{im}(d^2_{\alpha(w)}(\pi(n))) \subseteq \mathbb{R}[H(n)] \cong \mathbb{R}[X]/(X^n - 1)$ has codimension greater than one.

In retrospect, as has been pointed out to us by Jack Button, the study in [16] would have been much clearer, when the connection to Fox calculus and the even more classical subject of Alexander polynomials would have been observed from the start.

Let us end this section by drawing some further connections to related facts. In case that K is residually finite, one could also prove Theorem 3.3 using Lück's approximation theorem together with the fact that the second L^2-Betti number of a one-relator group is zero by a well-known result of Dicks and Linnell [9] – or the validity of the L^2-zero divisor conjecture for torsionfree nilpotent groups applied to H (see [52] for more background). However, our argument is much more explicit and does even give an effective estimate.

3.4 Finite groups of Lie type

In this section, we prove Theorem 3.2 using aspects presented in Section 3.2 and 3.3 – for convenience of the reader we decided to present the proof first in the case of unitary groups U_n, where the methods come into play in the most natural way. However, we will now use the same cohomological method as in Lemma 3.17 together with Lemmas 3.20 and 3.21 of Section 3.3, but instead of using the additive group of $\mathbb{R}$ as our coefficient group, we now will use groups of type $(\mathbb{F}_q[X]/(\chi))^\times$ for $\chi \in \mathbb{F}_q[X]$ some polynomial. Indeed, we will need the following modified version of Lemma 3.21, which is an easy consequence of it.

Corollary 3.26. *In the setting of Lemma 3.21, using coefficients in $\mathbb{Z}$ instead of the field k, there is a non-zero $c \in \mathbb{N}$ and $f \in \mathbb{N}$ such that for all large primes p there exists a subset $C \subseteq H(p)$ of at least $(p-f)^h$ coordinates so that the projection of the right ideal $y(p)\mathbb{Z}[H(p)]$ onto $\mathbb{Z}[C]$ contains the module $(c\mathbb{Z})[C]$.*

Proof. Applying Lemma 3.21 to $k = \mathbb{Q}$, we get a set $C \subseteq H(p)$ of coordinates of size $|C| \geq (p-f)^h$ such that $y(p)\mathbb{Q}[H(p)]$ projects surjectively on these. Hence we generate the unit vectors in $\mathbb{Q}[H(p)]/\mathbb{Q}[H(p) \setminus C]$. So multiplying by the least common multiple c of the denominators of the involved scalars, we obtain that the projection of $y(p)\mathbb{Z}[H(p)]$ onto the coordinates C still contains the module $(c\mathbb{Z})[C]$. □

Subsequently, we fix the symbol c to be the constant from Corollary 3.26. For the rest of this chapter, all polynomials from $k[X]$ that occur are meant to be *monic* polynomials. Recall from Section 0.1(e) that for a polynomial $\chi \in k[X]$ for a field k we write $F(\chi)$ for the Frobenius block associated to χ, that is multiplication by $\overline{X}$ in $k[X]/(\chi)$ with respect to the standard monomial basis. Similarly, for $\lambda \in \overline{k}$ write $J_e(\lambda)$ for the Jordan block of size e with respect to λ, that is multiplication by $\lambda + \overline{X}$ in $k[\lambda, X]/(X^e)$. Recall that a polynomial $\chi \in k[X]$ is called primary if it is the power of an irreducible polynomial, i.e., if the ideal it generates is primary. Recall that for an element $g \in \mathrm{End}(V)$, V is irreducible resp. indecomposable resp. cyclic as a $k[X]$-module, where X acts as g, if and only if $g \cong F(i)$ resp. $g \cong F(\chi)$ resp. $\bigoplus_{i=1}^{l} F(\chi_i) \cong F(\chi_1 \cdots \chi_l)$ for an irreducible polynomial $i \in k[X]$ resp. a primary polynomial $\chi \in k[X]$ resp. pairwise coprime primary polynomials $\chi_1, \ldots, \chi_l \in k[X]$.

3.4.1 The linear case

We start by proving Theorem 3.2 in the case when $G = \mathrm{GL}_n(q)$. We consider here the more general case that $G = \mathrm{GL}_n(k)$ for an arbitrary field k. So let $V = k^n$ be the natural module of G. We use the same approach as in Subsection 3.2.4, first approximating isotypic elements $g \in \mathrm{GL}(V)$ by word values, i.e., we first assume that V is the direct sum of isomorphic cyclic $k[X]$-submodules so that $g \cong F(\chi)^{\oplus c_\chi}$ for some polynomial $\chi \in k[X]$ of degree k, and then deducing the general case by using the Frobenius normal form and Jensen's inequality.

The isotypic case. So let χ, c_χ and k be as previously mentioned. We want to approximate $F(\chi)$-isotypic elements by word values with these parameters, so that $n = c_\chi k$. As in Subsection 3.2.4, we distinguish two cases, one in which k is small and one in which it is large (compared to c_χ).

Estimate for small k. In view of Corollary 3.26 we need the following auxiliary fact.

Lemma 3.27. *It holds that $F(\chi(X^c))^c \cong F(\chi)^{\oplus c}$.*

Proof. The block $F(\chi(X^c))$ is the matrix of multiplication by $\overline{X}$ in the ring $k[X]/(\chi(X^c))$, so that $F(\chi(X^c))^c$ is the multiplication by $\overline{X}^c$ in $k[X]/(\chi(X^c))$. But $k[X]/(\chi(X^c)) = \bigoplus_{i=0}^{c-1} \overline{X}^i \langle \overline{X}^c \rangle_k$ holds for dimension reasons, so that the claim follows. □

Now we use the same idea as in Lemma 3.17 with appropriate coefficient group. Consider the ring $R := k[X]/(\chi(X^c))$ and write $c_\chi = rc + s$ for $r \in \mathbb{N}$ and $0 \leq s < c$. Corollary 3.26 and Lemma 3.27 give us that in $R^\times \wr \mathrm{Sym}(r) \leq \mathrm{GL}_{ckr}(k)$ we have that $w(R^\times \wr \mathrm{Sym}(r))$ approximates the block diagonal matrix $(F(\chi(X^c))^{\oplus r})^c \cong F(\chi)^{\oplus cr}$ up to an error of $d(1/r)$. Hence, since the function d is concave, we obtain

$$d_{\mathrm{rk}}(g, w(\mathrm{GL}_n(k))) \leq \frac{cr}{c_\chi} d(1/r) + s/c_\chi < d(c/c_\chi) + c/c_\chi.$$

Estimate for large k. On the other hand, the matrices $F(\chi)$ and $F(X^k - 1)$ differ only in the last row, so by rank one. The last matrix is the permutation matrix of a k-cycle, which we can approximate by word values by the result for symmetric groups (Theorem 3.1). Hence $d_{\mathrm{rk}}(F(\chi), F(X^k - 1)) \leq 1/k$, implying that

$$d_{\mathrm{rk}}(g, w(\mathrm{GL}_n(k))) < d(1/k) + 1/k.$$

Global estimate for isotypic elements. Now we combine both estimates as in the proof for symmetric groups. Using the first estimate if $k \leq \sqrt{n/c}$ and the second in the opposite case, we obtain

$$d_{\mathrm{rk}}(g, w(\mathrm{GL}_n(q))) < d\left(\sqrt{c/n}\right) + \sqrt{c/n},$$

as wished. Subsequently, in analogy to the proof of Theorem 3.1 in Section 3.2, write d_{it} for the function of $1/n$ on the right.

The general case. Using the Frobenius normal form, we can write $g \cong \bigoplus_{k\in\mathbb{Z}_+} F(\chi_k)^{\oplus c_k}$, χ_k being the invariant factor of degree k and $c_k := c_{\chi_k}$.

Now we can finish the proof. Define the set $S \subseteq \mathbb{Z}_+$ as in Subsection 3.2.4. Writing $n_k := c_k k$, we get that

$$\begin{aligned} d_{\mathrm{rk}}(g, w(\mathrm{GL}_n(k))) &\leq \sum_{k\in S} \frac{n_k}{n} d_{\mathrm{it}}(1/n_k) \\ &\leq d_{\mathrm{it}}\left(\sum_{k\in S} 1/n\right) \leq d_{\mathrm{it}}\left(\sqrt{2/n}\right), \end{aligned}$$

as at the end of Subsection 3.2.4, as desired.

Remark 3.28. Similarly to the symmetric case, one verifies that such a bound is also attained for power words $w = x^p$ when $\mathrm{char}(k) = p$.

3.4.2 The case of quasisimple groups of Lie type stabilizing a form

We proceed by proving Theorem 3.2 for quasisimple groups of Lie type of unbounded rank which stabilize a form. Recall the notation of Section 0.2 and the setting from Section 0.1(f):

The group G is of the form $\mathrm{Sp}_{2m}(q)$, $\mathrm{GO}_{2m+1}(q)$, $\mathrm{GO}_{2m}^{\pm}(q)$ or $\mathrm{GU}_n(q)$ ($n \geq 2$, $m \geq 1$), $V = k^n$ is the natural module of G, and f is the form stabilized by G. If we are in Case $(*)$ of Section 0.1(f), i.e., G is orthogonal and $p = \mathrm{char}(k) = 2$, recall that Q denotes the quadratic form inducing f. In the unitary case, f is semilinear in the second entry with respect to the q-Frobenius endomorphism $\sigma\colon k \to k$; $x \mapsto x^q$. In the other cases, let σ denote the identity on k. The sign $\varepsilon \in \{\pm 1\}$ of f is defined to be $+1$ if f is symmetric bilinear or conjugate-symmetric sesquilinear, and to be -1 if f is alternating.

For a fixed $g \in G$, which we want to approximate by w-values, subsequently, consider V as a $k[X]$-module, where X acts as g. If we are not in Case $(*)$, a non-singular submodule of V is said to be *orthogonally indecomposable* if it is not an orthogonal direct sum of non-trivial proper submodules (with respect to the form f). In Case $(*)$, when $\mathrm{rad}(f)$ is one-dimensional, V is said to be orthogonally indecomposable, if $V/\mathrm{rad}(f)$ is orthogonally indecomposable with respect to $\overline{f}$ (the induced form on the quotient).

In analogy to the linear case, V (resp. $V/\mathrm{rad}(f)$ in Case $(*)$) is the orthogonal direct sum of such submodules. Hence, following the same strategy as in Subsection 3.4.1, we first consider the case when V is itself orthogonally indecomposable. To provide a pleasant presentation, we recall the classification of such modules V in the case that $p = \mathrm{char}(k) \neq 2$ or that g is non-unipotent for an arbitrary field k (all statements are well-known and are, e.g., used in Section 6 of [49]; see also [75]). When $p = 2$, k is finite, and g is unipotent, we refer to Section 3 of [25].

§1 Structure of orthogonally indecomposable modules

In this subsection, we recall the classification of orthogonally indecomposable $k[X]$-modules, where k is an arbitrary (not necessarily perfect) field to provide a concise reference of this topic. All the results are summed up for finite fields in Fact 3.40. If the reader is not interested in the subsequent technical details, we propose to skip this subsubsection.

Auxiliary facts. We start by collecting some auxiliary facts, which we will use in the classification.

Lemma 3.29. *Fix a k-vector space U. Let $R \subseteq \mathrm{End}(U)$ be a faithful cyclic representation of the abelian k-algebra R, i.e., there is a vector $u \in U$ such that $u.R = U$. Then $\mathbf{C}_{\mathrm{End}(U)}(R) = R$. In particular, if an abelian group $A \subseteq \mathrm{End}(U)$ acts on the vector space U with a cyclic vector, then $\mathbf{C}_{\mathrm{End}(U)}(A) = k[A]$.*

Proof. Let $c \in \mathbf{C}_{\mathrm{End}(U)}(R)$ and u be a cyclic vector as in the lemma. Then the value $u.c$ of c determines c, since for any $v \in U$ we find $r \in R$ such that $u.r = v$ and so $v.c = u.rc = u.cr$. Hence, if $u.c = 0$, we must have $c = 0$. For an arbitrary c, by assumption we find $r \in R$ such that $u.c = u.r$. Then $u.(c-r) = 0$ and as R is abelian, $c - r \in \mathbf{C}_{\mathrm{End}(U)}(R)$. But then $c - r = 0$ by the preceding argument, so that $c \in R$. $\square$

Remark 3.30. This is a perfect analog of the corresponding statement for transitive abelian permutation groups, see [37, page 158, Hilfssatz 3.1].

For a k-vector space U, subsequently, let U^* denote the *σ-semilinear functionals* on U. Equipping it with the dual action $(u^*.g)(u) \coloneqq u^*(u.g^{-1})$ for $u \in U$, $u^* \in U^*$, the space U^* becomes a $k[X]$-module, too. We will also need the following simple connection between the centralizer of an action and the forms it stabilizes.

Lemma 3.31. *Let H be a group acting on the k-vector spaces U and U'. Let $f\colon U \times U' \to k$ be an H-invariant non-singular bilinear resp. σ-sesquilinear form over k (in the last case, it is σ-semilinear in U'). Then any other such H-invariant form $h\colon U \times U' \to k$, which is not necessarily non-singular, is of the form $h = f(\bullet.c, \bullet)$ for an element $c \in \mathbf{C}_{\mathrm{End}(U)}(H)$.*

Proof. The form f is the same as an H-equivariant bijective linear map $\varphi_f\colon U \to U'^*$. Consider the same map φ_h for h.

Set $\psi \coloneqq \varphi_h \circ \varphi_f^{-1} \in \mathbf{C}_{\mathrm{End}(U'^*)}(H)$, which exists as φ_f is invertible by assumption. Define $d \coloneqq \psi^* \in \mathbf{C}_{\mathrm{End}(U')}(H)$ as the adjoint of ψ with respect to the natural pairing between U'^* and U', i.e., via $(u^*.\psi)(u) = u^*(u.d)$ for all $u \in U'$, $u^* \in U'^*$. Then $h(u,v) = \varphi_h(u)(v) = (\psi \circ \varphi_f(u))(v) = (\varphi_f(u).\psi)(v) = \varphi_f(u)(v.d) = f(u, v.d)$. Taking as c the adjoint of d with respect to f, we get the claim. $\square$

If G is orthogonal and $p = 2$, we still need the following fact.

Lemma 3.32. *Assume we are in Case $(*)$ of Section 0.1(f), i.e., G is orthogonal and $p = \mathrm{char}(k) = 2$. Let $Q, Q'\colon U \to k$ be quadratic forms inducing the symmetric bilinear form f on a k-vector space U. Then $Q - Q'$ is a semilinear form on U with respect to the 2-Frobenius. In particular, when Q is H-invariant for some group H acting on U, and H does not fix a hyperplane in U, then Q is uniquely determined.*

Proof. We have that $(Q-Q')(u+v) = Q(u) + f(u,v) + Q(v) - (Q'(u) + f(u,v) + Q'(v)) = (Q-Q')(u) + (Q-Q')(v)$, so $Q - Q'$ is additive and hence semilinear with respect to the square map. The second claim follows from the fact that in this case $\ker(Q - Q') \le U$ is an H-invariant subspace of codimension at most one. $\square$

The classification. Write $g = \bigoplus_{\chi \text{ primary}} F(\chi)^{\oplus c_\chi}$ in generalized Jordan normal form. Consider a 'block' $U \le V$ such that $g|_U$ acts as $F(\chi)$ for some $\chi = i^e \in k[X]$ primary, with i irreducible of degree d and $e \ge 1$. This means that the action of $g|_U$ on U is isomorphic to multiplication by $\overline{X}$ in the k-algebra $R_\chi \coloneqq k[X]/(\chi) \cong U$. In particular, the g-invariant subspaces of U form a chain $0 = U_0 < U_1 < \cdots < U_e = U$, where U_j corresponds to the ideal $(i^{e-j}) \subseteq R_\chi$ ($j = 0, \ldots, e$). By finiteness, we can assume that e is as large as possible. Let $\lambda \in \overline{k}$ be a root of i and set $K_\chi \coloneqq k[\lambda] \cong k[X]/(i)$ to be the residue field of R_χ. For a polynomial $r = a_0 + a_1 X + \cdots + a_{k-1}X^{k-1} + X^k \in k[X]$ with $a_0 \neq 0$ write r^* for the *dual polynomial* $a_0^{-\sigma} X^{\deg(r)} r^\sigma(X^{-1})$ with 'normalized' reversed coefficients twisted by σ. Call the polynomial r *self-dual* if $r = r^*$, i.e., when its set of roots is preserved under inversion and an extension of σ, or equivalently, the map $\alpha\colon K_\chi \to K_\chi$ defined by $\alpha|_k = \sigma$ and $\overline{X} = \lambda \mapsto \lambda^{-1}$ extends (uniquely) to an automorphism of K_χ. Note that χ is self-dual if and only if i is self-dual. The restricted form $h \coloneqq f|_U$ induces a map $\varphi_h\colon U \to U^*$. Let $W \coloneqq \mathrm{rad}(h) = U \cap U^\perp$ be its radical. Then h descends to a non-singular form $\overline{h}$ on $\overline{U} \coloneqq U/W$. As W is clearly g-invariant, i.e., a submodule, and the set of submodules of U is a chain $0 = U_0 < U_1 < \cdots < U_e = U$, where g acts as $F(i^j)$ on U_j ($j = 0, \ldots, e$), when $W = U_j$, we have that g acts as $F(i^{e-j})$ on the quotient $\overline{U} = U/W$. The form $\overline{h}$ induces an isomorphism of the g-modules $\overline{U}$ and $\overline{U}^*$. Now g also acts as $F((i^*)^{e-j})$ on $\overline{U}^*$. Hence, if $j < e$, i.e., U is not totally singular, we need to have that $i = i^*$, i.e., i and so χ is self-dual.

(A): Hence, if U is non-singular and we are not in Case $(*)$ of Section 0.1(f), then $U = V$ and χ is self-dual, since $U \perp U^\perp = V$ is an orthogonal decomposition of V. If we are in the Case $(*)$ and U is non-singular, either $W = 0$ and we are in the previous situation, or $W = \mathrm{rad}(f)$ is one-dimensional and Q is non-trivial on W. This implies that i has degree one, so $i = X - \lambda$. Then for $v \in W \setminus \{0\}$ we have $Q(\lambda v) = \lambda^2 Q(v) = Q(v) \neq 0$ implying that $\lambda = 1$, i.e., g is unipotent. Then also $V = U$.

So assume that U is singular. Then W must contain the minimal submodule U_1 of U. Fix a vector $u_1 \in U_1 \setminus \{0\}$.

(B): First assume that $U_1 \cap \mathrm{rad}(f) = 0$ (which is always true unless we are in Case $(*)$ of Section 0.1(f) and $n = \dim(V)$ is odd). Then by assumption we find another block $U' \leq V$ (with respect to the decomposition associated to the above generalized Jordan normal form of g) and a vector $u' \in U'$ such that $f(u_1, u') \neq 0$. This implies that $f|_{U \times U'} : U \times U' \to k$ is separating in U, as $U \cap U'^\perp \leq U$ is g-invariant, but does not contain U_1. Hence we obtain an injective g-equivariant linear map $\varphi\colon U \to U'^*$. This implies that g acts as $F(\chi')$ on U', where $\chi' = (i^*)^{e'}$ for some $e' \geq e$. But by assumption on e, we must also have $e' \leq e$, implying that φ must be an isomorphism, i.e., $U' \cong U^*$, and $f|_{U \times U'}$ is also separating in U'. Then the submodule $U \perp U'$ is non-singular, so, since V was indecomposable, $V = U \perp U'$, as $V = (U \perp U') \perp (U \perp U')^\perp$ is an orthogonal decomposition of it.

(C): Now assume that $\mathrm{rad}(f)$ is one-dimensional and $\mathrm{rad}(f) \leq U$ and we are not in (A) (i.e., we are in Case $(*)$ of Section 0.1(f) and $n = \dim(V)$ is odd). As $\mathrm{rad}(f)$ is g-invariant, this implies $U_1 = \mathrm{rad}(f)$. As in (A) it follows that g is unipotent. But since we are not in (A), we must have $U_2 \leq W$. Fix a vector $u_2 \in U_2 \setminus U_1$. Similarly to the above, we find another block $U' \leq V$ such that f descends to a non-singular form on $U/U_1 \times U'$, so that

g acts as $F((X+1)^e)$ on U and as $F((X+1)^{e-1})$ on U'.

From this starting point we classify all cases that occur. We distinguish three cases.

Case 1: χ is not self-dual. Then we are in paragraph (B) from above and $U' \cong U^* \ncong U$, so that V must be orthogonally indecomposable, since the only direct sum decomposition of V is then $U \oplus U'$ and neither of the summands, by the above reason, carries a non-singular form. Also the pairing between U and $U' \cong U^*$ (as $k[X]$-modules) is canonical. Hence f is given by $f(u \oplus u', v \oplus v') = u'(v) + \varepsilon v'(u)^\sigma$, where $u', v' \in U'$ are interpreted as elements from U^*.

Remark 3.33. When we consider the isomorphisms $U \cong R_\chi$ and $U' \cong U^* \cong R_{\chi^*}$, we can write the pairing $f|_{U \times U'} : U \times U' \cong R_\chi \times R_{\chi^*} \to k$ more explicitly as

$$R_\chi \times R_{\chi^*} \ni (u,v) \mapsto \ell_\chi(uv^\alpha),$$

where $\alpha\colon R_{\chi^*} \to R_\chi$ is the isomorphism given by $\alpha|_k = \sigma$ and $\overline{X} \mapsto \overline{X}^{-1}$, and $\ell_\chi\colon R_\chi \to k$ is the form ℓ constructed in the proof of Lemma 3.34 below. This is completely analogous to the construction of f in Case 2 below (cf. the proof of Lemma 3.34), where $\chi = \chi^*$ and so $U = U' \cong R \coloneqq R_\chi = R_{\chi^*}$ and hence $\alpha \in \mathrm{Aut}(R)$.

Case 1: Orthogonal case in characteristic two.* In this case, we still have to determine Q. Since f vanishes on U and U', we must have that Q is semilinear with respect to the 2-Frobenius endomorphism on them. Hence the subspace $\{u \in U \mid Q(u) = 0\}$ resp. $\{u' \in U' \mid Q(u') = 0\}$ is g-invariant and has codimension at most one in U resp. U'. Thus Q must vanish on U and U' if $d > 1$, since then every proper submodule of U resp. U' has codimension at least d. If $d = 1$ we have that $g|_U \cong F((X-\lambda)^e) \cong J_e(\lambda)$. Let $u_1, \ldots, u_e$ be a basis in which the element g is represented by $J_e(\lambda)$. Q must vanish on U_{e-1}, so that $Q(u_e) = Q(u_e.g) = Q(\lambda u_e + u_{e-1}) = \lambda^2 Q(u_e) + \lambda f(u_e, u_{e-1}) + Q(u_{e-1}) = \lambda^2 Q(u_e)$, so that either $Q \equiv 0$ on U or $\lambda = 1$, the latter being a contradiction to the fact that i is not self-dual. The same holds on U'. Hence $Q\colon V \to k$ is given by $Q(u \oplus u') = f(u,u')$ as $Q(u) = Q(u') = 0$ for $u \in U$, $u' \in U'$, and one easily verifies that this quadratic form Q is g-invariant and induces f.

Case 2: $\chi = \chi^*$ is self-dual and $\lambda^2 \neq 1$ if f is bilinear. For this case we need a preparatory lemma (Lemma 3.34 below).

Preparation for Case 2. Write $K \coloneqq K_\chi = k[\lambda]$, $R \coloneqq R_\chi$, and let α resp. τ be the automorphism of R resp. K such that $\alpha|_k = \tau|_k = \sigma$ and $\alpha\colon \overline{X} \mapsto \overline{X}^{-1}$ resp. $\tau\colon \lambda \mapsto \lambda^{-1}$ (so that α induces τ on the residue field K of R). Note that both are involutions. Recall from Section 0.2 that R_α resp. K_τ denote the subring of R resp. the subfield of K whose elements are fixed by α resp. τ. Also define $\mathrm{tr}_\tau^{\pm 1}\colon K \to K$ to be the map $x \mapsto x \pm x^\tau$ and $\mathrm{N}_\alpha\colon R \to R_\alpha$ resp. $\mathrm{N}_\tau\colon K \to K_\tau$ to be the map $r \mapsto rr^\alpha$ resp. $x \mapsto xx^\tau$. Recall from above that the action of g on U is isomorphic to multiplication by $\overline{X}$ in $R \cong U$.

Lemma 3.34. *In this situation there is a natural bijection between the equivalence classes of non-singular alternating resp. symmetric bilinear resp. σ-conjugate-symmetric sesquilinear forms on $U \cong R$ and the factor group*

$$R_\alpha^\times / \mathrm{N}_\alpha(R^\times) \cong K_\tau^\times / \mathrm{N}_\tau(K^\times).$$

Proof. Let $\ell\colon R \to k$ be a k-linear form such that $f\colon (u,v) \mapsto \ell(uv^\alpha)$ is a non-singular σ-sesquilinear form. It is easy to see that this is equivalent to $(i^{e-1}) \not\subseteq \ker(\ell)$. Then f is g-invariant as $u\overline{X}(v\overline{X})^\alpha = u\overline{X}v^\alpha\overline{X}^{-1} = uv^\alpha$ for all $u, v \in R$ and any other such (possibly singular) form is given by $f_c(u,v) := f(u.c, v) = \ell(cuv^\alpha)$ for $c \in R$ by Lemmas 3.29 and 3.31 as $U \cong R$ is a cyclic module over R. Moreover, the map $c \mapsto f_c$ is easily seen to be bijective.

Now let us construct a form ℓ explicitly such that the corresponding form f is non-singular alternating resp. symmetric bilinear resp. σ-conjugate-symmetric sesquilinear.

At first assume that K/k is separable (which is always the case when k is finite). Define $S := K[Y]/(Y^e) = k[\lambda, Y]/(Y^e)$ and $\beta \in \mathrm{Aut}(S)$ via $\overline{Y} \mapsto -\overline{Y}$ and $\beta|_K = \tau$. Now let $s = \lambda + s_1\overline{Y} + \cdots + s_{e-1}\overline{Y}^{e-1} \in S$ with $s_1 \neq 0$. Then s is conjugate to $\lambda + \overline{Y}$ in S via the k-automorphism $\overline{Y} \mapsto s_1\overline{Y} + \cdots + s_{e-1}\overline{Y}^{e-1}$ of S. So assume subsequently, w.l.o.g., that $s = \lambda + \overline{Y}$. Consider the homomorphism $\varphi\colon k[X] \to S$ given by $X \mapsto s$. Then $\varphi(i) = i(s) = \sum_{j=0}^{d-1} (D^j i)(\lambda)\overline{Y}^j$, where $D^j i$ is the jth *Hasse derivative* of i, so that, since i is separable, $\varphi(i) = u_1\overline{Y} + \cdots u_{e-1}\overline{Y}^{e-1}$ with $u_1 \neq 0$. This shows that $\varphi(\chi) = \varphi(i)^e = 0$, showing that $(\chi) \subseteq \ker(\varphi)$. But it is easy to check that φ is surjective as $K = k[\lambda]$, so that for dimension reasons we must have equality, i.e., $k[X]/(\chi) = R \cong S$ as k-algebras via $\overline{\varphi}\colon \overline{X} \mapsto s$. Next, dropping the assumption $s = \lambda + \overline{Y}$, we find appropriate coefficients for s such that α corresponds to β via $\overline{\varphi}$. For this we have to solve the equations $ss^\beta = 1$ and $s_0 = \lambda$ for $s = \sum_{j=0}^{e-1} s_j\overline{Y}^j$ inductively. This means

$$\left(\sum_{j=0}^{e-1} s_j\overline{Y}^j\right)\left(\sum_{j=0}^{e-1} (-1)^j s_j^\tau \overline{Y}^j\right) = 1,$$

together with $s_0 = \lambda$. The coefficient of $\overline{Y}^0$ in the above product is always one, since $\tau\colon \lambda \mapsto \lambda^{-1}$. Computing the coefficient of $\overline{Y}^k$ for $k \geq 1$ gives

$$\sum_{j=0}^{k} (-1)^{k-j} s_j s_{k-j}^\tau = s_0^\tau s_k + (-1)^k s_0 s_k^\tau + \sum_{j=1}^{k-1} (-1)^{k-j} s_j s_{k-j}^\tau = 0.$$

Abbreviate the last term by t_k and observe that $t_k^\tau = (-1)^k t_k$. This means that

$$\mathrm{N}_\tau(s_0)\,\mathrm{tr}_\tau^{(-1)^k}(s_k/s_0) = \mathrm{tr}_\tau^{(-1)^k}(s_k/s_0) = -t_k.$$

But the condition on t_k means precisely that

$$t_k \in \ker(\mathrm{tr}_\tau^{(-1)^{k-1}}) = \mathrm{im}(\mathrm{tr}_\tau^{(-1)^k}),$$

so that the system has a solution. Hence we have found s such that $\overline{\varphi}$ carries α into β. We can now define the K-linear form $\Lambda\colon S \cong R \cong U \to K$ by $S \ni u_0 + u_1\overline{Y} + \cdots + u_{e-1}\overline{Y}^{e-1} \mapsto u_{e-1}$ and the τ-sesquilinear form $S^2 \to K$ by $(u,v) \mapsto \Lambda(uv^\beta)$, so that

$$\Lambda(uv^\beta) = \sum_{j=0}^{e-1} (-1)^{e-1-j} u_j v_{e-1-j}^\tau.$$

has sign $(-1)^{e-1}$. Hence choosing $\delta \in K^\times$ such that $\delta^\tau = (-1)^{e-1}\varepsilon\delta$, i.e., $0 \neq \delta \in \ker(\mathrm{tr}_\tau^{(-1)^e\varepsilon})$, and setting $F(u,v) := \delta\Lambda(uv^\beta)$ for $u,v \in S$ gives a τ-sesquilinear form of the desired sign ε, so that $\ell := \mathrm{tr}_{K/k} \circ \delta\Lambda$ induces an appropriate form f on $U \cong S$, when K/k is separable. Note that by construction $\delta\Lambda(u^\beta) = \varepsilon(\delta\Lambda(u))^\tau$ for $u \in S \cong U$ and, similarly, $\ell(u^\alpha) = \varepsilon\ell(u)^\sigma$ for $u \in R \cong U$.

In the inseparable case, we have $i = i'(X^{p^a})$ for a separable polynomial i', where p^a is the degree of inseparability of K/k. Set $K_{\mathrm{sep}} := k[\lambda^{p^a}]$. Then R contains $R_{\mathrm{sep}} := \langle \overline{X}^{p^a} \rangle_k \cong k[Z]/(i'^e(Z)) \cong K_{\mathrm{sep}}[Y]/(Y^e) = k[\lambda^{p^a}, Y]/(Y^e) =: S_{\mathrm{sep}}$. Then define $\mathrm{tr}_{R/R_{\mathrm{sep}}}\colon R \to R_{\mathrm{sep}}$ by $u_0 + u_1\overline{X} + \cdots + u_{p^a-1}\overline{X}^{p^a-1} \mapsto u_0$, where $u_i \in R_{\mathrm{sep}}$ $(i = 0,\ldots,p^a-1)$, and set $\ell := \ell_{\mathrm{sep}} \circ \mathrm{tr}_{R/R_{\mathrm{sep}}}$, where $\ell_{\mathrm{sep}}\colon R_{\mathrm{sep}} \to k$ has already been defined in the separable case. Then ℓ induces a non-singular form f of the desired type on $U \cong R$, since for $0 \neq u = u_0 + u_1\overline{X} + \cdots + u_{p^a-1}\overline{X}^{p^a-1} \in R \cong U$ with $u_j \in R_{\mathrm{sep}}$ $(j = 0,\ldots,p^a-1)$ there is a $u_k \neq 0$. Then $\mathrm{tr}_{R/R_{\mathrm{sep}}}(u(\overline{X}^k)^\alpha) = u_k$, so that $f(u, \overline{X}^k v) \neq 0$ for appropriate $v \in R_{\mathrm{sep}} \subseteq R \cong U$. Also note that $\mathrm{tr}_{R/R_{\mathrm{sep}}}$ commutes with α (which stabilizes R_{sep}), so that still $\ell(u^\alpha) = \varepsilon\ell(u)^\sigma$ for all $u \in R \cong U$.

Finally, we have to prove the claimed correspondence. We already mentioned that $c \mapsto f_c$ is a bijection. Assume that f was already alternating resp. symmetric bilinear resp. σ-conjugate-symmetric sesquilinear. That f_c has the same property means now that $\ell(cuu^\alpha) = 0$ implying that $\ell(cu1) + \ell(c1u^\alpha) = \ell(c(u+u^\alpha)) = 0$ resp. $\ell(c(u-u^\alpha)) = 0$ resp. $\ell(cu) - \ell(cu^\alpha)^\sigma = 0$ for all $u \in R \cong U$.

In the first two cases, we see immediately that all $c \in R_\alpha$ satisfy the assumption. But since f is non-singular and $\dim_k(R_\alpha) = \dim_k(R)/2$, there are no other such c. In the last case, we can use the property of ℓ that $\ell(u)^\sigma = \ell(u^\alpha)$, so that the condition becomes $\ell((c - c^\alpha)u) = 0$, which means $c \in R_\alpha$ as well.

Changing the coordinates g-invariantly of a form $f(u,v) = \ell(uv^\alpha)$ means to multiply both entries by some $c \in R$. Hence we get a natural one-to-one correspondence between $R_\alpha^\times / N_\alpha(R^\times)$ and the non-singular forms f (in each of the three cases).

It remains to show that $R_\alpha^\times / \mathrm{N}_\alpha(R^\times) \cong K_\tau^\times / \mathrm{N}_\tau(K^\times)$. For this purpose consider the natural homomorphism of groups $\pi\colon R_\alpha^\times \to K_\tau^\times / \mathrm{N}_\tau(K^\times)$. That $u \in \ker(\pi)$ means that there is $v \in R^\times$ such that $u \equiv vv^\alpha$ modulo (i). Now let $1 \leq k \leq e-1$ and assume we have found $v \in R^\times$ such that $u_k \coloneqq u - \mathrm{N}_\alpha(v) \in (i^k)_\alpha$. We want to find $w \in (i^k)$ such that $u_{k+1} = u - \mathrm{N}_\alpha(v+w) \in (i^{k+1})_\alpha$. This means $u_k - v^\alpha w - vw^\alpha = u_k - \mathrm{N}_\alpha(v)\,\mathrm{tr}_\alpha(w/v) \in (i^{k+1})_\alpha$. But it is easy to check that α acts on $K = R/(i)$ as it acts on $(i^k)/(i^{k+1})$. Therefore, since $u_k \in (i^k)_\alpha$ and $\mathrm{N}_\alpha(v) \in R_\alpha^\times$, the system can be solved. Thus we find $v \in R^\times$ such that $u = \mathrm{N}_\alpha(v)$ and the proof is complete. □

Remark 3.35. Using for ℓ the trace $\mathrm{tr}_{R/k}$ is not a good choice, since this trace is zero if i is inseparable.

Remark 3.36. The proof shows that, if i is separable, all possible non-singular forms f are of the form $(u,v) \mapsto \mathrm{tr}_{K/k} \circ \delta\Lambda(uv^\beta)$ for suitable $0 \neq \delta \in \ker(\mathrm{tr}_\tau^{(-1)^e\varepsilon})$ as above. This is, since then $K_\tau^\times \subseteq R_\alpha^\times$ surjects on $K_\tau^\times / \mathrm{N}_\tau(K^\times)$ under the map π in the proof and δ is determined up to a factor in $K_\tau^\times$.

Now we apply Lemma 3.34.

Discussion of Case 2. We are either in paragraph (A) or (B) from above. We first show that (B) cannot occur. Assume we are in paragraph (B), i.e., $V = U \perp U'$ and $U \cong U' \cong U^*$. Then Lemma 3.34 gives that

$$f = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} c & 1 \\ \varepsilon & d \end{pmatrix}$$

for some $c, d \in R_\alpha$. By this notation, which we shall use also in the subsequent cases, we mean that $f(u_1 \oplus u_2, v_1 \oplus v_2) = \ell(u_1 f_{11} v_1 + u_1 f_{12} v_2 + u_2 f_{21} v_1 + u_2 f_{22} u v_2)$, where $\ell \colon R \cong U \to k$ is the linear form from the proof of Lemma 3.34. If c resp. d was a unit in R, we could split off U resp. U'. So assume $c, d \in (i)$. Then

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c & 1 \\ \varepsilon & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a^\alpha & 1 \end{pmatrix} = \begin{pmatrix} c + \varepsilon a + a^\alpha + a a^\alpha d & * \\ * & * \end{pmatrix},$$

so that taking a such that $K \ni \overline{a} \notin \ker(\mathrm{tr}_\tau^\varepsilon)$ gives a transformation such that $\{u \oplus ua \mid u \in U \cong R\}$ splits off. Here we used that α is non-trivial, since $\lambda \neq \lambda^{-1}$. Hence $U = V$, so we are actually in paragraph (A), and f is determined up to $K_\tau^\times / \mathrm{N}_\tau(K^\times)$ as described in Lemma 3.34 (so uniquely when k is finite, since then the norm $\mathrm{N}_\tau \colon K^\times \to K_\tau^\times$ is surjective).

Case 2: The orthogonal case in characteristic two.* In characteristic two we still need to determine Q.

At first assume that K/k is separable (which is certainly true if k is finite). In this case, from the proof of Lemma 3.34 and Remark 3.36 we get

$$f(u,v) = \mathrm{tr}_{K/k}\left(\delta \sum_{j=0}^{e-1} u_j v_{e-1-j}^\tau\right),$$

as $1 = -1$. Noting that $\operatorname{tr}_{K/k} = \operatorname{tr}_{K_\tau/k} \circ \operatorname{tr}_{K/K_\tau}$ and that

$$\begin{aligned}\operatorname{tr}_{K/K_\tau}\left(\delta \sum_{j=0}^{e-1} u_j v_{e-1-j}^\tau\right) &= \delta \sum_{j=0}^{e-1} (u_j v_{e-1-j}^\tau + u_j^\tau v_{e-1-j}) \\ &= H(u+v) - H(u) - H(v),\end{aligned}$$

where

$$H(u) \coloneqq \delta\Lambda(uu^\beta) = \delta \sum_{j=0}^{e-1} u_j u_{e-1-j}^\tau$$

is a g-invariant Hermitian form over K which induces the previous K_τ-bilinear form. Setting $Q \coloneqq \operatorname{tr}_{K_\tau/k} \circ H$ gives an appropriate quadratic form Q inducing f.

If K/k is inseparable and p^a is the degree of inseparability, then f is given by $f(u,v) = \sum_{i=0}^{p^a-1} f_{\mathrm{sep}}(u_i, v_i)$, as in the proof of Lemma 3.34, where $u = \sum_{i=0}^{p^a-1} u_i \overline{X}^i$, $v = \sum_{i=0}^{p^a-1} v_i \overline{X}^i$ with $u_i, v_i \in R_{\mathrm{sep}}$ $(i = 0, \ldots, p^a - 1)$, so that we can define Q by $Q(u) \coloneqq \sum_{i=0}^{p^a-1} Q_{\mathrm{sep}}(u_i)$, where Q_{sep} is the quadratic form we defined in the separable case.

It remains to prove that Q is uniquely determined. So assume Q' would be another g-invariant quadratic form inducing f. Arguing as in Case 1* for $Q - Q'$ instead of Q, we obtain that, if $Q - Q' \not\equiv 0$, we must have that g is unipotent, i.e., $\lambda = 1$, a contradiction. Thus Q is unique.

Case 3: $i = X \pm 1$ and f is bilinear. We can restrict to $i = X - 1$ by multiplying g by $-\operatorname{id}$ (which stabilizes f) if necessary.

Case 3.1: p is odd. Write $R \coloneqq R_\chi = k[X]/((X-1)^e)$ and $S \coloneqq k[Y]/(Y^e)$ as in Case 2 and note that $R \cong S$ via $\overline{X} \mapsto 1 + \overline{Y}$. Define α as there and set $R_{\pm 1} \coloneqq \operatorname{im}(\operatorname{tr}_\alpha^{\pm 1}) = \ker(\operatorname{tr}_\alpha^{\mp 1})$ to be the ± 1-eigenspace of $\alpha \colon R_\chi \to R_\chi$. In analogy to Lemma 3.34 we have the following.

Lemma 3.37. *In this situation there is a natural bijection between the equivalence classes of g-invariant alternating resp. symmetric bilinear forms on U and the quotient*

$$R_{(-1)^{e-1}} / \operatorname{N}_\alpha(R^\times) \text{ resp. } R_{(-1)^e} / \operatorname{N}_\alpha(R^\times).$$

Hence equivalence classes of g-invariant non-singular alternating resp. symmetric bilinear forms are in bijection with $k^\times/(k^\times)^2$, as $R_1^\times / \operatorname{N}_\alpha(R^\times) \cong k^\times/(k^\times)^2$, (which equals C_2 if k is finite) in the case e is even resp. odd, whereas all g-invariant alternating resp. symmetric forms on U in odd resp. even dimension are singular.

Proof. Find $s \in S$ such that $ss^\beta = 1$ and $a_0 = 1$ so that $\overline{X} \mapsto s$ defines an isomorphism $R \cong S$ carrying α to β as in the proof of Lemma 3.34 (here one uses that $2 \neq 0$). Define Λ as in the proof of Lemma 3.34 with $\tau = \operatorname{id}$, so that $f \colon S^2 \to k$; $(u,v) \mapsto \Lambda(uv^\alpha)$ has sign $(-1)^{e-1}$. Then by Lemma 3.31 any other possible form is given by f_c. Now f_c has sign $(-1)^{e-1}$ if and only if $c \in \operatorname{im}(\operatorname{tr}_\alpha) = R_1$ and sign $(-1)^e$ if $c \in \operatorname{im}(\operatorname{tr}_\alpha^{(-1)}) = R_{-1}$. $\square$

According to Lemma 3.37 we have the following two cases.

Case 3.1(a): $\varepsilon = (-1)^{e-1}$. Then we have that

$$f = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} c & 1 \\ 1 & d \end{pmatrix},$$

so that we can use the same trick as in Case 2 to obtain $U = V$. Indeed, since V is orthogonally indecomposable, $c, d \in (i)$, so that, if we apply the same coordinate transformation as there, we get

$$f' = \begin{pmatrix} f'_{11} & f'_{12} \\ f'_{21} & f'_{22} \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c & 1 \\ 1 & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a^\alpha & 1 \end{pmatrix} = \begin{pmatrix} c + a + a^\alpha + aa^\alpha d & * \\ * & * \end{pmatrix}.$$

Taking f'_{11} modulo (i) gives $2a_0 \in k$, where

$$a = \sum_{j=0}^{e-1} a_j \overline{Y}^j,$$

so that, taking $a_0 \neq 0$, we have found the g-invariant non-singular subspace $\{u \oplus ua \,|\, u \in U \cong R\}$, contradicting the assumption.

Remark 3.38. Hence, if $k = \mathbb{F}_q$ is finite, we get the two inequivalent possibilities that $f(u,v) = \ell(uv^\alpha)$ and $f(u,v) = \beta\ell(uv^\alpha)$ for $\beta \in k^\times$ a non-square. These correspond to $V_1(2k)$ and $V_\beta(2k)$ resp. $V_1(2k+1)$ and $V_\beta(2k+1)$ of Section 2.4 resp. 2.5 of [25]. Note that $V_\beta(l)^{\perp 2} = V_1(l)^{\perp 2}$ (use the trick in Section 3.4.6 of [78]). Hence, if g is isotypic of type $F((X \pm 1)^e)^{\oplus c}$, where $\varepsilon = (-1)^{e-1}$, V is either of type $V_1(e)^{\perp c}$ or $V_\beta(e) \perp V_1(e)^{\perp(c-1)}$ (cf. Propositions 2.3 and 2.4 of [25]).

Case 3.1(b): $\varepsilon = (-1)^e$. In this case, $V = U \oplus U'$ and due to Lemma 3.37 there is no non-singular g-invariant form on $U \cong U'$. We show that we can change the coordinates such that U and U' are totally singular. For this define $\deg(c) := \max\{j \in \{0, \dots, e\} \,|\, c \in (i^j)\}$. Write

$$f = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} a & b \\ -b^\alpha & d \end{pmatrix}$$

and assume, w.l.o.g., that $\deg(a) \le \deg(d)$ and $b \in 1 + (i)$ (at the start we may even assume that $b = 1$). Again we obtain

$$f' = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ -b^\alpha & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x^\alpha & 1 \end{pmatrix} = \begin{pmatrix} a - b^\alpha x + bx^\alpha + dxx^\alpha & b + dx \\ -(b+dx)^\alpha & d \end{pmatrix}.$$

We have that $R_{-1} = \mathrm{im}(\mathrm{tr}_\alpha^{-1}) = \mathrm{tr}_\alpha^{-1}(R_{-1})$, so that we find $x \in bR_{-1}$ such that $a - b^\alpha x + bx^\alpha = a - \mathrm{N}_\alpha(b)\,\mathrm{tr}_\alpha^{-1}(x/b) = 0$. Then the new entry $a' := f'_{11}$ has degree $\deg(dxx^\alpha) \ge \min\{e, \deg(d) + 2\}$ and still $f'_{12} = b' \in 1 + (i)$. Hence, repeating this procedure, eventually we obtain $a = c = 0$, so that in suitable coordinates $f_{12} = f_{21} = 0$.

Remark 3.39. If $k = \mathbb{F}_q$ is finite, in this subcase, we only get the representative $W(2l+1)$ resp. $W(2l)$ of Section 2.4 resp. 2.5 of [25] in the symplectic resp. orthogonal case.

Case 3.2: $p = 2$. In this case, to give a reasonable classification, we need some assumption on the field k. The case that k is quadratically closed, i.e., every quadratic polynomial has a solution, is considered in [33].

In the case that $k = \mathbb{F}_q$ is finite, we refer to Chapter 3 of [25], which shows that there are four types of indecomposables (namely $W(m)$, $W_\alpha(m)$, $V(2k)$, and $V_\alpha(2k)$, which satisfy certain relations given in Table 1 of [25]). Letting the transformation g act $F((X+1)^e)$-isotypically and choosing coordinates such that f (which is preserved by g) is in the normal form of [25, Theorem 3.1], we see that f restricted to all but constantly many Jordan blocks of g is of the form $W(e)$.

§2 The Frobenius normal form for elements $g \in G$

We wish to apply the same method as in Subsection 3.4.1, for which we need an analog of the Frobenius normal form for elements $g \in G$.

Write $g = h \perp u = h \perp u_1 \perp u_{-1}$, where $u_1, -u_{-1}$ are unipotent and h has only eigenvalues different from ± 1. This is possible by considering the Cases 1, 2, and 3 of indecomposables in the previous subsubsection.

We obtain a normal form for h in the same way as the Frobenius normal form is obtained from the generalized Jordan normal form: In the first summand we collect all orthogonally indecomposable summands from Case 1 resp. Case 2 of the form $F((ii^*)^e) = F(i^e) \oplus F(i^{*e})$ with $i \in k[X]$ irreducible and not self-dual resp. $F(i^e)$ with $i \in k[X]$ irreducible and self-dual (and $i \neq X \pm 1$) and e as large as possible. Then we split off this summand and proceed in the same way with the perpendicular complement.

For u_1 and u_{-1} we use the normal form provided by [25, Propositions 2.2, 2.3, 2.4, and Theorem 3.1].

We still need the following fact, which follows from the analysis of Cases 1, 2, and 3 of the previous subsubsection:

Fact 3.40. *Whenever $\chi \in k[X]$ is self-dual and is not divided by $X \pm 1$ in the bilinear case, then there exists a non-singular form f (coming from a quadratic form Q when $p = 2$) which is preserved by $F(\chi)$ (f is even unique up to linear equivalence). On the other hand, for all e we have that there is a form f (together with Q when $p = 2$) preserving $F((X \pm 1)^e)^{\oplus 2}$.*

§3 Proof of Theorem 3.2 for the remaining groups

We decompose $g \cong h \perp u_1 \perp u_{-1}$ as described above. Hence, using Jensen's inequality, we only need to consider two cases: (a) $g = h$ and (b) $g = \pm u$, where u is unipotent. Now we can apply the previous considerations to elements that are $F(\chi)$-isotypic for $\chi \in k[X]$ self-dual of degree k which is not divided by $X \pm 1$ in Case (a), and elements that are

$F((X \pm 1)^e)^{\oplus d}$-isotypic, where $d = 1$ or 2 and $k = de$ in Case (b). It is enough to consider these two isotypic cases. Again we derive an estimate for k small and k large.

Estimate for small k. In Case (a), we have that $g \cong F(\chi)^{c_\chi}$ for χ self-dual and there is up to equivalence only one form f preserved by $F(\chi)$ (which follows from the first part of Fact 3.40 above). We can approximate the linear map g by elements from $w(\langle \overline{X} \rangle \wr \mathrm{Sym}(r))$, where $\langle \overline{X} \rangle \subseteq R^\times = (k[X]/(\chi(X^c)))^\times$ and $c_\chi = rc + s$ for $0 \le s < r$, as in the estimate for small k in Subsection 3.4.1. But $\overline{X} \in R^\times$ preserves a non-singular form f as $\chi(X^c)$ is again self-dual (which again follows from Fact 3.40 above), so that also the group $\langle \overline{X} \rangle$ preserves such a form and we are done by uniqueness.

In Case (b), we have $g \cong F((X \pm 1)^e)^{\oplus d c_{d,e}}$. Since $k = de$ is small, $c_{d,e}$ is large and we can certainly assume it to be even. Observe that $F((X \pm 1)^e)^{\oplus 2}$ always supports a non-singular form, so that $F((X^c \pm 1)^e)^{\oplus 2}$ will do as well (by Fact 3.40). Hence we can use the same trick as in Case (a) and use Fact 3.40 in the unitary case and [25, Propositions 2.3, 2.4, and Theorem 3.1] in the bilinear case, which says that the form f is essentially determined by g up to a constant number of summands $F((X \pm 1)^e)^{\oplus d}$ (namely, with the notation used there, most of its blocks will be $U(e)$ in the unitary case and $V_1(e)$ or $W(e)$ in the bilinear case).

Estimate for large k. In this case, we assume that $g \cong F(\chi)$ in Case (a) (so $c_\chi = 1$) and $g \cong F((X \pm 1)^e)^{\oplus d}$ for $d = 1$ or 2. Here we want to apply the following simple fact.

Lemma 3.41. *Let $C > 0$ be a fixed constant. Assume that $V = X \oplus Y \oplus Z$, where X and Y are totally isotropic, $n - 2\dim(X), n - 2\dim(Y) \le C$, i.e., X and Y are close to a Witt subspace, and $\mathrm{codim}_X(X \cap X.g), \mathrm{codim}_Y(Y \cap Y.g) \le C$, i.e., X and Y are almost g-invariant. Then g can be approximated by word values.*

Proof. Note that $\dim(Y^\perp) = n - \dim(Y) \le \frac{n+C}{2}$, so that $\dim(X \cap Y^\perp) \le C$. Hence we can find $X' \le X$ of dimension at least $\dim(X) - C \ge \frac{n-3C}{2}$ such that $f|_{X' \times Y}$ is separating in X'. Hence, choosing $Y' \le Y$ which induces all σ-semilinear functionals X'^*, we can assume by passing from X to X' and Y to Y' that $f|_{X \times Y}$ is non-degenerate, so in particular $\dim(X) = \dim(Y)$.

Now let g' be an extension of $g|_{X \cap X.g^{-1}} : X \cap X.g^{-1} \to X \cap X.g$ to an invertible linear map $X \to X$. By Subsection 3.4.1, we find $h = w(x, y) \in w(\mathrm{GL}(X))$ such that $d_{\mathrm{rk},X}(g', h) \le d(1/\dim(X)) \le d(\frac{2}{n-C})$.

We extend h to all of V as follows: Extend $x, y \in \mathrm{GL}(X)$ to Y by taking their dual on Y, so that they fix X and Y setwise, and then extend them to V with Witt's lemma. Then set $h := w(x, y)$ on all of V. Write $Y.(g - h) = (Y.(g - h) \cap Y) \oplus W$. Then, since h fixes Y, W is injectively mapped by the natural map $Y.(g - h) \to (Y + Y.g)/Y$, but the last quotient, by assumption, had dimension at most C, so that $\dim(W) \le C$. Now $f(x.(g^{-1} - h^{-1}), y) = f(x, y.(g - h)) = 0$ for $x \in X$, $y \in Y$, when $x \in \ker(g^{-1} - h^{-1}) = (\ker(g - h)).h$ as $g^{-1} - h^{-1} = h^{-1}(h - g)g^{-1}$. But the vector space of all such $x \in X$ has dimension at least $\dim(X)(1 - d(1/\dim(X)) - C/\dim(X))$, which follows from the above estimate on $d_{\mathrm{rk},X}(g', h)$ and the fact that g and g' agree on $X \cap X.g^{-1}$.

Hence the dimension of $Y.(g-h)\cap Y$ is at most $\dim(X)(d(1/\dim(X))+C/\dim(X)))$, so that, using $\dim(W)\leq C$, the dimension of $Y.(g-h)$ is at most $\dim(X)(d(1/\dim(X))+2C/\dim(X))$. Hence the rank of $g-h$ is small on X and Y, so is small on V. This ends the proof. $\square$

Now V is the direct sum of orthogonally indecomposable modules, each type of which occurs at most once. Subsequently, we construct subspaces X, Y, and Z with the property required by Lemma 3.41. Write $\mathcal{S}_j$ $(j=1,2,3)$ for the orthogonally indecomposable summands of V described in Case j from above.

For each orthogonally indecomposable summand $S=U\oplus U^*\in\mathcal{S}_1$ of V as in Case 1 set $X_S\coloneqq U$, $Y_S\coloneqq U^*$, and $Z_S\coloneqq 0$. Then define $X_1\coloneqq\bigoplus_{S\in\mathcal{S}_1}X_S$, $Y_1\coloneqq\bigoplus_{S\in\mathcal{S}_1}Y_S$, and $Z_1\coloneqq 0$. Define χ_1 by the fact that all the summands from Case 1 grouped together act as $F(\chi_1)$.

In Case 2, for each $S=U\in\mathcal{S}_2$ we have that g acts as $F(\chi_S)=F(i_S^{e_S})$ on S, where i_S is irreducible of degree d_S and χ_S is of degree $k_S=d_Se_S$. Set $\chi_2\coloneqq\prod_{S\in\mathcal{S}_2}\chi_S\in k[X]$ and set $k_2\coloneqq\deg(\chi_2)$. The form f on $\bigoplus\mathcal{S}_2$ is given by $(u,v)\mapsto\ell(uv^\alpha)=\sum_{S\in\mathcal{S}_2}\ell_S(u_Sv_S^{\alpha_S})$, where $u=(u_S)_{S\in\mathcal{S}_2}$, $v=(v_S)_{S\in\mathcal{S}_2}$, and ℓ_S,α_S $(S\in\mathcal{S}_2)$ are as described in Case 2 above. Set $\overline{X}\coloneqq(\overline{X}_S)_{S\in\mathcal{S}_2}\in R\coloneqq\prod_{S\in\mathcal{S}_2}R_S=\prod_{S\in\mathcal{S}_2}k[X_S]/(\chi_S)$, $\alpha\coloneqq(\alpha_S)_{S\in\mathcal{S}_2}$, and recall that g acts on $V\cong R$ as multiplication by $\overline{X}$. Also recall that $R_\alpha=\{r\in R\,|\,r^\alpha=r\}$.

That the vectors $v,\dots,v.g^{l-1}$ for $v\in\bigoplus\mathcal{S}_2$ span a totally singular subspace hence means that $\ell(v\overline{X}^jv^\alpha)=\ell(\mathrm{N}_\alpha(v)\overline{X}^j)=0$ for $j=0,\dots,l-1$, where we write $\mathrm{N}_\alpha(v)=vv^\alpha$. Write $u=\mathrm{N}_\alpha(v)$. We demonstrate how to find such a vector v for which the vectors $v,\dots,v.g^{l-1}$ are linearly independent in the orthogonal case for $p=\mathrm{char}(k)\neq 2$ (i.e., $\sigma=\mathrm{id}$ and $\varepsilon=1$). The other two cases are similar. So assume that f is non-singular symmetric bilinear. Let $\lambda_S\in\overline{k}$ be a root of i_S. Define i'_S to be the minimal polynomial over k of $\lambda_S+\lambda_S^{-1}$ and let $\chi'_S\in k[X]$ be the minimal polynomial of $\overline{X}_S+\overline{X}_S^{-1}\in R_S$ $(S\in\mathcal{S}_2)$. Note that $i'_S(X+X^{-1})X^{d_S/2}=i_S$ and $\chi'_S(X+X^{-1})X^{k_S/2}=\chi_S$, so that $\deg(i'_S)=\deg(i_S)/2$ and $\deg(\chi'_S)=\deg(\chi_S)/2$. This holds as $\lambda_S\neq\lambda_S^{-1}$ are conjugate in $K_S=k[X]/(i_S)$, so that $2\mid d_S$ $(S\in\mathcal{S}_2)$. Hence the i'_S resp. χ'_S are pairwise coprime, since the i_S are coprime $(S\in\mathcal{S}_2)$. Define χ'_2 by $\chi'_2\coloneqq\prod_{S\in\mathcal{S}_2}\chi'_S\in k[X]$, so that $\deg(\chi'_2)=k_2/2$. Set $l\coloneqq\deg(\chi'_2)-1$. Note that R_α is a k-subalgebra of R of k-dimension $l+1=\deg(\chi'_2)$ and ℓ descends to a k-linear functional $R_\alpha\to k$. The minimal polynomial of $\overline{X}+\overline{X}^{-1}\in R_\alpha$ over k is χ'_2, so that $R_\alpha=k[\overline{X}+\overline{X}^{-1}]\cong k[X]/(\chi'_2)$ (here we used that the χ'_S are pairwise coprime for $S\in\mathcal{S}_2$). This implies that the $(\overline{X}+\overline{X}^{-1})^j$ and hence the $\overline{X}^j+\overline{X}^{-j}$ $(j=0,\dots,l-1)$ span an l-dimensional k-subspace of R_α and are hence linearly independent.

Now note that $\ell(\overline{X}^ju)=0$ is equivalent to $\ell(\overline{X}^{-j}u)=0$ when $u\in R_\alpha$, as ℓ has the property $\ell(x)^\sigma=\ell(x)=\varepsilon\ell(x^\alpha)=\ell(x^\alpha)$ and $\overline{X}^\alpha=\overline{X}^{-1}$ (as we are in the orthogonal case). But for $j\leq l-1$, $\overline{X}^j+\overline{X}^{-j}\neq 0$ from the previous observation, so that the two preceding equations are equivalent to $\ell((\overline{X}^j+\overline{X}^{-j})u)=0$. Now from the construction of ℓ one sees that f restricts to R_α as a non-singular form (here we use that we are in the orthogonal case and $p\neq 2$). Hence $W\coloneqq\langle(\overline{X}+\overline{X}^{-1})^j\,|\,j=0,\dots,l-1\rangle^\perp\cap R_\alpha$ is one-dimensional. Let $0 \neq u \in W$ be a generator of this subspace. We show that u is a unit in $R_\alpha \subseteq R$. Assume the contrary, namely that $u_S \in (i_S)$ for some $S \in \mathcal{S}_2$. Then the k-linear functional $\ell_u \colon R_\alpha \to k$ which is zero on all $R_{S',\alpha_{S'}}$ $(S \neq S' \in \mathcal{S}_2)$ and which equals $x \mapsto \ell(r i_S'^{e_S-1}(\overline{X}_S + \overline{X}_S^{-1})x)$ for $r \in R^\times_{S,\alpha_S}$ arbitrary must be a linear combination of the functionals $R_\alpha \to k$; $x \mapsto \ell((\overline{X} + \overline{X}^{-1})^j x)$ $(j = 0, \dots, l-1)$, since $\ell_u(u) = 0$ and the latter functionals span the space of all such functionals. This means that there is a polynomial $s \in k[X]$ of degree $l - 1 = \deg(\chi_2') - 2$ such that $s(\overline{X} + \overline{X}^{-1})$ is zero on $R_{S',\alpha_{S'}}$ for $S' \neq S$ and lies in $(i_S'^{e_S-1}(\overline{X}_S + \overline{X}_S^{-1})) = (i_S^{e_S-1}) \cap R_{S,\alpha_S}$ on R_{S,α_S}. This means that $\chi'_{S'} \mid s$ for $S' \neq S$ and $i_S'^{e_S-1} \mid s$. Hence, since the polynomials $i_T' \in k[X]$ $(T \in \mathcal{S}_2)$ are irreducible and pairwise coprime, to achieve an arbitrary r, we hence need that $s = (\chi_2'/i_S')s_0$, where $s_0 \in k[X]$ is arbitrary of degree less than $d_S/2$. Hence we would need in the worst case that $\deg(s) = l = \deg(\chi_2') - 1$, which is a contradiction. So we have that u is a unit in R_α and every such unit is in the image of the norm $\mathrm{N}_\alpha \colon R \to R_\alpha$, since k is a finite field, so that we find an appropriate vector $v \in R$ such that $\mathrm{N}_\alpha(v) = u$. Also, since v is a unit in R, the space $X_2 \coloneqq \langle v.g^j \mid j = 0, \dots, l-1 = \deg(\chi_2') - 2\rangle$ is actually of dimension l. Hence, setting $Y_2 \coloneqq X_2.g^l$ and choosing Z_3 appropriately, we are done in this case.

Now, in Case (a), we have that g acts as $F(\chi) = F(\chi_1\chi_2)$. Setting $X \coloneqq X_1 \oplus X_2$, $Y \coloneqq Y_1 \oplus Y_2$, and $Z \coloneqq Z_1 \oplus Z_2$ (and cutting of a further dimension if necessary when $p = 2$ by restricting to $Q = 0$), we can apply Lemma 3.41 to $F(\chi)$.

In Case (b), when $S \in \mathcal{S}_3$, one can easily extract almost invariant isotropic subspaces X_S, Y_S, and a space Z_S as needed in Lemma 3.41 from the explicit representations given in [25, Propositions 2.2, 2.3, 2.4, and Theorem 3.1] and sum them up as in Case (a).

Final proof. The final proof is now identical with the one given in the last two paragraphs of Subsection 3.4.1.

3.4.3 An alternative way of proving Theorem 3.1 using wreath products

The techniques of Section 3.3 and 3.4 provide an alternative proof of Theorem 3.1. The reason for this is the following lemma on isotypic elements.

Lemma 3.42. *Let $\sigma \in \mathrm{S}_n$ be k-isotypic of type (k^{c_k}). Then the following hold.*

(i) There exists a prime $p_0(w)$ such that, if all prime divisors of k are greater than $p_0(w)$, then, if $c_k \geq N(w)$, the element σ lies in the word image $w(\mathrm{S}_n)$.

(ii) There are constants $N(w)$, $m(w)$ such that also, if k has prime divisors $\leq p_0(w)$, for $m(w) \mid c_k$ and $c_k \geq N(w)m(w)$ the permutation σ lies in $w(\mathrm{S}_n)$. In particular, any permutation of type (k^{c_k}) can be approximated up to distance $m(w)/c_k$ by word values if $c_k \geq N(w)m(w)$.

(iii) If $c_k = 1$, we can approximate σ up to error $C(w)/n$ for a fixed constant $C(w) > 0$ (which is Theorem 1.3 of [44] without fixed constant).

Proof. Use the setting of Section 3.3 together with Corollary 3.26 of Section 3.4 with coefficient group the cyclic group C_k, i.e., look at the wreath product $\mathrm{C}_k \wr \mathrm{S}_{c_k}$.

For (i) we find the finite p-groups $H(p)$ such that $d^2(\pi(p))$ is non-trivial. Then, as $H(p)$ acts transitively on the coordinates of $C_2(\pi(p))$, the image $\mathrm{im}(d^2(\pi(p)))$ contains a vector $x(p) \in C^2(\pi(p)) \cong \mathbb{R}[H(p)]$ with all coefficients non-zero and integral (otherwise it would be contained in the union of the coordinate hyperplanes, which is impossible as $\mathbb{Z} \subseteq \mathbb{R}$ is infinite). Then, using coefficients in C_k, we get that $w(\mathrm{C}_k \wr \mathrm{S}_{p^h})$ contains an element of cycle type (k^{p^h}), when k is not divisible by a prime dividing one of the coefficients of $x(p)$. Now, if we use for p two distinct primes p_1, p_2, we can write every $c_k \geq N(w) := (p_1^h - 1)(p_2^h - 1)$ as a sum of numbers p_1^h, p_2^h, so that, taking disjoint unions of the previous construction, every σ of type (k^{c_k}) with $c_k \geq N(w)$ lies in the word image $w(\mathrm{S}_n)$.

To obtain (ii), use the same idea as in (i), but instead of using coefficients in C_k, take coefficients in C_{mk}, where $m = m(w)$ is defined to be the least common multiple of all coefficients of $x(p_1)$ and $x(p_2)$ from above. Doing the same construction, we obtain that any element σ of type (k^{c_k}) with $m \mid c_k$ and $c_k/m \geq N(w)$ lies in $w(\mathrm{C}_{mk} \wr \mathrm{S}_{c_k/m}) \subseteq w(\mathrm{S}_n)$. For the second part of (ii) write $c_k = rm + s$, where $0 \leq s < m$, and use the former construction on S_{mkr} and fill up by the identity.

For (iii) just take a fixed group $H(p)$ where $d^2(\pi(p))$ is non-trivial. Then pick an element with only non-zero coefficients as in (i). Plugging in the coefficient group C_r, we see that we differ in boundedly many points from a k-cycle, where $k = n = rp^h$. For arbitrary n write $n = rp^h + s$, where $0 \leq s < p^h$, use the former construction on rp^h points and fill up with the identity. □

Remark 3.43. We conjecture that $m = 1$, if k is odd, and $m = 2$ if k is even, can be taken (which would be the best possible outcome, as, when $w \in \mathbf{F}_2'$, its image $w(\mathrm{S}_n)$ lies in A_n).

The estimate for arbitrary elements follows as at the end of Subsection 3.2.4.

Chapter 4

Isomorphism questions for metric ultraproducts

4.1 Introduction

In [73, ?] Thom and Wilson discussed various properties of metric ultraproducts of finite simple groups. In particular, they asked the question which such ultraproducts can be isomorphic. In Theorem 2.2 of [73], a metric ultraproduct of alternating groups is distinguished from a metric ultraproduct of classical groups of Lie type, where the permutation degrees resp. dimensions of the natural module tend to infinity. This is done by considering the structure of centralizers of torsion elements in these groups (see Theorems 2.8 and 2.9 of [73]). In the case of a metric ultraproduct of classical groups of Lie type, in Theorem 2.8 of [73], investigating the structure of such centralizers of semisimple and unipotent torsion elements, Thom and Wilson even extract the 'limit characteristic' of the group. At the end of Section 2 of [73] they ask which metric ultraproducts of classical groups of different types can be isomorphic.

In this chapter, we will give an almost complete answer to this question in the case when the field sizes are bounded. We will show that for a metric ultraproduct of alternating or classical groups of Lie type of unbounded rank over fields of bounded size one can extract the Lie type (up to one exception). Also one can extract the 'limit field size'. Our results are summed up in Theorem 4.1 below. To state it, we first need to introduce some notation.

Recall the definitions from Section 0.1(a)–(f) and Section 0.3. Also recall the classification given in Subsection 3.4.2 §1. Let $\mathcal{H} = (H_i)_{i\in I}$ be a sequence of groups where either $H_i = \mathrm{S}_{n_i}$ is a symmetric group or $H_i = X_i(q_i)$ is a classical group of Lie type X_i over the field $\mathbb{F}_{q_i}$ (resp. $\mathbb{F}_{q_i^2}$ in the unitary case; $i \in I$). Here let each X_i be one of GL_{n_i}, Sp_{2m_i}, $\mathrm{GO}_{2m_i}^{\pm}$, GO_{2m_i+1} (q_i odd), or GU_{n_i} for suitable $m_i, n_i \in \mathbb{Z}_+$ ($i \in I$). Throughout, let $G := \mathcal{H}_{\mathcal{U}}^{\mathrm{met}}$ be the metric ultraproduct of these groups (see Definition 0.7) with respect to the normalized Hamming length function ℓ_{H} resp. the normalized rank length function

ℓ_{rk} (see Definition 0.12) when H_i is a symmetric resp. a classical linear group of Lie type. Assume that the permutation degrees resp. dimensions of the natural module n_i of H_i $(i \in I)$ tend to infinity along $\mathcal{U}$.

Note that, since $\mathcal{U}$ is an ultrafilter, we may assume that each group H_i is of the same type, i.e., all groups H_i are either symmetric, general linear, symplectic, general orthogonal, or general unitary groups. In these five distinct cases, we write $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}$, $\mathrm{Sp}_{\mathcal{U}}$, $\mathrm{GO}_{\mathcal{U}}$, or $\mathrm{GU}_{\mathcal{U}}$ for G. Also, when the field sizes q_i are bounded, we may assume that $q_i = q$ is constant $(i \in I)$, setting $q \coloneqq \lim_{\mathcal{U}} q_i$. Throughout, set $Z \coloneqq \mathbf{Z}(G)$ and $\overline{G} \coloneqq G/Z$.

If the groups H_i $(i \in I)$ are symmetric groups, then $Z = \mathbf{1}$ and $\overline{G} = G$. Now assume that all groups H_i are of type $X(q_i)$ $(i \in I$; i.e., they are not symmetric groups). Then $\overline{G} = G/Z = \overline{\mathcal{H}}_{\mathcal{U}}^{\text{met}}$ is the metric ultraproduct of the groups $\overline{H}_i \coloneqq H_i/\mathbf{Z}(H_i)$ with respect to the projective rank length function ℓ_{pr}. Recall from Theorem 1.3 that $\overline{G}$ is the unique simple quotient of G.

Similarly to the above, write $\mathrm{PGL}_{\mathcal{U}}$, $\mathrm{PSp}_{\mathcal{U}}$, $\mathrm{PGO}_{\mathcal{U}}$, or $\mathrm{PGU}_{\mathcal{U}}$ for $\overline{G}$ when all the groups H_i $(i \in I)$ are general linear, symplectic, general orthogonal, or general unitary groups. Moreover, in this case, if all the fields $\mathbb{F}_{q_i}$ $(i \in I)$ are equal to $\mathbb{F}_q$ (or $\mathbb{F}_{q^2}$ in the unitary case), we write $\mathrm{GL}_{\mathcal{U}}(q)$, $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, $\mathrm{GU}_{\mathcal{U}}(q)$ resp. $\mathrm{PGL}_{\mathcal{U}}(q)$, $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, $\mathrm{PGU}_{\mathcal{U}}(q)$ for G resp. $\overline{G}$. Also write $\mathbf{M}_{\mathcal{U}}$, $\mathbf{M}_{\mathcal{U}}(q)$ resp. $\mathrm{P}\mathbf{M}_{\mathcal{U}}$, $\mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$ for the metric ultraproduct of the spaces $\mathbf{M}_{n_i}(q_i)$, $\mathbf{M}_{n_i}(q)$ resp. $\mathrm{P}\mathbf{M}_{n_i}(q_i)$, $\mathrm{P}\mathbf{M}_{n_i}(q)$ with respect to the metrics d_{rk} and d_{pr} $(i \in I$; see Remark 0.10), so that $\mathrm{GL}_{\mathcal{U}} \subseteq \mathbf{M}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q) \subseteq \mathbf{M}_{\mathcal{U}}(q)$, $\mathrm{PGL}_{\mathcal{U}} \subseteq \mathrm{P}\mathbf{M}_{\mathcal{U}}$, $\mathrm{PGL}_{\mathcal{U}}(q) \subseteq \mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$. Throughout, if not stated otherwise, let $k = \mathbb{F}_q$ when G is classical non-unitary and $k = \mathbb{F}_{q^2}$ in the unitary case.

If all H_i $(i \in I)$ are symplectic, orthogonal, or unitary, write f_i for the forms stabilized (and Q_i for the quadratic form in the orthogonal case in characteristic two).

The main result of this chapter is now as follows.

Theorem 4.1. *Let $\overline{G} \cong \overline{G}_1 \cong \overline{G}_2$ with $G_j = X_{j\mathcal{U}_j}(q_j)$, where $X_j \in \{\mathrm{GL}, \mathrm{Sp}, \mathrm{GO}, \mathrm{GU}\}$ $(j = 1, 2)$. Then it holds that $q_1 = q_2$. Also we must have $X_1 = X_2$ or $\{X_1, X_2\} = \{\mathrm{Sp}, \mathrm{GO}\}$. Moreover, an ultraproduct $\overline{X}_{1\mathcal{U}_1}$ where the sizes q_i of the finite fields $\mathbb{F}_{q_i}$ $(i \in I_1)$ tend to infinity along $\mathcal{U}_1$ cannot be isomorphic to an ultraproduct $\overline{X}_{2\mathcal{U}_2}(q)$.*

Let us say some words about the proof of Theorem 4.1. Our strategy is to compute double centralizers of semisimple torsion elements of a fixed order $o \in \mathbb{Z}_+$ in the above metric ultraproducts. If the sizes q_i $(i \in I)$ of the fields $\mathbb{F}_{q_i}$ are bounded, it turns out that these are always finite abelian groups. Then we consider the maximal possible exponent, which such a double centralizer can have. It turns out that this data is enough to determine the limit field size q and the Lie type (up to the exception mentioned in Theorem 4.1). If the field sizes q_i $(i \in I)$ tend to infinity, a double centralizer of such a torsion element of order $o > 2$ is always infinite.

4.2 Description of conjugacy classes in $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q)$, and $\mathrm{PGL}_{\mathcal{U}}(q)$

In this section, we give a description of the conjugacy classes of groups of type $\mathrm{S}_{\mathcal{U}}$ or $\mathrm{PGL}_{\mathcal{U}}(q)$. We will make use of this in the subsequent sections. At first we consider $\mathrm{GL}_{\mathcal{U}}(q)$ instead of $\mathrm{PGL}_{\mathcal{U}}(q)$. Throughout this chapter, all polynomials from $k[X]$ that occur are meant to be *monic* polynomials.

Conjugacy classes in $\mathrm{S}_{\mathcal{U}}$ *and* $\mathrm{GL}_{\mathcal{U}}(q)$*.* For an integer $k \in \mathbb{Z}_+$ and a polynomial $\xi \in k[X]$ define $r_k(\sigma) := |\mathrm{fix}(\sigma^k)|/n$ resp. $r_\xi(g) := \dim(\ker(\xi(g)))/n$ for $\sigma \in \mathrm{S}_n$ resp. $g \in \mathbf{M}_n(q)$. Extend this definition to $\mathrm{S}_{\mathcal{U}}$ and $\mathbf{M}_{\mathcal{U}}(q)$ by setting $r_k(\sigma) := \lim_{\mathcal{U}} r_k(\sigma_i)$ and $r_\xi(g) := \lim_{\mathcal{U}} r_\xi(g_i)$ for $\sigma = \overline{(\sigma_i)}_{i\in I}$ resp. $g = \overline{(g_i)}_{i\in I}$. Both expressions are well-defined, since for $\sigma = \overline{(\sigma_i)}_{i\in I} = \overline{(\tau_i)}_{i\in I} \in \mathrm{S}_{\mathcal{U}}$ one has

$$\begin{aligned}\frac{1}{n_i}\left||\mathrm{fix}(\sigma_i^k)| - |\mathrm{fix}(\tau_i^k)|\right| &\leq d_{\mathrm{H}}(\sigma_i^k, \tau_i^k) \\ &\leq d_{\mathrm{H}}(\sigma_i^k, \sigma_i^{k-1}\tau_i) + \cdots + d_{\mathrm{H}}(\sigma_i\tau_i^{k-1}, \tau_i^k) \\ &= k d_{\mathrm{H}}(\sigma_i, \tau_i) \to_{\mathcal{U}} 0.\end{aligned} \tag{4.1}$$

Similarly, if $g = \overline{(g_i)}_{i\in I} = \overline{(h_i)}_{i\in I} \in \mathrm{GL}_{\mathcal{U}}(q)$ and $\xi = a_0 + a_1X + \cdots + a_{k-1}X^{k-1} + X^k \in k[X]$ we have

$$\begin{aligned}\frac{1}{n_i}\left|\dim(\ker\xi(g_i)) - \dim(\ker\xi(h_i))\right| &\leq d_{\mathrm{rk}}(\xi(g_i), \xi(h_i)) \\ &\leq \sum_{j=0}^{k} d_{\mathrm{rk}}(a_j g_i^j, a_j h_i^j) \\ &\leq \sum_{j=0}^{k} d_{\mathrm{rk}}(g_i^j, h_i^j) \\ &\leq \left(\sum_{j=0}^{k} j\right) d_{\mathrm{rk}}(g_i, h_i) = \binom{k+1}{2} d_{\mathrm{rk}}(g_i, h_i)\end{aligned} \tag{4.2}$$

and the latter tends to zero along $\mathcal{U}$. Here we used the same trick as in Estimate (4.1) above to bound $d_{\mathrm{rk}}(g_i^j, h_i^j)$ by $j d_{\mathrm{rk}}(g_i, h_i)$ $(j = 0, \ldots, k)$ in Estimate (4.2). Write $r(\sigma) := (r_k(\sigma))_{k\in\mathbb{Z}_+}$ and $r(g) := (r_\xi(g))_{\xi\in k[X]}$. Now define $q_k(\sigma)$ for $k \in \mathbb{Z}_+$ via the equality

$$\sum_{d|k} q_d(\sigma) = r_k(\sigma),$$

for all $k \in \mathbb{Z}_+$. Write $q(\sigma) := (q_k(\sigma))_{k\in\mathbb{Z}_+}$. Applying Möbius inversion, we obtain

$$q_k(\sigma) = \sum_{d|k} \mu(k/d) r_d(\sigma).$$

Alternatively, one can think of $q_k(\sigma)$ as the $\mathcal{U}$-limit of the normalized cardinality of the support of all k-cycles in σ_i $(i \in I)$, i.e.,

$$q_k(\sigma) = \lim_{\mathcal{U}} n_k(\sigma_i)/n_i = k \lim_{\mathcal{U}} c_k(\sigma_i)/n_i.$$

Similarly, for $\chi \in k[X]$ primary define $q_\chi(g)$ via the equality

$$r_\xi(g) = \sum_{\chi \text{ primary}} \frac{\deg(\gcd\{\chi, \xi\})}{\deg(\chi)} q_\chi(g),$$

for all polynomials $\xi \in k[X]$. Write $q(g) \coloneqq (q_\chi(g))_{\chi \text{ primary}}$. Alternatively, one can think of $q_\chi(g)$ as the $\mathcal{U}$-limit of the normalized dimensions of the (not unique) subspaces $V_\chi(g_i)$ $(i \in I)$, i.e.,

$$q_\chi(g) = \lim_{\mathcal{U}} n_\chi(g_i)/n_i = k_\chi \lim_{\mathcal{U}} c_\chi(g_i)/n_i,$$

where $k_\chi = \deg(\chi) = e \deg(i)$. This is because, when g acts as $F(\chi)$ and $\xi \in k[X]$, then $\dim(\ker(\xi(g))) = \deg(\gcd\{\chi, \xi\})$.

We claim that the conjugacy classes in $\mathrm{S}_{\mathcal{U}}$ resp. $\mathbf{M}_{\mathcal{U}}(q)$ are in one-to-one correspondence with all tuples $(q_k(\sigma))_{k \in \mathbb{Z}_+}$ resp. $(q_\chi(g))_{\chi \text{ primary}}$, where the only condition on the sequences are that

$$\sum_{k \in \mathbb{Z}_+} q_k(\sigma) \leq 1 \text{ resp. } \sum_{\chi \text{ primary}} q_\chi(g) \leq 1.$$

Here we let $\mathrm{GL}_{\mathcal{U}}(q)$ act on $\mathbf{M}_{\mathcal{U}}(q)$ by conjugation. The element g lies in $\mathrm{GL}_{\mathcal{U}}(q)$ if and only if $q_\chi(g) = 0$ for $\chi = X^e$ $(e \geq 1)$. Indeed, one sees easily that $r(\sigma)$ resp. $r(g)$ is conjugacy invariant, and so is $q(\sigma)$ resp. $q(g)$ for $\sigma \in \mathrm{S}_{\mathcal{U}}$ and $g \in \mathbf{M}_{\mathcal{U}}(q)$.

To see the converse, let $\sigma = \overline{(\sigma_i)}_{i \in I}, \tau = \overline{(\tau_i)}_{i \in I} \in \mathrm{S}_{\mathcal{U}}$ resp. $g = \overline{(g_i)}_{i \in I}, h = \overline{(h_i)}_{i \in I} \in \mathbf{M}_{\mathcal{U}}(q)$ be elements such that $q(\sigma) = q(\tau)$ resp. $q(g) = q(h)$.

Find a sequence $(N_i)_{i \in I}$ tending to infinity along $\mathcal{U}$ such that

$$\sum_{k=1}^{N_i} |q_k(\sigma) - q_k(\sigma_i)|, \ \sum_{k=1}^{N_i} |q_k(\tau) - q_k(\tau_i)| \to_{\mathcal{U}} 0$$

resp. such that

$$\sum_{\substack{\chi \text{ primary} \\ \deg(\chi) \leq N_i}} |q_\chi(g) - q_\chi(g_i)|, \ \sum_{\substack{\chi \text{ primary} \\ \deg(\chi) \leq N_i}} |q_\chi(h) - q_\chi(h_i)| \to_{\mathcal{U}} 0.$$

Then by the triangle inequality

$$\frac{1}{n_i} \sum_{k=1}^{N_i} |n_k(\sigma_i) - n_k(\tau_i)| \to_{\mathcal{U}} 0$$

resp.

$$\frac{1}{n_i} \sum_{\substack{\chi \text{ primary} \\ \deg(\chi) \leq N_i}} |\dim(V_\chi(g_i)) - \dim(V_\chi(h_i))| \to_{\mathcal{U}} 0.$$

Hence we can conjugate a big part of $\bigsqcup_{k=1}^{N_i} \Omega_k(\sigma_i)$ equivariantly to a big part of $\bigsqcup_{k=1}^{N_i} \Omega_k(\tau_i)$ resp. an almost fulldimensional part of

$$\bigoplus_{\substack{\chi \text{ primary} \\ \deg(\chi) \leq N_i}} V_\chi(g_i) \quad \text{to} \quad \bigoplus_{\substack{\chi \text{ primary} \\ \deg(\chi) \leq N_i}} V_\chi(h_i)$$

equivariantly (with no error in the limit; here again $V_\chi(g_i)$ resp. $V_\chi(h_i)$ are not unique). The remaining part of σ_i and τ_i resp. g_i and h_i can be modified into one big cycle resp. one big Frobenius block with no change of σ and τ resp. g and h, since $N_i \to_{\mathcal{U}} \infty$. Then we conjugate this cycle resp. Frobenius block onto the other.

The case of $\mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$. Let the group $k^\times = \mathbb{F}_q^\times$ act on all (monic) polynomials $\xi \in k[X]$ by $\xi.z := z^{-k_\xi}\xi(zX)$, where $k_\xi := \deg(\xi)$. Extend this action to all tuples $q = (q_\chi)_{\chi \text{ primary}}$ with $\sum_{\chi \text{ primary}} q_\chi \leq 1$ via

$$q.z = (q_{\chi.z})_{\chi \text{ primary}}$$

and denote by $\overline{q}$ the orbit $\mathrm{orb}_{k^\times}(q)$ of q under this action of $k^\times$.

Let $\overline{G} = \mathrm{PGL}_{\mathcal{U}}(q)$. We claim that the conjugacy classes of elements $\overline{g} \in \mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$ are classified by the bijection $\overline{g}^{\overline{G}} \mapsto \overline{(q_\chi(g))}_{\chi \text{ primary}}$, where g is any lift of $\overline{g}$ in $\mathbf{M}_{\mathcal{U}}(q)$ (here we exclude the tuple q where $q_X = 1$ and $q_\chi = 0$ otherwise).

Indeed, the map is well-defined, since any other h such that $\overline{h} = \overline{g} \in \mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$ is of the form zg for some $z \in k^\times$ (as $k = \mathbb{F}_q$ is finite), so that $(q_\chi(h))_{\chi \text{ primary}} = (q_\chi(zg))_{\chi \text{ primary}} = (q_{\chi.z}(g))_{\chi \text{ primary}} = q(g).z$. Also q is constant on conjugacy classes of $\mathbf{M}_{\mathcal{U}}(q)$ (under the action of $\mathrm{GL}_{\mathcal{U}}(q)$).

Conversely, if $q_\chi(h) = q_{\chi.z}(g) = q_\chi(zg)$ for some fixed $z \in k^\times$ and all $\chi \in k[X]$ primary, then from the above we derive that the elements g and $z^{-1}h$ of $\mathbf{M}_{\mathcal{U}}(q)$ are conjugate, so that $\overline{g}$ and $\overline{h}$ are conjugate in $\mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$. This proves the claim.

Remark 4.2. For G of type $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$ the conjugacy classes of elements $g \in G$ for which $\sum_{\chi \text{ primary}} q_\chi(g) = 1$ are still characterized by the tuples $(q_\chi(g))_{\chi \text{ primary}}$. The only necessary additional restriction on these tuples is that $q_\chi(g) = q_{\chi^*}(g)$, where χ^* is the dual polynomial of χ defined in Subsection 3.4.2 §1.

Indeed, assume $g = \overline{(g_i)}_{i\in I}, h = \overline{(h_i)}_{i \in I} \in G$, $q(g) = q(h)$ and

$$\sum_{\chi \text{ primary}} q_\chi(g) = \sum_{\chi \text{ primary}} q_\chi(h) = 1.$$

Then g is conjugate to h.

This holds, since on all but constantly many Frobenius blocks $F(\chi)$ of g_i resp. h_i ($i \in I$), where χ is a self-dual primary polynomial or of the form $\chi = \xi\xi^*$, where ξ is not self-dual primary (see Subsection 3.4.2 §1), the form f_i (and in characteristic two the quadratic form Q_i) is uniquely determined up to linear equivalence, so that we can map these blocks of g_i to such blocks of h_i and extend this partial map by Witt's lemma (Lemma 0.1).

Conversely, if we have a tuple $(q_\chi)_{\chi \text{ primary}}$ such that $\sum_{\chi \text{ primary}} q_\chi = 1$ and $q_\chi = q_{\chi^*}$, we can see from Fact 3.40 that there exists an element $g \in G$ such $q_\chi(g) = q_\chi$ for all $\chi \in k[X]$ primary.

Recall that $\overline{G} = G/Z$, where $Z = \mathbf{Z}(G)$. For a tuple $q = (q_\chi)_{\chi \text{ primary}}$ let $\overline{q}$ denote its orbit $\mathrm{orb}_Z(q)$ under the action of $Z \leq k^\times$. Then for the elements $\overline{g} \in \overline{G}$ (i.e., $\overline{G}$ is one of $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, or $\mathrm{PGU}_{\mathcal{U}}(q)$) such that $\sum_{\chi \text{ primary}} q_\chi(g) = 1$ for one lift $g \in G$ of $\overline{g}$ the same characterization as for $\mathrm{PGL}_{\mathcal{U}}(q)$ above holds by the same argument. Again we need to restrict the tuples $q = (q_\chi)_{\chi \text{ primary}}$ so that $q_\chi = q_{\chi^*}$ for all χ primary.

However, we conjecture that the above characterization for G of type $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$ or $\mathrm{GU}_{\mathcal{U}}(q)$ is false if

$$\sum_{\chi \text{ primary}} q_\chi(g) < 1$$

for an element $g \in G$.

Remark 4.3. It is easy to see that $(\chi.z)^* = \chi^*.z$ for $z \in Z$. Indeed, since $z^\sigma = z^{-1}$ for $z \in Z$, we have $(\chi.z)^* = (z^{-k}\chi(zX))^* = (z^k)^\sigma a_0^{-\sigma} X^k (z^{-k})^\sigma \chi^\sigma(z^\sigma X^{-1}) = a_0^{-\sigma} X^k \chi^\sigma((zX)^{-1}) = z^{-k}\chi^*(zX) = \chi^*.z$, where $\chi = a_0 + a_1 X + \cdots + a_{k-1}X^{k-1} + X^k$ and $\sigma \in \mathrm{Aut}(k)$ is defined as at the beginning of Subsection 3.4.2.

4.3 Characterization of torsion elements in $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q)$, and $\mathrm{PGL}_{\mathcal{U}}(q)$

In this section, we characterize torsion elements in metric ultraproducts of the above type. At first note that an invertible element in $\mathbf{M}_{\mathcal{U}}(q)$, i.e., an element of $\mathrm{GL}_{\mathcal{U}}(q)$, is algebraic over $k = \mathbb{F}_q$ if and only if it is torsion. Indeed, if g is torsion, then $g^o - 1 = 0$ for some integer $o \geq 1$. Conversely, if g is algebraic and invertible, let $\chi \in k[X]$ be the its minimal polynomial. Setting $o := |(k[X]/(\chi))^\times| < \infty$ one sees that $g^o = 1$ as g is invertible.

Here comes the promised characterization of torsion elements.

Lemma 4.4. *An element $\sigma \in \mathrm{S}_{\mathcal{U}}$ resp. $g \in \mathrm{GL}_{\mathcal{U}}(q)$ is torsion if and only if there is $N \in \mathbb{Z}_+$ such that*

$$\sum_{k=1}^{N} q_k(\sigma) = 1 \textit{ resp. } \sum_{\substack{\chi \text{ primary} \\ \deg(\chi) \leq N}} q_\chi(g) = 1.$$

An element $\overline{g} \in \mathrm{PGL}_{\mathcal{U}}(q)$ is torsion if and only if any lift $g \in \mathrm{GL}_{\mathcal{U}}(q)$ is torsion.

Proof. Indeed, if the above two conditions are fulfilled, then writing $o := \mathrm{lcm}\{1, \ldots, N\}$ resp. $o := \mathrm{lcm}\{|(k[X]/(\chi))^\times| \mid \chi \text{ primary}, \deg(\chi) \leq N\}$, we have $\ell_{\mathrm{H}}(\sigma_i^o) \to_{\mathcal{U}} 0$ resp.

$\ell_{\mathrm{rk}}(g_i^o) \to_{\mathcal{U}} 0$ meaning that $\sigma^o = 1$ resp. $g^o = 1$. Conversely, if we assume $\sigma^o = 1$ resp. $g^o = 1$, we get that $\ell_{\mathrm{H}}(\sigma_i^o) \to_{\mathcal{U}} 0$ resp. $\ell_{\mathrm{rk}}(g_i^o) \to_{\mathcal{U}} 0$, meaning that, asymptotically, the d-cycles in σ_i for $d \mid o$ support the whole set resp. all Frobenius blocks $F(\chi)$ for $\chi \mid X^o - 1$ primary support the whole vector space, so taking $N := o$ above, we get the converse direction.

The last statement follows, since the kernel of the surjective homomorphism $\mathrm{GL}_{\mathcal{U}}(q) \to \mathrm{PGL}_{\mathcal{U}}(q)$ equals $k^\times = \mathbb{F}_q^\times$, which is finite. Hence, if $g \in \mathrm{GL}_{\mathcal{U}}(q)$ represents $\overline{g} \in \mathrm{PGL}_{\mathcal{U}}(q)$ and the latter is of order $o < \infty$, we have that $\mathrm{ord}(g) \mid o(q-1) < \infty$. $\square$

4.4 Faithful action of $\mathrm{S}_{\mathcal{U}}$ and $\mathrm{PGL}_{\mathcal{U}}(q)$ on the Loeb space and the associated continuous geometry

In this section, we show that the groups $\mathrm{S}_{\mathcal{U}}$ and $\mathrm{PGL}_{\mathcal{U}}(q)$ faithfully act on natural associated objects. For this purpose we need the so-called *Loeb space*

$$L(n_i)_{i\in I} := (\mathcal{S}, \mu)$$

resp. its vector space analog, the *continuous geometry*

$$V(n_i)_{i\in I} := (\mathcal{V}, \dim)\,,$$

which are associated naturally to the metric ultraproduct $\mathrm{S}_{\mathcal{U}}$ resp. $\mathrm{PGL}_{\mathcal{U}}(q)$.

Here $\mathcal{S}$ resp. $\mathcal{V}$ equals $\prod_{i\in I} \mathcal{P}(\underline{n_i})$ resp. $\prod_{i\in I} \mathrm{Sub}(k^{n_i})$ modulo the equivalence relation $(S_i)_{i\in I} \sim (T_i)_{i\in I}$ resp. $(U_i)_{i\in I} \sim (V_i)_{i\in I}$ iff $\mu_i(S_i \triangle T_i) \to_{\mathcal{U}} 0$ resp. $\dim_i(U_i + V_i) - \dim_i(U_i \cap V_i) \to_{\mathcal{U}} 0$, where μ_i resp. $\dim_i$ is the normalized counting measure resp. dimension on $\underline{n_i}$ resp. k^{n_i} (and $A \triangle B$ denotes the symmetric difference of the sets A and B). Then one defines μ resp. dim by

$$\mu(S) = \mu(\overline{(S_i)}_{i\in I}) := \lim_{\mathcal{U}} \mu_i(S_i)$$

and

$$\dim(V) = \dim(\overline{(V_i)}_{i\in I}) := \lim_{\mathcal{U}} \dim_i(V_i).$$

It is easy to check that both are well-defined in this way. Also the operations $\cup, \cap$ resp. $+, \cap$ are inherited to $\mathcal{S}$ resp. $\mathcal{V}$ in a natural way, e.g., $\overline{(S_i)}_{i\in I} \cap \overline{(T_i)}_{i\in I} := \overline{(S_i \cap T_i)}_{i\in I}$. Write $S \subseteq T$ resp. $U \leq V$ iff $\mu(S\cap T) = \mu(S)$ resp. $\dim(U \cap V) = \dim(U)$. Call a permutation of $\mathcal{S}$ resp. $\mathcal{V}$ an automorphism iff it preserves μ resp. dim and the relation $\subseteq$ resp. $\leq$.

Then one observes that $\mathrm{S}_{\mathcal{U}}$ resp. $\mathrm{PGL}_{\mathcal{U}}(q)$ is faithfully represented as group of automorphisms of $(\mathcal{S}, \mu)$ resp. $(\mathcal{V}, \dim)$.

At first we consider the case $\overline{G} = \mathrm{S}_{\mathcal{U}}$. Indeed, assume for fixed $\sigma = \overline{(\sigma_i)}_{i\in I} \in \mathrm{S}_{\mathcal{U}}$ that $S.\sigma = S$ for all $S \in \mathcal{S}$. Then take $S = \overline{(S_i)}_{i\in I}$, where $S_i \subseteq \underline{n_i}$ is taken in the following way. For each k-cycle $c \subseteq \underline{n_i}$ ($k > 1$; here seen as a set) we pick $s_c \in c$ and define S_i by $S_i \cap c = \{s_c, s_c.\sigma_i^2, \ldots, s_c.\sigma_i^{2(\lfloor k/2\rfloor - 1)}\}$ and $S_i \cap \Omega_1(\sigma_i) = \emptyset$. Then $S_i \triangle S_i.\sigma_i = \emptyset$

and $\mu_i(S_i) \geq 1/3|\mathrm{supp}(\sigma_i)|$. This means that S is fixed by σ if and only if $\mathrm{supp}(\sigma) := \overline{(\mathrm{supp}(\sigma_i))_{i\in I}} = \overline{(\emptyset)}$ has measure zero. But this means $\sigma = \mathrm{id}$ in the metric ultraproduct $\mathrm{S}_{\mathcal{U}}$.

Now consider the case $\overline{G} = \mathrm{PGL}_{\mathcal{U}}(q)$. Here, similarly, assume for fixed $g = \overline{(g_i)_{i\in I}} \in \mathrm{GL}_{\mathcal{U}}(q)$ that $V.g = V$ for all $V \in \mathcal{V}$. Then take $V = \overline{(V_i)_{i\in I}}$ in the following way: The linear transformation g_i is a direct sum of Frobenius blocks $F(\chi)$, where $\chi \in k[X]$ runs through all (monic) primary polynomials. For each such block $b \leq k_i^{n_i}$ of dimension $k_b > 1$ (here seen as a subspace) of g_i select a cyclic vector v_b. Then define V_i by $V_i = \bigoplus_{b,k_b>1} \langle v_b, v_b.g_i^2, \ldots, v_b.g_i^{2(\lfloor k_b/2\rfloor -1)}\rangle$. Then one observes that $V_i \cap V_i.g_i = 0$, so that $\dim_i(V_i + V_i.g_i) - \dim_i(V_i \cap V_i.g_i) = 2\dim_i(V_i)$. This shows that $q_\chi(g) = 0$ for all $\chi \in k[X]$ primary of degree $k_\chi > 1$. But one observes that, if $q_{(X-\lambda)}(g), q_{(X-\mu)}(g) \geq \varepsilon > 0$ for $\lambda \neq \mu$ elements of k, we can use the following construction: Let $e_{i1}, \ldots, e_{ik_i} \in V_{X-\lambda}(g_i)$ and $f_{i1}, \ldots, f_{ik_i} \in V_{X-\mu}(g_i)$ such that $\lim_{\mathcal{U}} k_i/n_i = \varepsilon$. Define $V_i := \langle e_{ij} + f_{ij} \mid j = 1, \ldots, k_i\rangle$ $(i \in I)$. Assume $v \in V_i \cap V_i.g_i$, then there exists numbers $\alpha_1, \ldots, \alpha_{k_i}, \beta_1, \ldots, \beta_{k_i} \in k$ such that

$$v = \sum_{j=1}^{k_i} \alpha_j(e_{ij} + f_{ij}) = \sum_{j=1}^{k_i} \beta_j(\lambda e_{ij} + \mu f_{ij}).$$

This gives that

$$\sum_{j=1}^{k_i} (\alpha_j - \beta_j\lambda)e_{ij} = \sum_{j=1}^{k_i} (\beta_j\mu - \alpha_j)f_{ij},$$

so that by disjointness of $V_{X-\lambda}(g_i)$ and $V_{X-\mu}(g_i)$ both sides are zero and so, since the e_{ij}, f_{ij} $(j = 1, \ldots, k_i)$ are linearly independent, we get that $\alpha_j - \beta_j\lambda = \beta_j\mu - \alpha_j = 0$, so that, since $\lambda \neq \mu$, we obtain $\alpha_j = \beta_j = 0$. Hence $v = 0$ and $V_i \cap V_i.g_i = 0$. But $\lim_{\mathcal{U}} \dim(V_i)/n_i \geq \varepsilon$, yielding the same contradiction as above. Therefore we must have $g = \lambda\,\mathrm{id}$ (as $k = \mathbb{F}_q$ is finite) in the metric ultraproduct $\mathrm{GL}_{\mathcal{U}}(q)$, i.e., $\mathrm{PGL}_{\mathcal{U}}(q)$ is faithfully represented.

Remark 4.5. The above statement about $\mathrm{PGL}_{\mathcal{U}}(q)$ holds for any such metric ultraproduct of groups $\mathrm{PGL}_{n_i}(k_i)$ where the fields k_i are not restricted with the same proof. Here the kernel of the action $\mathrm{GL}_{\mathcal{U}} \to \mathrm{Aut}(\mathcal{V}, \dim)$ is given by $\prod_{\mathcal{U}} k_i^\times$ (the algebraic ultraproduct of these groups).

Remark 4.6. Hence, if the sequence of subsets $(S_i)_{i\in I}$ resp. subspaces $(V_i)_{i\in I}$ is almost stabilized by each element of a subgroup H of $\overline{G} = \mathrm{S}_{\mathcal{U}}$ resp. $\overline{G} = \mathrm{PGL}_{\mathcal{U}}(q)$ (or of $G = \mathrm{GL}_{\mathcal{U}}(q)$), we can restrict H to $S := \overline{(S_i)_{i\in I}}$ resp. $V := \overline{(V_i)_{i\in I}}$.

Remark 4.7. For an element $\sigma = \overline{(\sigma_i)_{i\in I}}$ set $\Omega_k(\sigma) := \overline{(\Omega_k(\sigma_i))_{i\in I}} \in \mathcal{S}$ for $k \in \mathbb{Z}_+$. Similarly, for a *semisimple* element $g = \overline{(g_i)_{i\in I}} \in \mathrm{GL}_{\mathcal{U}}(q)$ (see the next section for the definition of semisimple elements) set $V_\chi(g) := \overline{(V_\chi(g_i))_{i\in I}} \in \mathcal{V}$ for $\chi \in k[X]$ primary. Note that these definitions are independent of the chosen representatives (for the uniqueness of $V_\chi(g)$ we need that g is semisimple, since then $V_\chi(g_i) = \ker(\chi(g_i))$ for a suitable representative $(g_i)_{i\in I}$ of g).

Remark 4.8. Call a sequence of subsets $(B_i)_{i\in I} \subseteq k^{n_i}$ a basis of $V \in \mathcal{V}$ if there is a representative $(V_i)_{i\in I}$ of V such that B_i is a basis of V_i ($i \in I$).

Remark 4.9. Call $V \in \mathcal{V}$ totally singular if it has a representative $(V_i)_{i\in I}$ such that V_i is totally singular ($i \in I$).

4.5 Centralizers in $\mathrm{S}_{\mathcal{U}}$, $\mathrm{GL}_{\mathcal{U}}(q)$, $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, and $\mathrm{GU}_{\mathcal{U}}(q)$

In this section, we provide tools (Lemmas 4.10 and 4.11) to compute centralizers of certain elements from the metric ultraproducts $\mathrm{S}_{\mathcal{U}}$ and $\mathrm{GL}_{\mathcal{U}}(q)$. We will use this in Section 4.6 to compute centralizers of elements in $\mathrm{PGL}_{\mathcal{U}}(q)$.

Centralizers of elements in $G = \mathrm{S}_{\mathcal{U}}, \mathrm{GL}_{\mathcal{U}}(q)$. Note that for $\sigma = \overline{(\sigma_i)_{i\in I}} \in \mathrm{S}_{\mathcal{U}}$ resp. $g = \overline{(g_i)_{i\in I}} \in \mathrm{GL}_{\mathcal{U}}(q)$ we have $\prod_{\mathcal{U}} \mathbf{C}(\sigma_i) \leq \mathbf{C}(\sigma)$ resp. $\prod_{\mathcal{U}} \mathbf{C}(g_i) \leq \mathbf{C}(g)$ (subsequently, by this notation we mean the metric ultraproduct of subgroups of the H_i ($i \in I$)). In the following lemma, we characterize when the above inclusion is actually an equality in the case of $\mathrm{S}_{\mathcal{U}}$.

Lemma 4.10. *An element $\sigma \in \mathrm{S}_{\mathcal{U}}$ satisfies $\sum_{k\in\mathbb{Z}_+} q_k(\sigma) = 1$ if and only if for each choice of a representative $(\sigma_i)_{i\in I}$ of σ the centralizer $\mathbf{C}(\sigma)$ equals $\prod_{\mathcal{U}} \mathbf{C}(\sigma_i)$.*

Before proving Lemma 4.10, we turn to $\mathrm{GL}_{\mathcal{U}}(q)$. An element $g \in \mathrm{GL}_{\mathcal{U}}(q)$ is called *semisimple* if it has a representative $(g_i)_{i\in I}$ such that $g_i \in \mathrm{GL}_{n_i}(q)$ is semisimple, i.e., of order prime to q.

Lemma 4.11. *A semisimple element $g \in \mathrm{GL}_{\mathcal{U}}(q)$ satisfies $\sum_{\chi \text{ primary}} q_\chi(g) = 1$ if and only if for each choice of a representative $(g_i)_{i\in I}$ of g where each g_i is semisimple ($i \in I$) the centralizer $\mathbf{C}(g)$ equals $\prod_{\mathcal{U}} \mathbf{C}(g_i)$.*

To prove Lemmas 4.10 and 4.11, we need the following auxiliary result.

Lemma 4.12. *The following are true:*

(i) Assume $\sigma \in \mathrm{Sym}(\underline{n})$ is of order k and $S \subseteq \underline{n}$ has normalized counting measure $\mu(S)$. Then S contains a σ-invariant subset T of measure $\mu(T) \geq 1 - k(1-\mu(S))$.

(ii) Assume $g \in \mathrm{GL}(V)$ for a k-vector space V and that the minimal polynomial of g over k has degree k. Assume $U \leq V$, then there exists a g-invariant subspace of U of codimension at most $k\,\mathrm{codim}(U)$.

Proof. (i): Observe that the biggest σ-invariant subset of S is equal to $T = \bigcap_{i\in\mathbb{Z}} S.\sigma^i$. But since $\sigma^k = \mathrm{id}$ by assumption, we see that actually $T = \bigcap_{i=0}^{k-1} S.\sigma^i$. Hence, since $\mu(S.\sigma^i) = \mu(S)$ for all $i \in \mathbb{Z}$, we have that $\mu(T) \geq 1 - k(1-\mu(S))$.

(ii): Similarly to the above, the biggest g-invariant subspace contained in U is $W = \bigcap_{i\in\mathbb{Z}} U.g^i$. Now $v \in \bigcap_{i=0}^{k-1} U.g^i$ means that $v, \ldots, v.g^{-(k-1)} \in U$. But then $v.g^{-k} = -\frac{1}{a_0}(a_1 v.g^{-(k-1)} + \cdots + a_{k-1} v.g^{-1} + v) \in U$, where $\chi = a_0 + a_1 X + \cdots + a_{k-1}X^{k-1} + X^k$ is the minimal polynomial of g. Note that $a_0 = (-1)^k \det(g) \neq 0$. This shows that actually $W = \bigcap_{i=0}^{k-1} U.g^i$, so that $\operatorname{codim}(W) \leq k \operatorname{codim}(U)$. □

Remark 4.13. The bounds in Lemma 4.12 are sharp. E.g., take σ of type (k^{c_k}) and set $n = c_k k$. Take S of size $n - s$ such that for precisely $s \leq c_k$ k-cycles of σ, S contains $k-1$ elements of each of them and all elements of the remaining k-cycles. Then the set T constructed in Lemma 4.12 has size $n - ks$. In (ii) we can use a similar construction.

Now we are able to prove the Lemmas 4.10 and 4.11.

Proof of Lemmas 4.10 and 4.11. At first we prove Lemma 4.10. Assume that $\sigma = \overline{(\sigma_i)}_{i\in I}$, $\tau = \overline{(\tau_i)}_{i\in I} \in \mathrm{S}_{\mathcal{U}}$ commute and assume that $\sum_{k=1}^{\infty} q_k(\sigma) = 1$. Find a sequence $(N_i)_{i\in I}$ tending to infinity along $\mathcal{U}$ such that

$$\lim_{\mathcal{U}} \sum_{k=1}^{N_i} q_k(\sigma_i) = 1 \text{ and } \binom{N_i+1}{2} \ell_{\mathrm{H}}([\sigma_i, \tau_i]) \to_{\mathcal{U}} 0.$$

Recall that $C_k(\sigma_i)$ denotes the set of k-cycles of σ_i (see Section 0.1(d)). Call a k-cycle of σ_i *bad* if it is not mapped σ_i-equivariantly to another k-cycle of σ_i by τ_i. Collect the set of bad k-cycles of σ_i in $C'_k(\sigma_i)$. For each bad k-cycle of σ_i we get at least one non-fixed point of $[\sigma_i, \tau_i]$, so that $|C'_k(\sigma_i)|/n_i \leq \ell_{\mathrm{H}}([\sigma_i, \tau_i])$ for all $k \in \mathbb{Z}_+$. Hence, if we change τ_i such that all bad k-cycles of σ_i are mapped accurately for $k \leq N_i$ and all k-cycles for $k > N_i$ are mapped identically, we get a permutation τ'_i such that

$$\begin{aligned} d_{\mathrm{H}}(\tau_i, \tau'_i) &\leq \frac{1}{n_i} \sum_{k=1}^{N_i} k|C'_k(\sigma_i)| + \sum_{k=N_i+1}^{\infty} q_k(\sigma_i) \\ &\leq \left(\sum_{k=1}^{N_i} k\right) \ell_{\mathrm{H}}([\sigma_i, \tau_i]) + \sum_{k=N_i+1}^{\infty} q_k(\sigma_i) \\ &= \binom{N_i+1}{2} \ell_{\mathrm{H}}([\sigma_i, \tau_i]) + \sum_{k=N_i+1}^{\infty} q_k(\sigma_i). \end{aligned}$$

By the assumption $\sum_{k=1}^{\infty} q_k(\sigma) = 1$, the last term in the above estimate tends to zero along $\mathcal{U}$. Hence $\tau = \overline{(\tau_i)}_{i\in I} = \overline{(\tau'_i)}_{i\in I}$ and $[\sigma_i, \tau'_i] = 1$.

Conversely, assume that $\sum_{k=1}^{\infty} q_k(\sigma) < 1$. Choose the sequence $(N_i)_{i\in I}$ such that $\lim_{\mathcal{U}} \sum_{k=1}^{N_i} q_k(\sigma_i) = \sum_{k=1}^{\infty} q_k(\sigma)$ and $\lim_{\mathcal{U}} N_i/n_i = 0$.

For each $i \in I$ change σ_i to σ'_i such that the k-cycles of σ'_i are the same as in σ_i for $1 \leq k \leq N_i$ and the other k-cycles of σ'_i ($k > N_i$; if they exist) are grouped into one big K_i-cycle so that $d_{\mathrm{H}}(\sigma_i, \sigma'_i)$ is minimal possible. It is easy to see that then still $d_{\mathrm{H}}(\sigma_i, \sigma'_i) \to_{\mathcal{U}} 0$ as $N_i \to_{\mathcal{U}} \infty$. Now σ'_i eventually has precisely one K_i-cycle for $K_i > N_i$. Obtain σ''_i by dividing this K_i-cycle (if it exists) into two $\lfloor K_i/2 \rfloor$-cycles and at most one fixed point so that $d_{\mathrm{H}}(\sigma'_i, \sigma''_i) \leq 3/n_i$ is minimal. Note that $K_i/n_i = 1 - \sum_{k=1}^{N_i} q_k(\sigma_i) \to_{\mathcal{U}} \varepsilon > 0$, so that $\lfloor K_i/2 \rfloor > N_i$ along $\mathcal{U}$, as $\lim_{\mathcal{U}} N_i/n_i \to_{\mathcal{U}} 0$ by assumption.

Now consider the restriction of the centralizers $\mathbf{C}(\sigma_i')$ and $\mathbf{C}(\sigma_i'')$ to the support of the unique K_i-cycle of σ_i' (which certainly both fix setwise by the previous inequality). The first group is isomorphic to C_{K_i}, whereas the second is isomorphic to $\mathrm{C}_{\lfloor K_i/2\rfloor} \wr \mathrm{C}_2$. Taking the metric ultraproducts of these groups restricted to this support (in the sense of Remark 4.6), we get an abelian group in the first case, and a non-abelian group in the second case. Hence, in at least one case, $\prod_{\mathcal{U}} \mathbf{C}(\sigma_i') \neq \mathbf{C}(\sigma)$ or $\prod_{\mathcal{U}} \mathbf{C}(\sigma_i'') \neq \mathbf{C}(\sigma)$.

Now we prove Lemma 4.11. Assume that $g = \overline{(g_i)}_{i\in I}, h = \overline{(h_i)}_{i\in I} \in \mathrm{GL}_{\mathcal{U}}(q)$ commute, i.e., $[g,h] = \mathrm{id}$, that g and each g_i $(i \in I)$ is semisimple, and assume that

$$\sum_{\chi \text{ irreducible}}^{\infty} q_\chi(g) = 1.$$

Note that semisimplicity implies that for each Frobenius block $F(\chi)$ in the generalized Jordan normal form of g_i, $\chi = i^1$ is irreducible. Choose the sequence $(N_i)_{i\in I}$ such that

$$\lim_{\mathcal{U}} \sum_{\substack{\chi \text{ irreducible}\\ \deg(\chi)\leq N_i}} q_\chi(g_i) = 1 \text{ and } \left(\sum_{\substack{\chi \text{ irreducible}\\ \deg(\chi)\leq N_i}} \deg(\chi) \right) \ell_{\mathrm{rk}}([g_i,h_i]) \to_{\mathcal{U}} 0.$$

Define $U_i := \ker([g_i,h_i] - \mathrm{id})$. Fix an irreducible polynomial $\chi \in k[X]$ and apply Lemma 4.12(ii) inside $V := V_\chi(g_i)$ to the subspace $U := U_i \cap V_\chi(g_i)$ to get a g_i-invariant subspace $W = W_{i\chi} \leq U$ such that $\mathrm{codim}_V(W_{i\chi}) \leq k_\chi \,\mathrm{codim}(U_i)$, where $k_\chi = \deg(\chi)$. Note here that $V_\chi(g_i) = \ker(\chi(g_i))$ is unique, since g_i is semisimple. This large-dimensional subspace $W_{i\chi}$ is mapped accurately by h_i, as g_i commutes with h_i on it. Define h_i' to be equal to h_i on each $W_{i\chi}$ and complete it on each $V_{i\chi}$ to a map commuting with g_i for $\deg(\chi) \leq N_i$ (here we use semisimplicity of g_i). On $V_\chi(g_i)$ with $\deg(\chi) > N_i$ set h_i' to be the identity. As in the proof for $\mathrm{S}_{\mathcal{U}}$ above, it follows that $d_{\mathrm{rk}}(h_i,h_i') \to_{\mathcal{U}} 0$ and $[g_i,h_i'] = 1$.

Conversely, assume that $\sum_{\chi \text{ irreducible}}^{\infty} q_\chi(g) < 1$. Choose the sequence $(N_i)_{i\in I}$ such that

$$\lim_{\mathcal{U}} \sum_{\substack{\chi \text{ irreducible}\\ \deg(\chi)\leq N_i}}^{N_i} q_\chi(g_i) = \sum_{\chi \text{ irreducible}}^{\infty} q_\chi(g) \text{ and } \lim_{\mathcal{U}} N_i/n_i = 0.$$

For each $i \in I$ change g_i into g_i' such that all Frobenius blocks $F(\chi)$ for χ irreducible of degree at most N_i are left unchanged and all bigger Frobenius blocks (if there is any such block) are grouped into one big Frobenius block $F(\varphi)$ of size K_i (for φ irreducible). Define g_i'' in the same way, but split the Frobenius block $F(\varphi)$ (if it exists) into two or three blocks, two of which are $F(\phi)$ for ϕ irreducible of degree $\lfloor K_i/2\rfloor$ and, if K_i is odd, one block of size one, which is the identity. Then, as above, the centralizer of g_i' restricted to the large Frobenius block $F(\varphi)$ of it, equals $\mathbf{C}(g_i') \cong (k[X]/(\varphi))^\times$, whereas the centralizer $\mathbf{C}(g_i'')$ restricted to the same subspace is non-abelian (again in the sense of Remark 4.6). Also one sees that their metric ultraproducts are non-isomorphic, similarly to the case of permutations. The proof is complete. □

Remark 4.14. If G is one of $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$ and a semisimple $g \in G$ is represented by $(g_i)_{i\in I}$ and $\sum_{\chi \text{ irreducible}} q_\chi(g) = 1$, one can adapt the above argument for $\mathrm{GL}_{\mathcal{U}}(q)$ to see that still $\mathbf{C}(g) = \prod_{\mathcal{U}} \mathbf{C}(g_i)$ when all g_i are semisimple.

Indeed, from Subsection 3.4.2 §1 it follows that in the space $W_{i\chi} + W_{i\chi^*}$ (where $W_{i\chi}, W_{i\chi^*}$ are constructed as above) we can still find a big, i.e., almost fulldimensional, g_i-invariant non-singular subspace $W'_{i\chi,\chi^*}$. Then the form f_i (and Q_i in the orthogonal case in characteristic two) on $W'^{\perp}_{i\chi,\chi^*} \cap (V_\chi(g_i)+V_{\chi^*}(g_i))$ and $(W'_{i\chi,\chi^*}.h_i)^\perp \cap (V_\chi(g_i)+V_{\chi^*}(g_i))$ are isomorphic (which again follows from Subsection 3.4.2 §1), so that we can still complete our partial maps to h'_i $(i \in I)$.

As a consequence of Lemma 4.4 together with Lemmas 4.10 and 4.11, and Remark 4.14, we get the following corollary.

Corollary 4.15. *If $\sigma \in \mathrm{S}_{\mathcal{U}}$ resp. a semisimple element $g \in \mathrm{GL}_{\mathcal{U}}(q)$, $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$ is torsion, then $\mathbf{C}(\sigma)$ resp. $\mathbf{C}(g)$ is equal to $\prod_{\mathcal{U}} \mathbf{C}(\sigma_i)$ resp. $\prod_{\mathcal{U}} \mathbf{C}(g_i)$ for each representative $(\sigma_i)_{i\in I}$ resp. $(g_i)_{i\in I}$ of σ resp. g, where we require all g_i $(i \in I)$ to be semisimple.*

4.6 Centralizers in $\mathrm{PGL}_{\mathcal{U}}(q)$, $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, and $\mathrm{PGU}_{\mathcal{U}}(q)$

Now we can deduce the structure of centralizers of semisimple elements from $\mathrm{PGL}_{\mathcal{U}}(q)$, i.e., elements that lift to semisimple elements in $\mathrm{GL}_{\mathcal{U}}(q)$. Let $g = \overline{(g_i)_{i\in I}} \in \mathrm{GL}_{\mathcal{U}}(q)$ be a semisimple element which maps to $\overline{g} \in \mathrm{PGL}_{\mathcal{U}}(q) = \mathrm{GL}_{\mathcal{U}}(q)/k^\times$. Here g_i is also assumed to be semisimple $(i \in I)$.

Assume that $h = \overline{(h_i)_{i\in I}} \in \mathrm{GL}_{\mathcal{U}}(q)$ is such that $[g,h] = \mu\,\mathrm{id}$ for $\mu \in k^\times$, then $g^h = \mu g$, so that $q(g) = q(g^h) = q(\mu g) = q(g).\mu$, i.e., $\mu \in \mathrm{stab}_{k^\times}(q(g))$. Now let $\nu \in \mathrm{stab}_{k^\times}(q(g)) \leq k^\times$ be a generator of this cyclic group.

It is now easy to see that the conformal centralizer $\mathbf{C}_{\mathrm{conf}}(g) := \{h \in \mathrm{GL}_{\mathcal{U}}(q) \mid \text{there is } \mu \in k^\times \text{ such that } g^h = \mu g\}$ is an extension $\mathbf{C}(g).\mathrm{stab}_{k^\times}(q(g)) = \mathbf{C}(g).\langle\nu\rangle$ of $\mathbf{C}(g)$ by $\mathrm{stab}_{k^\times}(q(g))$. Hence $\mathbf{C}(\overline{g}) = (\mathbf{C}(g).\langle\nu\rangle)/k^\times$.

Remark 4.16. The analog statement of Lemma 4.11 is false in $\mathrm{PGL}_{\mathcal{U}}(q)$. Indeed, take a semisimple element $\overline{g} \in \mathrm{PGL}_{\mathcal{U}}(q)$ such that for a lift $g \in \mathrm{GL}_{\mathcal{U}}(q)$ the group $\mathrm{stab}_{k^\times}(q(g))$ is non-trivial. Choose a representative $(g_i)_{i\in I}$ of $g \in \mathrm{GL}_{\mathcal{U}}(q)$ such that $q_\chi(g_i) \neq q_\xi(g_i)$ for all $\chi, \xi \in k[X]$ distinct irreducible and g_i is semisimple $(i \in I)$. Then $\mathbf{C}(g_i)$ stabilizes each subspace $V_\chi(g_i) = \ker(\chi(g_i)) \leq k^{n_i}$. But this means that, if $h \in C := \prod_{\mathcal{U}} \mathbf{C}_{\mathrm{conf}}(g_i)$, we have that $g^h = g$, so that $C/k^\times$ is properly contained in $\mathbf{C}(\overline{g})$ (namely, $\mathbf{C}(\overline{g})/(C/k^\times) \cong \mathrm{stab}_{k^\times}(q(g))$, which is non-trivial).

Remark 4.17. For the groups $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, and $\mathrm{PGU}_{\mathcal{U}}(q)$ the same structure for $\mathbf{C}(\overline{g})$ holds, where $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$ resp. $\mathrm{GU}_{\mathcal{U}}(q)$ play the role of $\mathrm{GL}_{\mathcal{U}}(q)$. The possible scalars $\mu \in k^\times$ are restricted to $\mu \in \{\pm 1\}$ in the symplectic or orthogonal case, and to $\mu^{q+1} = 1$ in the unitary case.

4.7 Double centralizers of torsion elements

In this section, we compute the double centralizers of (semisimple) torsion elements of the groups $\overline{G}$ of type $\mathrm{S}_{\mathcal{U}}$, $\mathrm{PGL}_{\mathcal{U}}(q)$, $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, and $\mathrm{PGU}_{\mathcal{U}}(q)$. Note that for $g \in G$ a group element $\mathbf{C}(\mathbf{C}(g)) = \mathbf{Z}(\mathbf{C}(g))$, since $g \in \mathbf{C}(g)$, so that $\mathbf{C}(\mathbf{C}(g)) \leq \mathbf{C}(g)$. Set $\mathbf{C}^2(g) := \mathbf{C}(\mathbf{C}(g))$ and $\mathbf{C}^2_{\mathrm{conf}}(g) := \mathbf{C}_{\mathrm{conf}}(\mathbf{C}_{\mathrm{conf}}(g))$ to be the *double centralizer* resp. *double conformal centralizer* of g. Here $\mathbf{C}_{\mathrm{conf}}(g) := \{h \in G | [g,h] \in \mathbf{Z}(G)\}$.

4.7.1 The case $\mathrm{S}_{\mathcal{U}}$

Let $\sigma = \overline{(\sigma_i)}_{i \in I} \in \mathrm{S}_{\mathcal{U}} = \overline{G}$ be torsion of order o. Then $\sum_{k|o} q_k(\sigma) = 1$ by Lemma 4.4. By Corollary 4.15 we have that $\mathbf{C}(\sigma) = \prod_{\mathcal{U}} \mathbf{C}(\sigma_i)$. But $\mathbf{C}(\sigma_i)$ has a subgroup

$$\prod_{k|o} \mathrm{C}_k \wr \mathrm{Sym}(c_k(\sigma_i))$$

which is dense in it along $\mathcal{U}$, so that $C := \mathbf{C}(\sigma) = \prod_{\mathcal{U}} \prod_{k|o} \mathrm{C}_k \wr \mathrm{Sym}(c_k(\sigma_i))$.

At first, for simplicity, assume that σ_i is isotypic of type $(k^{c_{ik}})$ (so that $n_i = c_{ik}k$). Assume that $\tau = \overline{(\tau_i)}_{i \in I} \in \mathbf{Z}(C)$ and $\tau_i = (a_{ij}).\varphi_i \in \mathrm{C}_k \wr \mathrm{Sym}(c_{ik})$. Assume that $\lim_{\mathcal{U}} |\mathrm{supp}(\varphi_i)|/c_{ik} = \varepsilon > 0$. Then we can conjugate φ_i by $\phi_i \in \mathrm{Sym}(c_{ik}) \leq \mathrm{C}_k \wr \mathrm{Sym}(c_{ik}) = \mathbf{C}(\sigma_i)$ such that $\lim_{\mathcal{U}} d_{\mathrm{H}}(\varphi_i\phi_i, \phi_i\varphi_i) \geq \varepsilon > 0$. But this leads to the contradiction

$$\lim_{\mathcal{U}} d_{\mathrm{H}}(\tau_i\phi_i, \phi_i\tau_i) \geq \varepsilon > 0.$$

Hence we may assume that $\varphi_i = \mathrm{id}$, applying a small change to τ_i along $\mathcal{U}$ if necessary ($i \in I$). Now assume that $\lim_{\mathcal{U}} |\{j \,|\, a_{ij} = c\}|/c_{ik} = \varepsilon \in (0,1)$. Then we find permutations $\phi_i \in \mathrm{Sym}(c_{ik}) \leq \mathrm{C}_k \wr \mathrm{Sym}(c_{ik}) = \mathbf{C}(\sigma_i)$ such that $d_{\mathrm{H}}(\tau_i, \tau_i^{\phi_i}) = |\{j \,|\, a_{ij} \neq a_{ij.\phi_i}\}|/c_{ik} \geq \min\{\varepsilon, 1-\varepsilon\} > 0$. Hence we can assume that all a_{ij} are equal. This shows that, in this case, $\mathbf{Z}(\mathbf{C}(\sigma))$ is the metric ultraproduct $\prod_{\mathcal{U}} \mathrm{C}_k \cong \mathrm{C}_k$ where C_k in the ith component is generated by the element σ_i itself ($i \in I$).

In the general case, we obtain that

$$\mathbf{Z}(\mathbf{C}(\sigma)) = \prod_{k|o, q_k(\sigma)>0} \prod_{\mathcal{U}} \mathrm{C}_k \cong \prod_{k|o, q_k(\sigma)>0} \mathrm{C}_k.$$

This holds, because $\sigma \in \mathbf{C}(\sigma)$, so that, when $\tau \in \mathbf{Z}(\mathbf{C}(\sigma))$, it must commute with σ. But this implies that $\lim_{\mathcal{U}} |\Omega_k(\sigma_i) \triangle \Omega_k(\sigma_i).\tau_i| = 0$, so that τ must stabilize the isotypic components of σ (in the sense of Remark 4.6), and we can apply the above argument.

4.7.2 The case $\mathrm{PGL}_{\mathcal{U}}(q)$, $\mathrm{PSp}_{\mathcal{U}}(q)$, $\mathrm{PGO}_{\mathcal{U}}(q)$, and $\mathrm{PGU}_{\mathcal{U}}(q)$

Recall that $k = \mathbb{F}_q$ when G is $\mathrm{GL}_{\mathcal{U}}(q)$, $\mathrm{Sp}_{\mathcal{U}}(q)$, or $\mathrm{GO}_{\mathcal{U}}(q)$, and $k = \mathbb{F}_{q^2}$ when $G = \mathrm{GU}_{\mathcal{U}}(q)$. Set $d = 1$ in the first three cases and $d = 2$ when G is unitary over $\mathbb{F}_{q^2}$.

Recall that $Z = k^\times$ when $G = \mathrm{GL}_{\mathcal{U}}(q)$, $Z = \{\pm 1\} \subseteq k^\times$ when $G = \mathrm{Sp}_{\mathcal{U}}(q)$ or

$G = \mathrm{GO}_{\mathcal{U}}(q)$, and $Z = \{z \in k^\times \mid z^{q+1} = 1\} \subseteq k^\times = \mathbb{F}_{q^2}^\times$ when $G = \mathrm{GU}_{\mathcal{U}}(q)$. Also, recall that, if G is not of shape $\mathrm{GL}_{\mathcal{U}}(q)$, we have $z^\sigma = z^{-1}$ for $z \in Z$, where $\sigma\colon k \to k$ is the identity in the symplectic and orthogonal case, and the q-Frobenius endomorphism $x \mapsto x^q$ when $G = \mathrm{GU}_{\mathcal{U}}(q)$ (cf. the beginning of Subsection 3.4.2). Let $g = \overline{(g_i)_{i\in I}} \in G \leq \mathrm{GL}_{\mathcal{U}}(k)$ be semisimple, with g_i ($i \in I$) semisimple such that $\overline{g} \in \overline{G} \leq \mathrm{PGL}_{\mathcal{U}}(k) = \mathrm{GL}_{\mathcal{U}}(k)/k^\times$ is torsion of order dividing o, i.e., there is $\mu \in k^\times$ such that $g^o = \mu\,\mathrm{id}$. This implies $\mu \in Z$. Then

$$\sum_{\substack{\chi \text{ irreducible} \\ X^o \equiv \mu\,(\chi)}} q_\chi(g) = 1$$

by Lemma 4.4. Set $P \coloneqq \{\chi \in k[X] \mid \chi \text{ (monic) irreducible, } \chi \mid X^o - \mu\}$, $T \coloneqq \mathrm{stab}_Z(q(g))$, $K_\chi \coloneqq k[X]/(\chi)$ for $\chi \in k[X]$ irreducible (as in Subsection 3.4.2 §1), and $c_{i\chi} \coloneqq c_\chi(g_i)$ ($i \in I$). Hence, similarly to the above, we have

$$\mathbf{C}(g) = \prod_{\mathcal{U}} \prod_{\substack{\chi \text{ irreducible} \\ X^o \equiv \mu\,(\chi) \\ q_\chi(g) > 0}} \mathbf{M}_{c_{i\chi}}(K_\chi),$$

the centralizer being computed in $\mathbf{M}_{\mathcal{U}}(k)$. Now, by Section 4.6 we 'know' the structure of $\mathbf{C}_{\mathrm{conf}}(g) \leq G$. For $\chi \in k[X]$ irreducible consider the g-invariant subspace $V \coloneqq V_{\overline{\chi}}(g) \coloneqq \bigoplus_{\xi\in\overline{\chi}} V_\xi(g) \in \mathcal{V}$, where $\overline{\chi} \coloneqq \mathrm{orb}_T(\chi)$ is the orbit of χ under T (see Remark 4.7 for the definition of $V_\xi(g) \in \mathcal{V}$). Set $l_\chi \coloneqq |\overline{\chi}|$ and $m_\chi \coloneqq |T|/l_\chi$. Note that $m_\chi = |\mathrm{stab}_T(\chi)|$, and so $m_\chi = \max\{m \mid |T| \mid \exists \chi' : \chi = \chi'(X^m)\}$. The restriction of the action of $\mathbf{C}_{\mathrm{conf}}(g)/Z$ to $V_{\overline{\chi}}(g)$ is given by

$$\left(\left(\prod_{\mathcal{U}} \prod_{\xi\in\overline{\chi}} \mathbf{C}\left(g|_{V_\xi(g)}\right)\right) \rtimes T\right)/Z.$$

We will explain this below.

Definition of the action of T. In this situation $t \in T \leq Z \leq k^\times$ acts as the map φ_t which is constructed as follows: Find K_ξ-bases $(B_{\xi,i})_{i\in I}$ of each $V_\xi(g)$ ($\xi \in \overline{\chi}$ for all representatives χ of orbits of the action of T on the irreducible polynomials; see Remark 4.8) and compatible bijections $\alpha_{\xi_1,\xi_2,i}\colon B_{\xi_1,i} \to B_{\xi_2,i}$ ($i \in I$; i.e., $\alpha_{\xi_2,\xi_3,i} \circ \alpha_{\xi_1,\xi_2,i} = \alpha_{\xi_1,\xi_3,i}$ for all $\xi_1, \xi_2, \xi_3 \in \overline{\chi}$, all χ, and all $i \in I$). If G comes from groups preserving a form, we still find bijections $\bullet^*\colon B_{\xi,i} \to B_{\xi^*,i}$ such that $b^{**} = b$, the pairing f_i restricted to $K_\xi b \times K_{\xi^*} b^* \to k$ is non-singular, the pairing f_i restricted to $K_\xi b \times K_{\xi'} b'$ is zero for all $b \in B_{\xi,i}$, $b' \in B_{\xi',i}$, $b' \neq b^*$, and such that $\bullet^*$ commutes with the maps $\alpha_{\xi_1,\xi_2,i}$ ($i \in I$). Such bases exist by the classification in Subsection 3.4.2 §1. The last condition can be fulfilled, since $(\xi.t)^* = \xi^*.t$ for all $\xi \in k[X]$ and $t \in T$ by Remark 4.3. Define $\varphi_{t,i}$ by $\varphi_{t,i}|_{\langle B_{\xi,i}\rangle_{K_\xi}} : \langle B_{\xi,i}\rangle_{K_\xi} \to \langle B_{\xi.t,i}\rangle_{K_{\xi.t}}$ to be the field isomorphism $\varphi_{\xi,t}\colon K_\xi = k[X]/(\xi) \to K_{\xi.t} = k[X]/(\xi.t)$; $\overline{X} \mapsto t\overline{X}$ applied to each K_ξ-multiple of a basis vector in $B_{\xi,i}$, i.e.,

$$\varphi_{t,i}\left(\sum_{b\in B_{\xi,i}} \lambda_b b\right) = \sum_{b\in B_{\xi.t,i}} \varphi_{\xi,t}(\lambda_b)\alpha_{\xi,\xi.t,i}(b).$$

Doing this for all representatives χ of orbits of the action of T on the irreducible polynomials $\chi \in k[X]$ with $\chi \mid X^o - \mu$, this defines, up to a small error in the rank metric, a map $\varphi_{t,i}\colon k^{n_i} \to k^{n_i}$, so set φ_t to be $\overline{(\varphi_{t,i})_{i\in I}}$.

The action of T preserves the forms f_i (and Q_i; $i \in I$). Assume G is not $\mathrm{GL}_{\mathcal{U}}(q)$. Then one verifies that T preserves the forms f_i ($i \in I$): According to Subsection 3.4.2 §1 for $b \in B_{\xi,i}$ the form $f_i|_{U\times U^*} : U \times U^* := K_\xi b \times K_{\xi^*}b^* \to k$ is given as $U \times U^* \cong K_\xi \times K_{\xi^*} \ni (u, v) \mapsto \beta \operatorname{tr}_{K_\xi/k}(uv^\alpha)$ (where $\alpha\colon K_{\xi^*} \to K_\xi$ as remarked in Remark 3.33, noting that $K_\xi = R_\xi$ as $e = 1$, since g is semisimple, and β is either one or a standard non-square in $k^\times$; the latter is only needed in Case 3.1 of Subsection 3.4.2 §1 when G is orthogonal and $b = b^*$; but we can even neglect this case by Remark 3.38; so $\beta = 1$). In Case 3.2 of Subsection 3.4.2 §1, i.e., $p = 2$, so f_i is alternating and thus $b \neq b^*$, we can assume additionally that $Q(\lambda b + \mu b^*) = \lambda\mu \in k$, as all but at most one irreducible block have this shape $W(1)$ (cf. [25, page 8 and Theorem 3.1]).

Hence for $(u, v) \in K_\xi \times K_{\xi^*} \cong U \times U^*$ we obtain $f_i(u.t, v.t) = \operatorname{tr}_{K_{\xi.t}/k}(\varphi_t(u)\varphi_t(v)^\alpha) = \operatorname{tr}_{K_{\xi.t}/k}(\varphi_t(u)\varphi_t(v^\alpha)) = \operatorname{tr}_{K_{\xi.t}/k}(\varphi_t(uv^\alpha)) = \operatorname{tr}_{K_\xi/k}(uv^\alpha) = f_i(u, v)$. This holds, since the action of T commutes with α and φ_t is a field isomorphism. The former is verified as follows: Let $v \in k[X]$. Then $\varphi_t(v(\overline{X}))^\alpha = v(t\overline{X})^\alpha = v^\sigma(t^\sigma\overline{X}^{-1}) = v^\sigma(t^{-1}\overline{X}^{-1}) = \varphi_t(v^\sigma(\overline{X}^{-1})) = \varphi_t(v^\alpha)$, as desired, since by definition of Z we have $t^\sigma = t^{-1} \in Z$. Here $\overline{X}$ is the image of X in $K_\xi = k[X]/(\xi)$.

Now let us fix $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$. We want to understand the shape of h.

Step 1: h stabilizes each $V_\chi(g)$ ($\chi \in k[X]$ irreducible). Assume that $h \in \mathbf{C}^2_{\mathrm{conf}}(g|_V) \leq \mathbf{C}_{\mathrm{conf}}(g|_V)$ does not stabilize each subspace $V_\xi(g)$ of V ($\xi \in \overline{\chi}$). Write $\overline{\chi} = \{\xi_1, \ldots, \xi_l\}$ and assume that $V_{\xi_1}(g).h = V_{\xi_2}(g)$. Take $f = (M_1, M_2, *, \ldots, *) \in \mathbf{C}(g|_V) \leq \mathbf{C}_{\mathrm{conf}}(g|_V)$, where the jth component of f acts on $V_{\xi_j}(g)$ ($j = 1, \ldots, l$), then

$$f^h = h^{-1}fh = (*, M_1^h, *, \ldots, *).$$

Now there are three cases according to the classification in Subsection 3.4.2 §1: If $G = \mathrm{GL}_{\mathcal{U}}(q)$, we can take $M_2 = 1_{V_{\xi_2}(g)}$ and M_1 far away from $k^\times \mathrm{id}_{V_{\xi_1}(g)}$. Then $[f, h] = (*, M_1^h, *, \ldots, *)$ is far away from $k^\times \mathrm{id}_V$. If G is one of $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$, ξ_1 is not self-dual and $\xi_2 \neq \xi_1^*$, we can do the same as before. When $\xi_1^* = \xi_2$ in this case, we must have $M_2 = (M_1^{-\sigma})^\top$, so that $[f, h] = (*, M_1^{\sigma\top}M_1^h, *, \ldots, *)$. Again we can choose $M_1 \in \prod_{\mathcal{U}} \mathrm{GL}_{c_{i\xi_1}}(K_{\xi_1})$ such that $(M_1^\sigma)^\top M_1^h$ is far away from Z. In the last case, $\xi_1 = \xi_1^*$ is self-dual. Then again M_1 and M_2 are independent of each other and we can choose $M_2 = 1_{V_{\xi_2}(g)}$. The only restriction on M_1 is that it lies in $\prod_{\mathcal{U}} \mathrm{GU}_{c_{i\xi_1}}(K_{\xi_1})$ if

$\xi_1 \neq X \pm 1$ or G is $\mathrm{GU}_{\mathcal{U}}(q)$ (see Case 2 of Subsection 3.4.2 §1) resp. $M_1 \in \prod_{\mathcal{U}} X_{c_{i\xi_1}}(k)$ in the opposite case when $\xi_1 = X \pm 1$, where $G = X_{\mathcal{U}}(q)$ ($X = \mathrm{Sp}$ or GO; see Case 3.1 of Subsection 3.4.2 §1), so again we can choose M_1 such that $[f,h] = (*, M_1^h, *, \ldots, *)$ is far away from Z. In all cases, we get a contradiction. This shows that $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$ fixes each $V_\chi(g) \in \mathcal{V}$ ($\chi \in k[X]$ irreducible).

Assume now that $h|_{V_\chi(g)} = M.\alpha$, where α corresponds to an element of $T^{l_\chi} = \{t^{l_\chi} \mid t \in T\}$ which induces a non-trivial field automorphism on K_χ.

Step 2: The automorphism α equals the identity id_{K_χ}. Then for $\lambda \in K_\chi$ we have $(\lambda\,\mathrm{id})^h = \lambda^\alpha\,\mathrm{id} = (\lambda^\alpha\lambda^{-1})\lambda\,\mathrm{id}$. This implies that for all $\lambda \in K_\chi^\times$ stabilizing the forms f_i (or Q_i; $i \in I$) on $V_\chi(g)$ we have $\lambda^\alpha\lambda^{-1} \in Z \leq k^\times$. When $G = \mathrm{GL}_{\mathcal{U}}(q)$ or χ is not self-dual, there is no restriction on λ (of course, if G is one of $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$, then if h acts as M on $V_\chi(g)$, it must act as $(M^{-\sigma})^\top$ on $V_{\chi^*}(g)$). Hence, in this case, for each $\lambda^\times \in K_\chi$ there exists $\kappa_\lambda \in k^\times$ such that $\lambda^\alpha\lambda^{-1} = \kappa_\lambda$. However, then every vector $\lambda \in K_\chi$ is an eigenvector of the k-linear map α, which forces $\alpha = \mathrm{id}_{K_\chi}$, since $1 \in K_\chi$ is fixed, a contradiction.

In the opposite case, G is one of $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$ and χ is self-dual. Then we are in Case 2 and 3 of Subsection 3.4.2 §1. Let $\tau\colon K_\chi \to K_\chi$ be the map defined there, i.e., $\tau|_k = \sigma$ and $\tau\colon \lambda \mapsto \lambda^{-1}$, where $\lambda \in K_\chi$ is the root of χ. Then $\tau^2 = \mathrm{id}_{K_\chi}$ and $\tau = \mathrm{id}_{K_\chi}$ if and only if we are in Case 3 of Subsection 3.4.2 §1. Here, if we are in Case 2, $\mathbf{C}(g)|_{V_\chi(g)}$ is an ultraproduct of unitary groups over the field K_χ equipped with the involution τ. In Case 3, $\mathbf{C}(g)|_{V_\chi(g)}$ is an ultraproduct of symplectic resp. orthogonal groups over $K_\chi = k$. We proceed as follows: Find totally singular K_χ-subspaces $U = \overline{(U_i)}_{i\in I}, U' = \overline{(U_i')}_{i\in I}, U'' = \overline{(U_i'')}_{i\in I} \in \mathcal{V}$ of $V_\chi(g)$ (in the sense of Remark 4.9) such that $U \oplus U' = U \oplus U'' = V_\chi(g)$, $U' \cap U'' = 0$ and $\dim(U) = \dim(U') = \dim(U'') = \dim(V_\chi(g))/2$. W.l.o.g., we may assume that $\dim_{K_\chi}(U_i) = \dim_{K_\chi}(U_i') = \dim_{K_\chi}(U_i'')$ and that the restrictions $f_i|_{U_i\times U_i'}$ and $f_i|_{U_i\times U_i''}$ are non-degenerate ($i \in I$; as we may by modifying U_i, U_i', and U_i'' a little if necessary). Then define $f' = \overline{(f_i')}_{i\in I}, f'' = \overline{(f_i'')}_{i\in I} \in \mathbf{C}(g) \leq G$ such that f_i' and f_i'' act $F(\varphi)$-isotypically on U_i and such that f_i' resp. f_i'' act $F(\varphi^*)$-isotypically on U_i' resp. U_i'' ($i \in I$) for a fixed irreducible polynomial $\varphi \in K_\chi[X]$ which is not self-dual with respect to τ. Then $f'^h|_{V_\chi(g)} = z'\, f|'_{V_\chi(g)}$ and $f''^h|_{V_\chi(g)} = z''\, f''|_{V_\chi(g)}$ for $z', z'' \in Z$. Note that $q_{\varphi.z'^{-1}}(z'f'|_{V_\chi(g)}) = q_\varphi(f'|_{V_\chi(g)}) = 1/2$ and $q_{\varphi.z''^{-1}}(z''f''|_{V_\chi(g)}) = q_\varphi(f''|_{V_\chi(g)}) = 1/2$, and $\varphi.z'^{-1}$ and $\varphi.z''^{-1}$ are both also not self-dual, since $\varphi \in K_\chi[X]$ is not self-dual and $z'^{-1}, z''^{-1} \in Z$, so that $z'^{-\tau} = z'^{-\sigma} = z'$ and $z''^{-\tau} = z''^{-\sigma} = z''$, whence, e.g., $(\varphi.z'^{-1})^* = \varphi^*.z'^{-1} \neq \varphi.z'^{-1}$. Then h must stabilize the decompositions $V_\chi(g) = U \oplus U' = U \oplus U''$, so it must stabilize U. But on the h-invariant totally isotropic subspace U, we can do the same argument as above for $G = \mathrm{GL}_{\mathcal{U}}(q)$ to see that $\alpha = \mathrm{id}_{K_\chi}$.

Hence we have obtained that $h|_{V_\chi(g)} = M \in \prod_{\mathcal{U}} \mathbf{M}_{c_{i\chi}}(K_\chi)$, so that $h \in \mathbf{C}(g)$.

Step 3: We have that $h|_{V_\chi(g)} = M = \lambda\,\mathrm{id}$ for $\lambda \in K_\chi$. According to Subsection 3.4.2 §1 we can find V_i ($i \in I$) such that $\overline{(V_i)}_{i\in I} = V_\chi(g)$ such that either all V_i are totally singular

(Case 1 of Subsection 3.4.2 §1; if χ is not self-dual; this includes the case $G = \mathrm{GL}_{\mathcal{U}}(q)$) or H_i preserves a unitary form (Case 2) or a symplectic or orthogonal form (Case 3) over K_χ on V_i ($i \in I$). Note from the classification in Subsection 3.4.2 §1 that orthogonally indecomposable blocks involving a Frobenius block of size ≥ 2 are non-central in the ambient projective linear classical group. This shows that $q_\xi(M) = 0$ for all $\xi \in K_\chi[X]$ of degree ≥ 2. Assume now that there exist distinct $\lambda, \mu \in K_\chi^\times$ such that $q_{X-\lambda}(M), q_{X-\mu}(M) \geq \varepsilon > 0$. If $G = \mathrm{GL}_{\mathcal{U}}(q)$ or we are in Case 1, $F(X-\lambda) \oplus F(X-\mu) = \mathrm{diag}(\lambda, \mu) \in \mathrm{GL}_2(K_\chi)$ is mapped to a non-central element in $\mathrm{PGL}_2(K_\chi)$, so that by the assumption, since we have 'many' of these blocks, $h|_{V_\chi(g)}$ would not commute modulo scalars with all of $\mathbf{C}(g)|_{V_\chi(g)} \cong \prod_{\mathcal{U}} \mathbf{M}_{c_{i\chi}}(K_\chi)$. In Case 2, we use the same argument for a block of shape $\mathrm{diag}(\lambda, \lambda^{-\tau}, \mu, \mu^{-\tau})$ acting on a four-dimensional (K_χ, τ)-unitary space. In Case 3, we use the same argument with a block $\mathrm{diag}(\lambda, \lambda^{-1}, \mu, \mu^{-1})$ acting on a four-dimensional symplectic or orthogonal space. In total we get that $M = \lambda\,\mathrm{id}$ for $\lambda \in K_\chi$. If we are in Case 2 of Subsection 3.4.2 §1, we have the additional assumption that $\mathrm{N}_\tau(\lambda) = 1$, where $\mathrm{N}_\tau \colon K_\chi \to K_{\chi,\tau}$ is the norm defined there. In Case 3 of Subsection 3.4.2 §1, we must have $\lambda^2 = 1$.

Step 4: The precise shape of $C := \mathbf{C}^2_{\mathrm{conf}}(g)$. We know now that $h|_{V_\chi(g)} = \lambda_\chi(h)\,\mathrm{id}$ for each irreducible $\chi \in k[X]$ and so h commutes with all of $\mathbf{C}(g)$. In order that $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$, we still need to check that $[h, T] \subseteq Z$. Now choose $t \in T$ to be a generator and $z \in Z$ and assume that h z-commutes with t, i.e., $[h, \varphi_t] = z\,\mathrm{id}$. Let $\chi \in k[X]$ run through a system of representatives of the orbits of the action of T and $\bullet^*$ on the irreducible polynomials (the action of $\bullet^*$ is only used when G is not $\mathrm{GL}_{\mathcal{U}}(q)$). This means $zh = h^{\varphi_t}$, so since $h|_{V_\chi(g)} = \lambda_\chi(h)\,\mathrm{id}$, we must have $h|_{V_{\chi.t}(g)} = z^{-1}\varphi_{\chi,t}(\lambda_\chi(h))\,\mathrm{id}_{V_{\chi.t}}$, so that h is determined on all of $V = V_{\overline{\chi}}(g)$ by $\lambda_\chi(h)$. In this situation the only condition that needs to be satisfied is that $h|_{V_\chi(g)} = h|_{V_{\chi.t^{l_\chi}}(g)} = \lambda_\chi(h)\,\mathrm{id} = z^{-l_\chi}\varphi_{\chi,t^{l_\chi}}(\lambda_\chi(h))\,\mathrm{id}$. Note that

$$\varphi_{\chi,t^l_\chi} \colon K_\chi \cong \mathbb{F}_{q^{dk_\chi}} \to K_\chi \cong \mathbb{F}_{q^{dk_\chi}} \quad (d = 1, 2)$$

is given by $x \mapsto q^{dk_\chi/m_\chi}$, so that the previous condition becomes

$$z^{l_\chi} = (\lambda_\chi(h))^{q^{dk_\chi/m_\chi} - 1}. \tag{4.3}$$

Hence we can write C as follows. When $G = \mathrm{GL}_{\mathcal{U}}(q)$, we have

$$C = \left\{ h = \bigoplus_{\substack{\chi \text{ irreducible} \\ X^o \equiv \mu\,(\chi) \\ q_\chi(g) > 0}} \lambda_\chi(h)\,\mathrm{id}_{V_\chi(g)} \;\middle|\; \exists z \in Z : \lambda_{\chi.t}(h) = z^{-1}\varphi_{\chi,t}(\lambda_\chi(h)) \text{ for all } \chi \right\}. \tag{4.4}$$

Here the condition from Equation (4.4) is equivalent to Equation (4.3) for χ running through a system of representatives of the action of T on the set $P \coloneqq \{\chi \in k[X]$ irreducible $\mid \chi$ divides $X^o - \mu$ and $q_\chi(g) > 0\}$. For G one of $\mathrm{Sp}_{\mathcal{U}}(q)$, $\mathrm{GO}_{\mathcal{U}}(q)$, or $\mathrm{GU}_{\mathcal{U}}(q)$ we have

$$C = \left\{ \bigoplus_{\substack{\chi \text{ irreducible} \\ X^o \equiv \mu\,(\chi) \\ q_\chi(g)>0}} \lambda_\chi(h)\, \mathrm{id}_{V_\chi(g)} \,\middle|\, R \right\}. \tag{4.5}$$

where the condition R is that there exists $z \in Z$ such that $\lambda_{\chi.t}(h) = z^{-1}\varphi_{\chi,t}(\lambda_\chi(h))$ for all $\chi \in P$ (as in the previous case) and $\lambda_\chi(h)\lambda_{\chi^*}(h)^\alpha = 1$ for all $\chi \in P$, where $\alpha\colon R_{\chi^*} = K_{\chi^*} \to R_\chi = K_\chi$ is defined as in Remark 3.33. If G is one of $\mathrm{Sp}_{\mathcal{U}}(q)$ or $\mathrm{GO}_{\mathcal{U}}(q)$ and $\chi = \chi^* \neq X \pm 1$ is self-dual, this means k_χ is even and $\lambda_\chi(h)^{q^{k_\chi/2}+1} = 1$. Also, in this case, if $\chi = X \pm 1$ it means $\lambda_\chi(h)^2 = 1$. If $G = \mathrm{GU}_{\mathcal{U}}(q)$ and $\chi = \chi^*$, this means that k_χ is odd (since α needs to induce $\sigma : x \mapsto x^q$ on $k = \mathbb{F}_{q^2}$) and $\lambda_\chi(h)^{q^{k_\chi}+1} = 1$.

4.8 Distinction of metric ultraproducts

Now we want to distinguish all (simple) metric ultraproducts $\overline{G} = \overline{X}_{\mathcal{U}}(q)$ for distinct pairs (X, q), where $X \in \{\mathrm{GL}, \mathrm{Sp}, \mathrm{GO}, \mathrm{GU}\}$ and q is a prime power (all but $\mathrm{PSp}_{\mathcal{U}_1}(q)$ and $\mathrm{PGO}_{\mathcal{U}_2}(q)$ as mentioned in Theorem 4.1). For a group H define the quantity

$$e_H(o) \coloneqq \max_{h \in H : h^o = 1_H} \exp(\mathbf{C}^2(h)).$$

Clearly, when $H \cong L$, we have $e_H(o) = e_L(o)$ for all values $o \in \mathbb{Z}_+$. Our strategy is to compute $e_H(o)$ for the groups $H = \overline{G}$, where $\overline{G} = \overline{X}_{\mathcal{U}}(q)$ as above, for certain values of o to distinguish these groups (with the only exception: $\mathrm{PSp}_{\mathcal{U}_1}(q) \cong \mathrm{PGO}_{\mathcal{U}_2}(q)$?).

4.8.1 Computation of $e_{\overline{G}}(o)$ when $\gcd\{o, p\} = \gcd\{o, |Z|\} = 1$

If o is coprime to $|Z|$ (and by semisimplicity of $g \in G$ coprime to p), from Subsection 4.7.2 we can compute $e_{\overline{G}}(o)$. Note that in this situation, when $g^o = \mu \in Z$, we can replace g by $g' = \lambda g \in G$ such that $g'^o = 1$, choosing $\lambda \in Z$ such that $\lambda^o = \mu^{-1}$, since the homomorphism $Z \to Z$; $x \mapsto x^o$ is then bijective. So assume, w.l.o.g., $g^o = 1$. Then $P \subseteq Q \coloneqq \{\chi \in k[X]$ irreducible $\mid \chi$ divides $X^o - 1\}$.

The case $G = \mathrm{GL}_{\mathcal{U}}(q)$. From Equation (4.4) we see that, the bigger the group T is, for an element $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$, the more restrictions are imposed to the scalars $\lambda_\chi(h) \in K_\chi^\times$ ($\chi \in P$). Also, the bigger the set P is, the 'bigger' is the group $\mathbf{C}^2_{\mathrm{conf}}(g)$, i.e., there are more components. Hence, to optimize the exponent of $\mathbf{C}^2(\overline{g}) = \mathbf{C}^2_{\mathrm{conf}}(g)/Z$, we choose g such that $P = Q$ and $0 < q_{X-1}(g) \neq q_\chi(g) > 0$ for all $\chi \in P \setminus \{X - 1\}$. Namely, then $T = \mathrm{stab}_Z(q(g))$ must fix the polynomial $X - 1$, so that we must have $T = \mathbf{1}$. Set

$$f_q(o) \coloneqq \min\{q^e - 1 \mid o \text{ divides } q^e - 1\}.$$

Equation (4.4) then gives

$$e_{\overline{G}}(o) = \exp(\mathbf{C}^2(\overline{g})) = \exp(\mathbf{C}^2_{\mathrm{conf}}(g)/Z) = \begin{cases} 1 & \text{if } o = 1 \\ f_q(o) & \text{if } o > 1 \end{cases}. \tag{4.6}$$

Let us demonstrate Equation (4.6). The first equality in it holds by the previous argument. When $o = 1$, we have $\overline{g} = 1_{\overline{G}}$ and so

$$\mathbf{C}^2(\overline{g}) = \mathbf{Z}(\overline{G}) = \mathbf{1},$$

so that $e_{\overline{G}}(1) = 1$. Now assume $o > 1$. For each $\chi \in P$, if $\lambda \in \overline{k}^{\times}$ is a root of χ, the condition that $\chi \mid X^o - 1$ is equivalent to $\lambda^o = 1$. Also $K_\chi = k[\lambda]$. Let $\mu \in \overline{k}^{\times}$ be an element of order o with minimal polynomial $\xi \in k[X]$. Then, if λ is a root of $\chi \in P$, we must have $\lambda^o = 1$ and thus $\lambda = \mu^f$ for some $f \in \mathbb{N}$. Hence $K_\chi = k[\lambda] = k[\mu^f] \subseteq k[\mu] = K_\xi$, so that in Equation (4.4) we have $\mathrm{ord}(\lambda_\chi(h)) \mid |K_\chi^\times| \mid |K_\xi^\times| = f_q(o)$. This shows that

$$\exp(\mathbf{C}^2(\overline{g})) \mid \exp(\mathbf{C}^2_{\mathrm{conf}}(g)) \mid \mathrm{lcm}\{|K_\chi^\times| \mid \chi \text{ divides } X^o - 1\} = |K_\xi^\times| = |k[\mu]^\times| = f_q(o).$$

To show the equality $\exp(\mathbf{C}^2(\overline{g})) = f_q(o)$, take $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$ such that $\lambda_{X-1}(h) = 1$ and $\lambda_\xi(h)$ has order $f_q(o) = |K_\xi^\times|$ in $\overline{k}^{\times}$. Then, when $\overline{h}^l = 1_{\overline{G}}$, we must have $h^l \in Z$. But $\lambda_{X-1}(h)^l = 1$, so that, since $q_{X-1}(g) > 0$, it follows that $h^l = 1_G$. Then $\lambda_\xi(h)^l = 1$, so that $\exp(\mathbf{C}^2(\overline{g})) \geq l \geq \mathrm{ord}(\lambda_\xi(h)) = f_q(o)$. This completes the proof.

The case $G = \mathrm{Sp}_{\mathcal{U}}(q)$ or $\mathrm{GO}_{\mathcal{U}}(q)$. As in the linear case, Equation (4.5) shows that the optimal exponent of $\mathbf{C}^2(\overline{g})$ is obtained when $P = Q$ and $0 < q_{X-1}(g) \neq q_\chi(g) > 0$ for all $\chi \in P \setminus \{X - 1\}$, so that $T = \mathbf{1}$. Set

$$f'_q(o) := \begin{cases} q^{e/2} + 1 & \text{if } f_q(o) = q^e - 1, e \text{ is even and } o \mid q^{e/2} + 1 \\ f_q(o) & \text{otherwise} \end{cases}. \tag{4.7}$$

Equation (4.5) then gives

$$e_{\overline{G}}(o) = \exp(\mathbf{C}^2(\overline{g})) = \exp(\mathbf{C}^2_{\mathrm{conf}}(g)/Z) = \begin{cases} 1 & \text{if } o = 1 \\ 2 & \text{if } o = 2 \\ f'_q(o) & \text{if } o > 2 \end{cases}. \tag{4.8}$$

We demonstrate Equation (4.8). If $o = 1$, we obtain, as in the linear case, that $\mathbf{C}^2(\overline{g}) = \mathbf{1}$ and so $e_{\overline{G}}(1) = 1$. If $o = 2$, $g^2 = 1$ and so $P = \{X - 1, X + 1\}$. From Equation (4.5) we see that, if $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$, we have $\lambda_{X-1}(h)^2 = \lambda_{X+1}(h)^2 = 1$, so that $h^2 = 1$. Also, defining h by $\lambda_{X-1}(h) := 1$ and $\lambda_{X+1}(h) = -1$, we obtain $h \notin Z$, so $\mathrm{ord}_{\overline{G}}(\overline{h}) = e_{\overline{G}}(2) = 2$ ($1 \neq -1$, since the case $p = 2$ does not occur due to the condition $\gcd\{o, p\} = 1$). Assume now that $o > 2$. As in the linear case, for each $\chi \in P$, if $\lambda \in \overline{k}$ is a root of χ, the condition that

$\chi \mid X^o - 1$ is equivalent to $\lambda^o - 1$. Choose $\mu \in \overline{k}^\times$ of order o and let $\xi \in P$ be its minimal polynomial. Then, as previously, if λ is a root of $\chi \in P$, we have $\lambda = \mu^f$ for some $f \in \mathbb{N}$. There are two cases:

In the first case, ξ is not self-dual. This means that μ and μ^{-1} are not conjugate in K_ξ/k. If they were conjugate, say by an automorphism α, i.e., $\mu^\alpha = \mu^{-1}$, then $\alpha \in \mathrm{Gal}(K_\xi/k)$ needs to be the unique involution (since $\mu \neq \mu^{-1}$ as $o > 2$) given by $x \mapsto x^{q^{k_\xi/2}}$; in particular, $e = k_\xi$ would need to be even. Hence this case is equivalent to either $e = k_\xi$ being odd or $\mu\mu^\alpha = \mu^{q^{k_\xi/2}+1} \neq 1$, i.e., $o \nmid q^{e/2} + 1 = q^{k_\xi/2} + 1$. This is precisely the opposite of the first case in Equation (4.7). Here for an element $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$ we can choose $\lambda_\xi(h) \in K_\xi^\times = k[\mu]^\times$ arbitrarily ($\lambda_{\xi^*}(h)$ is then determined by $\lambda_\xi(h)$). Arguing as in the linear case, we obtain $\exp(\mathbf{C}^2(\overline{g})) = f_q(o)$. Indeed, for $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$, as above, $\mathrm{ord}(\lambda_\chi(h)) \mid f_q(o)$ and defining h such that $\lambda_{X-1}(h) = 1$ and $\lambda_\xi(h)$ has order $f_q(o)$, we see that $\mathrm{ord}_{\overline{G}}(\overline{h}) = e_{\overline{G}}(o) = f_q(o)$.

In the opposite case, ξ is self-dual and $\xi \neq X \pm 1$ as $o > 2$. Then $e = k_\xi$ needs to be even and
$$\mu\mu^\alpha = \mu^{q^{k_\xi/2}+1} = \lambda_\xi(h)\lambda_\xi(h)^\alpha = \lambda_\xi(h)^{q^{k_\xi/2}+1} = 1,$$
where α is the involution $x \mapsto x^{q^{k_\xi/2}}$ of $K_\xi \cong \mathbb{F}_{q^{k_\xi}}$ from Subsection 3.4.2 §1 Case 2. This means that $o \mid q^{e/2} + 1$ and we are in the first case of Equation (4.7). Note that for each $\chi \in P$ the map α restricts to an automorphism of each $K_\chi \subseteq K_\xi$ of order dividing two (as all the fields are finite). Then $\alpha|_{K_\chi} = \mathrm{id}$ if and only if $k_\chi \mid k_\xi/2$, and $\alpha|_{K_\chi}$ is the unique involution of K_χ if k_ξ/k_χ is odd. Now, if $\lambda \in \overline{k}^\times$ is a root of χ, then in the first case $\lambda^2 = 1$, and in the second case $\lambda^{q^{k_\chi/2}+1} = 1$; so all $\chi \in P$ are self-dual. Hence, if $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$, for each $\chi \in P$ one of $\lambda_\chi(h)^2 = 1$ or $\lambda_\chi(h)^{q^{k_\chi/2}+1} = 1$ must hold. But $2 \mid q^{k_\xi/2} + 1$ if $p > 2$, and $q^{k_\chi/2} + 1 \mid q^{k_\xi/2} + 1$ in the second case, since k_ξ/k_χ is then odd. Hence $\exp(\mathbf{C}^2_{\mathrm{conf}}(g)) \mid q^{e/2} + 1 = q^{k_\xi/2} + 1 = f'_q(o)$. Defining $h \in \mathbf{C}^2_{\mathrm{conf}}(g)$ such that $\lambda_{X-1}(h) = 1$ and $\lambda_\xi(h)$ has order $f'_q(o)$, we see that $\mathrm{ord}_{\overline{G}}(\overline{h}) = f'_q(o)$.

The case $G = \mathrm{GU}_{\mathcal{U}}(q)$. Here, as well, Equation (4.5) shows that the optimal exponent of $\mathbf{C}^2(\overline{g})$ is obtained when $P = Q$ and $0 < q_{X-1}(g) \neq q_\chi(g) > 0$ for all $\chi \in P \setminus \{X - 1\}$, so that $T = \mathbf{1}$. Set
$$f''_q(o) := \begin{cases} q^e + 1 & \text{if } f_{q^2}(o) = q^{2e} - 1,\ e \text{ is odd and } o \mid q^e + 1 \\ f_{q^2}(o) & \text{otherwise} \end{cases}.$$

Equation (4.5) gives
$$e_{\overline{G}}(o) = \exp(\mathbf{C}^2(\overline{g})) = \exp(\mathbf{C}^2_{\mathrm{conf}}(g)/Z) = \begin{cases} 1 & \text{if } o = 1 \\ f''_q(o) & \text{if } o > 1 \end{cases}. \tag{4.9}$$

Again $e_{\overline{G}}(1) = 1$ is clear. If $o > 1$, take $\mu \in \overline{k}^\times$ of order o with minimal polynomial ξ. Then the argument proceeds as in the bilinear case. But the condition that ξ is self-dual is here equivalent to μ being conjugate to μ^{-1} in K_ξ by an automorphism α such that $\alpha|_k = \sigma$; $x \mapsto x^q$. This forces $e = k_\xi$ to be odd and $\mu^{q^{dk_\xi/2}+1} = 1$, i.e., $o \mid q^e + 1 = q^{k_\xi} + 1$.

4.8.2 Proof of Theorem 4.1

Set $G_j := X_{j\mathcal{U}_j}(q_j)$, $Z_j := \mathbf{Z}(G_j)$, and $\overline{G}_j := G_j/Z_j$ $(j = 1, 2)$. Let p_j be the characteristic of the field $\mathbb{F}_{q_j}$ $(j = 1, 2)$. Assume that $\overline{G} \cong \overline{G}_1 \cong \overline{G}_2$. We start by showing that $p_1 = p_2$.

Determining the characteristic p. Choose o large enough and coprime to p_1, p_2, $|Z_1|$, $|Z_2|$. Then from Equations (4.6), (4.8), and (4.9) we see that $e_{\overline{G}}(o)$ is of the form $q_1^{e_1} \pm 1$ and $q_2^{e_2} \pm 1$. If $e_{\overline{G}}(o) = q_1^{e_1} - 1 = q_2^{e_2} - 1$ or $e_{\overline{G}}(o) = q_1^{e_1} + 1 = q_2^{e_2} + 1$, we have $q_1^{e_1} = q_1^{e_2}$, so that $p_1 = p_2$ by the uniqueness of the prime factorization. So, w.l.o.g., we have $e_{\overline{G}}(o) = q^{e_1} - 1 = q_2^{e_2} + 1$ for infinitely many o, and so for infinitely many pairs $(e_1, e_2) \in \mathbb{Z}_+^2$. If $p_1 \neq p_2$ we get a contradiction to Corollary 1.8 of [18]. Hence $p_1 = p_2 =: p$.

Determining q^d. We can now assume that $q_j = p^{e_j}$ $(j = 1, 2)$. Choose $j \in \{1, 2\}$ and set $X := X_j$, $q := q_j$, and $d := d_j$. Consider the quantity $f := \gcd\{e_{\overline{G}}(o) \,|\, o \in O\}$, where

$$O := \{o \in \mathbb{Z}_+ \,|\, 2 < o \text{ coprime to } p, |Z_1|, |Z_2|;\ e_{\overline{G}}(o) \equiv -1 \text{ modulo } p^3\}$$

From Equations (4.6), (4.8), and (4.9) it follows that for every element $o \in O$ the number $e_{\overline{G}}(o)$ is either of the form $q^{de} - 1$ or $q^e + 1$. But the second case is excluded by the condition that $e_{\overline{G}}(o) \equiv -1$ modulo p^3. Hence $q^d - 1 \mid e_{\overline{G}}(o) = q^{de} - 1$ and so $q^d - 1 \mid f$.

For a prime r set $t_r := \frac{q^r-1}{q-1}$. Then for distinct primes r and s we have

$$\gcd\{t_r, t_s\} = \gcd\left\{\frac{q^r-1}{q-1}, \frac{q^s-1}{q-1}\right\} = \frac{1}{q-1}\gcd\{q^r-1, q^s-1\} = \frac{q^{\gcd\{r,s\}}-1}{q-1} = 1.$$

Hence the numbers t_r (r prime), being pairwise coprime, have arbitrarily large prime divisors. Take for $r > 2$ a prime such that t_r has a prime divisor $o > p, |Z_1|, |Z_2|, q^d - 1$. Then o is coprime to p, $|Z_1|$, and $|Z_2|$, so that by Equations (4.6), (4.8), and (4.9) we have $e_{\overline{G}}(o) \mid f_{q^d}(o) \mid q^{dr} - 1$, as $o \mid q^r - 1 \mid q^{dr} - 1$. Hence the number $f_{q^d}(o)$ must be one of $q^{dr} - 1$ or $q^d - 1$, the latter being excluded by the condition $o > q^d - 1$; so $f_{q^d}(o) = q^{dr} - 1$. If $X = \mathrm{GL}$, $X = \mathrm{Sp}$, or $X = \mathrm{GO}$, since r is odd and $d = 1$, Equations (4.6) and (4.8) show that we must have $e_{\overline{G}}(o) = f_{q^d}(o) = q^{dr} - 1 = q^r - 1 \equiv -1$ modulo p^3. Hence, in this case, $o \in O$. If $X = \mathrm{GU}$, it could be that $e_{\overline{G}}(o) = q^r + 1$, when $o \mid q^r + 1$. However, $\gcd\{q^r + 1, t_r\} \mid \gcd\{q^r + 1, q^r - 1\} \mid 2$ and t_r is always odd, so that $\gcd\{q^r + 1, t_r\} = 1$ and hence, as $o \mid t_r$, also $\gcd\{q^r + 1, o\} = 1$. This shows that here also $e_{\overline{G}}(o) = f_{q^d}(o) = q^{dr} - 1 = q^{2r} - 1 \equiv -1$ modulo p^3. Therefore again $o \in O$.

Applying this argument for two different primes r, say r_1 and r_2, which produces two different primes o, say o_1 and o_2, we get $f = \gcd\{e_{\overline{G}}(o) \,|\, o \in O\} \mid \gcd\{e_{\overline{G}}(o_1), e_{\overline{G}}(o_2)\} = \gcd\{q^{dr_1} - 1, q^{dr_2} - 1\} = q^{\gcd\{dr_1, dr_2\}} - 1 = q^d - 1$.

Altogether, we have shown that $f = q^d - 1$. Plugging in $j = 1, 2$, we obtain $q_1^{d_1} - 1 =$

$q_2^{d_2} - 1$ implying that $q_1^{d_1} = q_2^{d_2}$.

Now we exclude all remaining possible isomorphisms but $\mathrm{PSp}_{\mathcal{U}_1}(q) \cong \mathrm{PGO}_{\mathcal{U}_2}(q)$.

Proof that $\mathrm{PGL}_{\mathcal{U}_1}(q) \not\cong \mathrm{PSp}_{\mathcal{U}_2}(q)$ *and* $\mathrm{PGL}_{\mathcal{U}_1}(q) \not\cong \mathrm{PGO}_{\mathcal{U}_2}(q)$. Let $G_1 = \mathrm{GL}_{\mathcal{U}_1}(q)$ and $G_2 = X_{\mathcal{U}_2}(q)$, where $X = \mathrm{Sp}$ or GO. Set

$$o := \begin{cases} \frac{q^2+1}{2} & \text{if } p > 2 \\ q^2+1 & \text{if } p = 2 \end{cases}. \tag{4.10}$$

Note that $o > 2$ is coprime to p, $|Z_1| = q-1$, and $|Z_2| = |\{\pm 1\}|$. Hence by Equation (4.8) we have $e_{\overline{G}}(o) = e_{\overline{G}_2}(o) = q^2+1$. Indeed, $o \mid q^2+1 \mid q^4-1$. But $o \nmid q^f - 1$ for f properly dividing 4, since then $o \mid q^2-1$, but it is easy to see that $\gcd\{o, q^2-1\} = 1$. This shows $f_q(o) = q^4 - 1$ and $e_{\overline{G}}(o) = e_{\overline{G}_2}(o) = f'_q(o) = q^2+1$. But then by $\overline{G}_1 \cong \overline{G}_2$ we obtain $e_{\overline{G}}(o) = e_{\overline{G}_1}(o) = f_q(o) = q^4 - 1 > q^2+1 = e_{\overline{G}_2}(o) = e_{\overline{G}}(o)$, a contradiction.

Proof that $\mathrm{PSp}_{\mathcal{U}_1}(q^2) \not\cong \mathrm{PGU}_{\mathcal{U}_2}(q)$ *and* $\mathrm{PGO}_{\mathcal{U}_1}(q^2) \not\cong \mathrm{PGU}_{\mathcal{U}_2}(q)$. Let $G_1 = X_{\mathcal{U}_1}(q^2)$, where $X = \mathrm{Sp}$ or GO, and $G_2 = \mathrm{GU}_{\mathcal{U}_2}(q)$. Define o as in Equation (4.10). Note that $o > 2$ is coprime to p, $|Z_1| = |\{\pm 1\}|$, and $|Z_2| = q+1$. Then by Equation (4.8) we have $e_{\overline{G}}(o) = e_{\overline{G}_1}(o) = f'_{q^2}(o) = q^2+1$ (as above). But by Equation (4.9) we obtain that $e_{\overline{G}} = e_{\overline{G}_2}(o) = f''_q(o) = q^4 - 1 > q^2 + 1 = e_{\overline{G}_1}(o) = e_{\overline{G}}(o)$, since $e = 2$ is even, a contradiction.

Proof that $\mathrm{PGL}_{\mathcal{U}_1}(q^2) \not\cong \mathrm{PGU}_{\mathcal{U}_2}(q)$. Let $G_1 = \mathrm{GL}_{\mathcal{U}_1}(q^2)$ and $G_2 = \mathrm{GU}_{\mathcal{U}_2}(q)$. Set

$$o := \begin{cases} \frac{q^5+1}{5(q+1)} & \text{if } q \equiv -1 \text{ modulo } 5 \\ \frac{q^5+1}{q+1} & \text{otherwise} \end{cases}.$$

Note that o is coprime to p, $|Z_1| = q^2-1$, and $|Z_2| = q+1 \mid q^2-1$. Indeed, $\gcd\{o, q+1\} \mid \gcd\{\frac{q^5+1}{q+1}, q+1\} = \gcd\{5, q+1\} \mid 5$. But $5 \nmid o$, so that $\gcd\{o, q+1\} = 1$. Similarly, $\gcd\{o, q-1\} \mid \gcd\{q^5+1, q-1\} \mid \gcd\{2, q-1\} \mid 2$. But o is always odd, so $\gcd\{o, q-1\} = 1$. We have that $o \mid q^{10}-1$, so that from Equation (4.6) we obtain that $e_{\overline{G}}(o) = e_{\overline{G}_1}(o) = f_{q^2}(o)$ is either $q^{10}-1$ or q^2-1. But clearly $q^2 - 1 < o$, so that we must have $f_{q^2}(o) = q^{10}-1$. But Equation (4.9) gives that $e_{\overline{G}}(o) = e_{\overline{G}_2}(o) = f''_q(o) = q^5+1 < q^{10}-1 = f_q(o) = e_{\overline{G}_1}(o)$, a contradiction.

Remark 4.18. If $q_i \to_{\mathcal{U}} \infty$, then double centralizers of semisimple torsion elements are infinite groups.

Remark 4.19. If q is even, then $\mathrm{PSp}_{\mathcal{U}_1}(q) \cong \mathrm{PGO}_{\mathcal{U}_2}(q)$ is possible due to the isomorphism $\mathrm{Sp}_{2m}(q) \cong \mathrm{GO}_{2m+1}(q)$. Also it seems hard to distinguish a group $\mathrm{PSp}_{\mathcal{U}_1}(q)$ from a group $\mathrm{PGO}_{\mathcal{U}_2}(q)$ for q odd.

Index of Symbols

Set theory

Symbol	Explanation
$x \in X$	x is an element of the set X
$x \notin X$	x is not an element of the set X
$A \subseteq B$	A is a subset of B
$A \subset B$	A is a proper subset of B
$A \supseteq B$	A is a superset of B
$A \supset B$	A is a proper superset of B
$\lvert X \rvert$	the cardinality of the set X
$A \cap B$	the intersection of the sets A and B
$\bigcap \mathcal{S}$	the intersection of the sets in $\mathcal{S}$
$\bigcap_{i \in I} S_i$	the intersection of the sets S_i $(i \in I)$
$A \cup B$	the union of the sets A and B
$\bigcup \mathcal{S}$	the union of the sets in $\mathcal{S}$
$\bigcup_{i \in I} S_i$	the union of the sets S_i $(i \in I)$
$A \sqcup B$	the disjoint union of the sets A and B
$\bigsqcup \mathcal{S}$	the disjoint union of the sets in $\mathcal{S}$
$\bigsqcup_{i \in I} S_i$	the disjoint union of the sets S_i $(i \in I)$
$A \setminus B$	the difference of the sets A and B
$A \triangle B$	the symmetric difference of the sets A and B
$A \times B$	the Cartesian product of the sets A and B
$\prod \mathcal{S}$	the Cartesian product of the sets in $\mathcal{S}$
$\prod_{i \in I} S_i$	the Cartesian product of the sets S_i $(i \in I)$
X^n	the nth power of the set X
$\mathcal{P}(X)$	the power set of the set X
$f\colon A \to B$	f is a map from the set A to the set B
$f\colon A \hookrightarrow B$	f is an injective map from the set A to the set B
$f\colon A \twoheadrightarrow B$	f is a surjective map from the set A to the set B
$f(x)$	the value at x of the function f
$X/\sim$	the quotient of the set X by the equivalence relation $\sim$

Standard sets

Symbol	Explanation
$\mathbb{N}$	the set of natural numbers $\{0, 1, 2, \ldots\}$
$\underline{n}$	the set $\{1, \ldots, n\}$ ($n \in \mathbb{N}$)
$\mathbb{Z}$	the set of integers $\{\ldots, -2, -1, -0, 1, 2, \ldots\}$
$\mathbb{Z}_+$	the set of positive integers $\{1, 2, \ldots\}$
$\mathbb{Q}$	the set of rational numbers $\{p/q \mid p \in \mathbb{Z}, q \in \mathbb{Z}_+\}$
$\mathbb{R}$	the set of real numbers
$\mathbb{C}$	the set of complex numbers

Arithmetic

Symbol	Explanation
$\log(x)$	the natural logarithm of x
$\min S$	the minimum of the set S
$\max S$	the maximum of the set S
$a \leq b$	a is less than or equal to b
$a < b$	a is less than b
$a \geq b$	a is greater than or equal to b
$a > b$	a is greater than b
$\|x\|$	the absolute value of $x \in \mathbb{C}$
a^b	a raised to the power b
$\sum_{i \in I} x_i$	sum of the x_i ($i \in I$)
$\sup_{i \in I} x_i$	supremum of the x_i ($i \in I$)
$\gcd S$	greatest common divisor of the elements of S
$\operatorname{lcm} S$	least common multiple of the elements of S

Basic group theory

Symbol	Explanation
1_G	the neutral element of a group G
g^{-1}	the inverse of the group element g
gh	the product of the group elements g and h
g^h	the conjugate of the group element g by the group element h; $g^h = h^{-1}gh$
$\overline{g}$	the image of a group element g under a natural homomorphism

g^G	the conjugacy class of the group element g inside G; $g^G = \{g^h \mid h \in G\}$
$\mathrm{ord}(g)$	the order of the group element g
$\lvert G\rvert$	the order of the group G
$\exp(G)$	the exponent of the group G
$\mathbf{1}$	the trivial group
C_k	the cyclic group of order $k \in \mathbb{Z}_+$
$G \cong H$	the groups G and H are isomorphic
$\mathrm{Aut}(G)$	the automorphism group of the group G
$H \leq G$	H is a subgroup of G
$N \trianglelefteq G$	N is a normal subgroup of G
$H \cap I$	the intersection of the subgroups H and I
$G \times H$	the direct product of the groups G and H
$H \wr G$	the wreath product of the groups H and G
$N.H$	an extension of the group N by the group H
$\prod \mathcal{H}$	the direct product of the groups in $\mathcal{H}$
$\prod_{i\in I} H_i$	the direct product of the groups H_i $(i \in I)$
$\varphi\colon G \to H$	φ is a group homomorphism from G to H
$\varphi\colon G \hookrightarrow H$	φ is an injective group homomorphism from G to H
$\varphi\colon G \twoheadrightarrow H$	φ is a surjective group homomorphism from G to H
$\mathrm{im}(\varphi)$	the image of the homomorphism φ
$\ker(\varphi)$	the kernel of the homomorphism φ
$\overline{G} = G/N$	the quotient of the group G by its normal subgroup N
$\langle S\rangle$	the subgroup generated by the set S (inside an ambient group)
$\langle\!\langle S\rangle\!\rangle, \langle\!\langle S\rangle\!\rangle_G$	the normal subgroup generated by the set S (inside the group G)
S^{*k}	the k-fold product of $S \subseteq G$; $S^{*k} = \{s_1 \cdots s_k \mid s_1, \dots, s_k \in S\}$
ST	the product of $S, T \subseteq G$; $ST = \{st \mid s \in S, t \in T\}$
$\mathbf{F}$	a free group (of arbitrary rank)
$\mathrm{rk}(\mathbf{F})$	the rank of a free group
$\mathbf{F}_r$	the free group of rank r (freely generated by $x_1, \dots, x_r$)
$\langle S \mid R\rangle$	the group generated by the set S with relations R

Group actions

Symbol	Explanation
$x.g$	the image of the point x under the action of the group element g
$\mathrm{stab}_G(x)$	the stabilizer of the point x in the group G
$\mathrm{orb}_G(x)$	the orbit of the point x under the group G

Commutators

Symbol	Explanation
$[g,h]$	the commutator of the group element g and h; $[g,h]=g^{-1}h^{-1}gh$
$[S,g]$	the subset $\{[s,g] \mid s \in S\}$ of an ambient group
$[g,S]$	the subset $\{[g,s] \mid s \in S\}$ of an ambient group
$[H,L]$	the subgroup generated by all commutators $[h,l]$, $h \in H$, $l \in L$

Subgroups

Symbol	Explanation
$\mathbf{Z}(G)$	the center of the group G
$\mathbf{C}(S), \mathbf{C}_G(S)$	the centralizer of the set S (inside the group G)
$G', G^{(1)}$	the commutator subgroup of G
$G^{(i)}$	the ith term in the derived series of G ($i \in \mathbb{N}$)
$\gamma_i(G)$	the ith term in the lower central series of G ($i \in \mathbb{Z}_+$)
$\gamma_\omega(G)$	the subgroup $\bigcap_{i\in\mathbb{Z}_+} \gamma_i(G)$ of G

Permutations

Symbol	Explanation
$\mathrm{Sym}(\Omega)$	the symmetric group on the finite set Ω
$\mathrm{Alt}(\Omega)$	the alternating group on the finite set Ω
S_n	the symmetric group of degree n
A_n	the alternating group of degree n
$\widetilde{\mathrm{A}}_n$	the Schur covering group of the finite alternating group A_n
$\mathrm{id}, \mathrm{id}_\Omega$	the identity permutation (on the set Ω)
$\mathrm{supp}(\sigma)$	the support of the permutation σ
$c_k(\sigma)$	the number of k-cycles of the permutation σ
$(k^{c_k})_{k\in\mathbb{Z}_+}$	the cycle type of a permutation
$C_k(\sigma)$	the set of k-cycles of the permutation σ
$\Omega_k(\sigma)$	the support of all the k-cycles of the permutation σ
$n_k(\sigma)$	the size of $\Omega_k(\sigma)$; $n_k(\sigma)=kc_k(\sigma)$

Vector spaces and linear maps

Symbol	Explanation
k	the ground field
0	the zero vector
$-v$	the additive inverse of the vector v
$u+v$	the sum of the vectors u and v
λv	the vector v multiplied by the scalar λ
$\dim(V)$	the dimension of the vector space V
$\operatorname{codim}(U)$	the codimension of a subspace U of V
$U \cong V$	the vector spaces U and V are isomorphic
$U \leq V$	U is a vector subspace of V
$U < V$	U is a proper vector subspace of V
$\operatorname{Sub}(V)$	the set of vector subspaces of V
$U \cap W$	the intersection of the vector subspaces U and W
$U \oplus V$	the direct sum of the vector spaces U and V
$\bigoplus_{i\in I} V_i$	the direct sum of the vector spaces V_i $(i \in I)$
$\varphi\colon U \to V$	φ is a linear map from U to V
$\varphi\colon U \hookrightarrow V$	φ is an injective linear map from U to V
$\varphi\colon U \twoheadrightarrow V$	φ is a surjective linear map from U to V
$\varphi \oplus \psi$	the direct sum of the linear maps φ and ψ
$\bigoplus_{i\in I} \varphi_i$	the direct sum of the linear maps φ_i $(i \in I)$
$\operatorname{im}(\varphi)$	the image of the linear map φ
$\ker(\varphi)$	the kernel of the linear map φ
$\operatorname{rk}(\varphi)$	the rank of the linear map φ
V/W	the quotient vector space of V by the subspace W
$\langle v_1, \dots, v_n\rangle$	the subspace generated by the vectors $v_1, \dots, v_n$
$\det(g)$	the determinant of the linear endomorphism g
$\operatorname{tr}(g)$	the trace of the linear endomorphism g
$F(\chi)$	the Frobenius block of a (monic) polynomial $\chi \in k[X]$
$J_e(\lambda)$	the Jordan block of size $e \in \mathbb{Z}_+$ with eigenvalue $\lambda \in k$
$c_\chi(g)$	the number of Frobenius blocks $F(\chi)$ the generalized Jordan normal form of g
$V_\chi(g)$	the $F(\chi)$-isotypic vector space part in the generalized Jordan normal form of g
$n_\chi(g)$	the dimension of $V_\chi(g)$; $n_\chi(g) = \deg(\chi)c_\chi(g)$
$\mathrm{id}, \mathrm{id}_V$	the identity (on the vector space V)
$\operatorname{diag}(\lambda_1, \dots, \lambda_n)$	the diagonal matrix with entries $\lambda_1, \dots, \lambda_n$ on the diagonal
$\operatorname{End}(V), \mathbf{M}(V)$	the ring of endomorphisms of the vector space V
$\mathbf{PM}(V)$	the projective space associated to $\mathbf{M}(V)$

$\mathbf{M}_n(k)$	the endomorphism ring $\mathbf{M}(k^n)$
$\mathrm{GL}(V)$	the general linear group of V; $\mathrm{GL}(V) = \mathbf{M}(V)^\times$
$\mathrm{SL}(V)$	the special linear group of V
$\mathrm{PGL}(V)$	the projective general linear group of V
$\mathrm{PSL}(V)$	the projective special linear group of V
$\mathrm{GL}_n(k)$	the general linear group of degree n over the field k
$\mathrm{SL}_n(k)$	the special linear group of degree n over the field k
$\mathrm{PGL}_n(k)$	the projective general linear group of degree n over the field k
$\mathrm{PSL}_n(k)$	the projective special linear group of degree n over the field k

Classical groups of Lie type and spaces with form

Symbol	Explanation
$\mathbb{F}_q$	the finite field with q elements
$p = \mathrm{char}(\mathbb{F}_q)$	the characteristic of the defining field
V	the natural module
$\mathrm{GL}_n(q)$	the group $\mathrm{GL}_n(\mathbb{F}_q)$
$\mathrm{SL}_n(q)$	the group $\mathrm{SL}_n(\mathbb{F}_q)$
$\mathrm{PGL}_n(q)$	the group $\mathrm{PGL}_n(\mathbb{F}_q)$
$\mathrm{PSL}_n(q)$	the group $\mathrm{PSL}_n(\mathbb{F}_q)$
σ	the sesquilinear map $x \mapsto x^q$
$f(u,v)$	the σ-sesquilinear form f applied to the vectors u and v
$Q(v)$	the quadratic form Q applied to the vector v
$u \perp v$	the vectors u and v are perpendicular, i.e., $f(u,v) = 0$
$U \perp W$	the spaces U and W are perpendicular, i.e., $u \perp w$ for all $u \in U$, $w \in W$
$Z = U \perp W$	Z is the orthogonal direct sum of U and W, i.e., $Z = U \oplus W$ and $U \perp W$
$U^\perp$	the perpendicular space of U; $U^\perp = \{v \in V \mid u \perp v \text{ for all } u \in U\}$
$\mathrm{rad}(f)$	the radical of f; $\mathrm{rad}(f) = V^\perp$
$\mathrm{rad}(Q)$	the radical of Q; $\mathrm{rad}(Q) = \mathrm{rad}(f) \cap \{v \in V \mid Q(v) = 0\}$
$\mathrm{GI}(V,f)$	the full isometry group of (V,f)
$\mathrm{GI}(V,Q)$	the full isometry group of (V,Q)
$\mathrm{Sp}_{2m}(q)$	the symplectic group of degree $2m$ over $\mathbb{F}_q$
$\mathrm{GO}_{2m+1}(q)$	the general orthogonal group of degree $2m+1$ over $\mathbb{F}_q$
$\mathrm{GO}^\varepsilon_{2m}(q)$	the general orthogonal group of degree $2m$ of ε type over $\mathbb{F}_q$ $(\varepsilon = \pm)$
$\mathrm{SO}_{2m+1}(q)$	the special orthogonal group of degree $2m+1$ over $\mathbb{F}_q$
$\mathrm{SO}^\varepsilon_{2m}(q)$	the special orthogonal group of degree $2m$ of ε type over $\mathbb{F}_q$ $(\varepsilon = \pm)$
$\Omega_{2m+1}(q)$	the kernel of the spinor norm (or quasideterminant for $p = 2$)
$\Omega^\varepsilon_{2m}(q)$	the kernel of the spinor norm (or quasideterminant for $p = 2$; $\varepsilon = \pm$)
$\mathrm{SO}_3(\mathbb{R})$	the special orthogonal group of degree three over $\mathbb{R}$

$\mathrm{GU}_n(q)$	the general unitary group of degree n over $\mathbb{F}_{q^2}$
$\mathrm{SU}_n(q)$	the special unitary group of degree n over $\mathbb{F}_{q^2}$
U_n	the general unitary group of degree n over $\mathbb{C}$
SU_n	the special unitary group of degree n over $\mathbb{C}$
$\mathrm{PSp}_{2m}(q)$	the projective symplectic group of degree $2m$ over $\mathbb{F}_q$
$\mathrm{PGO}_{2m+1}(q)$	the projective general orthogonal group of degree $2m+1$ over $\mathbb{F}_q$
$\mathrm{PGO}^{\varepsilon}_{2m}(q)$	the projective general orthogonal group of degree $2m$ of ε type over $\mathbb{F}_q$ ($\varepsilon = \pm$)
$\mathrm{PSO}_{2m+1}(q)$	the projective special orthogonal group of degree $2m+1$ over $\mathbb{F}_q$
$\mathrm{PSO}^{\varepsilon}_{2m}(q)$	the projective special orthogonal group of degree $2m$ of ε type over $\mathbb{F}_q$ ($\varepsilon = \pm$)
$\mathrm{P}\Omega_{2m+1}(q)$	the group $\Omega_{2m+1}(q)/\{\pm \mathrm{id}\}$
$\mathrm{P}\Omega^{\varepsilon}_{2m}(q)$	the group $\Omega^{\varepsilon}_{2m}(q)/\{\pm \mathrm{id}\}$ ($\varepsilon = \pm$)
$\mathrm{PGU}_n(q)$	the projective general unitary group of degree n over $\mathbb{F}_{q^2}$
$\mathrm{PSU}_n(q)$	the projective special unitary group of degree n over $\mathbb{F}_{q^2}$

Ring and field theory

Symbol	Explanation
$0, 1$	the neutral element with respect to addition resp. multiplication
$a + b$	the sum of the ring elements a and b
ab	the product of the ring elements a and b
a^{-1}	the inverse of the unit a
$R^{\times}$	the units of the ring R
rR	the right ideal generated by $r \in R$
(r)	the principal ideal generated by $r \in R$
R/I	the quotient of the ring R by its ideal I
r^{α}	the image of the ring element r under the ring automorphism α
R_{α}	the fixed ring of the ring automorphism α on R
I_{α}	the subset of the ideal I fixed pointwise by α
$\mathrm{Aut}(R)$	the automorphism group of the ring R
S^{+k}	the k-fold sum of the subset S of a ring; $S^{+k} = \{s_1 + \cdots + s_k \mid s_1, \ldots, s_k \in S\}$
$\mathbf{C}(S), \mathbf{C}_R(S)$	the centralizer of the set S (in the ring R)
$R[X_1, \ldots, X_n]$	the polynomial ring over R with free commuting variables $X_1, \ldots, X_n$
$\deg(r)$	the degree of the polynomial $r \in R[X]$
$\Phi_n(X)$	the nth cyclotomic polynomial (of degree $\varphi(n)$)
$\overline{k}$	the algebraic closure of the field k
$k[G]$	the group algebra of G

$\mathrm{tr}_{L/K}(x)$	the trace of $x \in L$ with respect to the subfield K
$\mathrm{Gal}(L/K)$	the Galois group of the Galois extension L/K

Ultraproducts and norms

Symbol	Explanation
I	an index set
$\mathcal{U}$	a non-principal ultrafilter
$\mathcal{H}_{\mathcal{U}}$	the algebraic ultraproduct of the sequence $\mathcal{H}$ of groups along the ultrafilter $\mathcal{U}$
$\mathcal{H}_{\mathcal{U}}^{\mathrm{met}}$	the metric ultraproduct of the sequence $\mathcal{H}$ of normed groups along the ultrafilter $\mathcal{U}$
$\lim_{\mathcal{U}}$	the limit operator associated to the ultrafilter $\mathcal{U}$
ℓ, ℓ_H	a length function (on the group H)
$\ell_{\mathrm{d}}, d_{\mathrm{d}}$	the discrete length function (and the associated metric)
$\ell_{\mathrm{c}}, d_{\mathrm{c}}$	the normalized conjugacy length function (and the associated metric)
$\ell_{\mathrm{H}}, d_{\mathrm{H}}$	the normalized Hamming length function (and the associated metric)
$\ell_{\mathrm{rk}}, d_{\mathrm{rk}}$	the normalized rank length function (and the associated metric)
$\ell_{\mathrm{pr}}, d_{\mathrm{pr}}$	the normalized projective rank length function (and the associated metric)
$\ell_{\mathrm{Cay},S}, d_{\mathrm{Cay},S}$	the Cayley length function for the set S (and the associated metric)

Special symbols in Chapter 1

Symbol	Explanation
$\mathcal{H} = (H_i)_{i \in I}$	a sequence of quasisimple groups H_i $(i \in I)$
$S(H)$	the quasiscalars of the quasisimple group H (see page 20)
A_0	the subgroup of $G = \prod \mathcal{H}$ defined on page 28
A_1	the subgroup of $G = \prod \mathcal{H}$ defined on page 28
N_0	the subgroup of $G = \prod \mathcal{H}$ defined in Theorem 1.3 (see page 18)
N_1	the subgroup of $G = \prod \mathcal{H}$ defined on page 28
N_{rk}	the subgroup of $G = \prod \mathcal{H}$ defined in Theorem 1.3 (see page 18)
N_{pr}	the subgroup of $G = \prod \mathcal{H}$ defined in Theorem 1.3 (see page 18)
Φ, Ψ	maps defined in Theorem 1.3(iii) (see page 19)
$\mathcal{L}$	a poset defined on page 26
$\sim$	an equivalence relation on $\mathcal{L}$ defined on page 26
$(\mathcal{L}/\sim, \leq)$	a linear order defined on page 26

L	a subset of $\mathcal{L}/\sim$ defined on page 26
ct	the convergence type (see page 26)
α, β	maps defined in Lemma 1.22 (see page 27)
$\mathrm{GO}(U)$	the general orthogonal group on the submodule U
$\mathrm{SO}(U)$	the special orthogonal group on the submodule U

Special symbols in Chapter 2

Symbol	Explanation
$\mathcal{C}$	a class of finite groups
$\mathcal{C}^{\mathrm{P}}$	the class of finite products of $\mathcal{C}$-groups
$\mathcal{C}^{\mathrm{SP}}$	the class of subgroups of finite products of $\mathcal{C}$-groups
Fin	the class of all finite groups
Alt	the class of all finite alternating groups
$\mathbf{Ab}_d$	the class of all finite d-generated abelian groups ($d \in \mathbb{N}$)
Nil	the class of all finite nilpotent groups
Sol	the class of all finite solvable groups
PSL	the class of all finite simple groups $\mathrm{PSL}_n(q)$
$\ell_{\mathbf{F}}$	the word length function on the free group $\mathbf{F}$ (see page 38)
$B_\varrho(\mathbf{F})$	the ϱ-ball of $\mathbf{F}$ around the identity $1_{\mathbf{F}}$ in the word metric (see page 38)
$\overline{S}$	the closure of the set S in the pro-$\mathcal{C}$ topology (see page 38)
$\widehat{\mathbf{F}}$	the profinite completion of $\mathbf{F}$

Special symbols in Chapter 3

Symbol	Explanation
w	a non-trivial word
$w(g_1, \ldots, g_r)$	the word map w applied to the tuple $(g_1, \ldots, g_r)$ of group elements
tr_w	the trace of the word map w (see page 62)
O	Landau's O notation
Ω	Landau's Ω notation
$\mathrm{rad}(x)$	the radical of the number $x \in \mathbb{Z}_+$ (see page 57)
$\pi_y(x)$	the y-part of x ($x, y \in \mathbb{Z}_+$; see page 57)
L_q	the projective line over $\mathbb{F}_q$
$[a : b]$	a point on the line L_q ($(a, b) \in \mathbb{F}_q^2 \setminus \{(0, 0)\}$)
$w(G)$	the image of the word map associated to w on the group G
K	the one-relator group $\mathbf{F}_2/\langle\langle w \rangle\rangle$ (see page 67)

X	the Cayley complex of the one-relator group K (see page 67)
Γ	the Cayley graph $\mathrm{Cay}(K, \{x, y\})$
$X(\pi)$	the quotient of X by π (see page 67)
$\Gamma(\pi)$	the Cayley graph $\mathrm{Cay}(G, \{g, h\})$ (see page 67)
$X_1(\pi), X_2(\pi)$	the 1- and 2-skeleton of $X(\pi)$
$(C_\bullet, d_\bullet)$	the chain complex defined on page 68
$(C^\bullet, d^\bullet)$	the cochain complex defined on page 68
$(C_\bullet(\pi), d_\bullet(\pi))$	the chain complex defined on page 67
$(C^\bullet(\pi), d^\bullet(\pi))$	the cochain complex defined on page 67
$\mathrm{supp}(y)$	the support of $y = \sum_{v\in G} \lambda_v v \in k[G]$; $\mathrm{supp}(y) = \{v \in G \mid \lambda_v \neq 0\}$
r^*	the dual polynomial of r (see page 79)
$\mathrm{tr}_\tau^{\pm 1}$	the map $x \mapsto x \pm x^\tau$ (see page 80)
$\mathrm{tr}_\alpha^{\pm 1}$	the map $x \mapsto x \pm x^\alpha$ (see page 84)
N_τ	the map $x \mapsto xx^\tau$ (see page 80)
N_α	the map $r \mapsto rr^\alpha$ (see page 80)
$D^j i$	the jth Hasse derivative of the polynomial i; $D^j(X^n) = \binom{n}{j} X^{n-j}$
U^*	the space of σ-semilinear functionals on the vector space U
$\mathrm{GU}(\ell^2 G)$	the general unitary group of the Hilbert space $\ell^2 G$

Special symbols in Chapter 4

Symbol	Explanation
$\mathcal{H} = (H_i)_{i\in I}$	a sequence of groups H_i ($i \in I$; see page 91)
$G = \mathcal{H}_\mathcal{U}^{\mathrm{met}}$	the metric ultraproduct of the groups H_i ($i \in I$; see page 91)
$\overline{G} = G/\mathbf{Z}(G)$	the unique simple quotient of G
$\mathrm{S}_\mathcal{U}$	a metric ultraproduct of groups S_{n_i} ($i \in I$; see page 92)
$\mathrm{GL}_\mathcal{U}$	a metric ultraproduct of groups $\mathrm{GL}_{n_i}(q_i)$ ($i \in I$; see page 92)
$\mathrm{GL}_\mathcal{U}(q)$	a metric ultraproduct of groups $\mathrm{GL}_{n_i}(q)$ ($i \in I$; see page 92)
$\mathrm{Sp}_\mathcal{U}$	a metric ultraproduct of groups $\mathrm{Sp}_{2m_i}(q_i)$ ($i \in I$; see page 92)
$\mathrm{Sp}_\mathcal{U}(q)$	a metric ultraproduct of groups $\mathrm{Sp}_{2m_i}(q)$ ($i \in I$; see page 92)
$\mathrm{GO}_\mathcal{U}$	a metric ultraproduct of groups $\mathrm{GO}_{2m_i}^\pm(q_i)$ or (for q_i odd) $\mathrm{GO}_{2m_i+1}(q_i)$ ($i \in I$; see page 92)
$\mathrm{GO}_\mathcal{U}(q)$	a metric ultraproduct of groups $\mathrm{GO}_{2m_i}^\pm(q)$ or (for q odd) $\mathrm{GO}_{2m_i+1}(q)$ ($i \in I$; see page 92)
$\mathrm{GU}_\mathcal{U}$	a metric ultraproduct of groups $\mathrm{GU}_{n_i}(q_i)$ ($i \in I$; see page 92)
$\mathrm{GU}_\mathcal{U}(q)$	a metric ultraproduct of groups $\mathrm{GU}_{n_i}(q)$ ($i \in I$; see page 92)
$\mathrm{PGL}_\mathcal{U}$	a metric ultraproduct of groups $\mathrm{PGL}_{n_i}(q_i)$ ($i \in I$; see page 92)
$\mathrm{PGL}_\mathcal{U}(q)$	a metric ultraproduct of groups $\mathrm{PGL}_{n_i}(q)$ ($i \in I$; see page 92)
$\mathrm{PSp}_\mathcal{U}$	a metric ultraproduct of groups $\mathrm{PSp}_{2m_i}(q_i)$ ($i \in I$; see page 92)

$\mathrm{PSp}_{\mathcal{U}}(q)$	a metric ultraproduct of groups $\mathrm{PSp}_{2m_i}(q)$ ($i \in I$; see page 92)
$\mathrm{PGO}_{\mathcal{U}}$	a metric ultraproduct of groups $\mathrm{PGO}^{\pm}_{2m_i}(q_i)$ or (for q_i odd) $\mathrm{PGO}_{2m_i+1}(q_i)$ ($i \in I$; see page 92)
$\mathrm{PGO}_{\mathcal{U}}(q)$	a metric ultraproduct of groups $\mathrm{PGO}^{\pm}_{2m_i}(q)$ or (for q odd) $\mathrm{PGO}_{2m_i+1}(q)$ ($i \in I$; see page 92)
$\mathrm{PGU}_{\mathcal{U}}$	a metric ultraproduct of groups $\mathrm{PGU}_{n_i}(q_i)$ ($i \in I$; see page 92)
$\mathrm{PGU}_{\mathcal{U}}(q)$	a metric ultraproduct of groups $\mathrm{PGU}_{n_i}(q)$ ($i \in I$; see page 92)
$\mathbf{M}_{\mathcal{U}}$	the metric ultraproduct of the rings $\mathbf{M}_{n_i}(q_i)$ ($i \in I$; see page 92)
$\mathbf{M}_{\mathcal{U}}(q)$	the metric ultraproduct of the rings $\mathbf{M}_{n_i}(q)$ ($i \in I$; see page 92)
$\mathbf{M}_{\mathcal{U}}(k)$	the metric ultraproduct of the rings $\mathbf{M}_{n_i}(k)$ for a field k
$\mathrm{P}\mathbf{M}_{\mathcal{U}}$	the metric ultraproduct of the spaces $\mathrm{P}\mathbf{M}_{n_i}(q_i)$ ($i \in I$; see page 92)
$\mathrm{P}\mathbf{M}_{\mathcal{U}}(q)$	the metric ultraproduct of the spaces $\mathrm{P}\mathbf{M}_{n_i}(q)$ ($i \in I$; see page 92)
f_i, Q_i	the sesquilinear resp. quadratic form stabilized by H_i ($i \in I$; see page 92)
$r_k(\sigma), r_\xi(g)$	quantities defined on page 93
$r(\sigma), r(g)$	the tuple $(r_k(\sigma))_{k\in\mathbb{Z}_+}$ resp. $(r_\chi(\sigma))_{\chi\in k[X]}$
$q_k(\sigma), q_\chi(g)$	quantities defined on page 93
$q(\sigma), q(g)$	the tuple $(q_k(\sigma))_{k\in\mathbb{Z}_+}$ resp. $(q_\chi(\sigma))_{\chi \text{ primary}}$
$(\mathcal{S}, \mu)$	the Loeb space defined on page 97
$(\mathcal{V}, \dim)$	the continuous geometry defined on page 97
$\Omega_k(\sigma), V_\chi(g)$	the extension of the corresponding expressions to $\mathcal{S}$ resp. $\mathcal{V}$ (see Remark 4.7; page 98)
$\mathbf{C}_{\mathrm{conf}}(g)$	the conformal centralizer of $g \in G$; $\mathbf{C}_{\mathrm{conf}}(g) = \{h \in G \mid gh = zhg \text{ for some } z \in \mathbf{Z}(G)\}$
$\mathbf{C}^2(g), \mathbf{C}^2_{\mathrm{conf}}(g)$	the double (conformal) centralizer of g (see page 103)
$e_H(o)$	the quantity defined on page 108 for a group H and $o \in \mathbb{Z}_+$

Index

$(S, \varepsilon, \delta_\bullet)$-homomorphism, 33
$F(\chi)$-isotypic linear map, 10
$\mathcal{C}$-approximable abstract group, 33
$\mathcal{C}$-approximable topological group, 34
$\mathcal{C}$-group, 33
$\mathcal{C}$-separable normal subgroup, 38
ϱ-ball in a normed group, 38
k-fold product of a subset, 7
k-isotypic permutation, 8
q-Frobenius map, 10

algebraic closure of a field, 13
algebraic ultraproduct, 13
almost simple group, 44
alternating bilinear form, 10
alternating group, 8
automorphism group, 7

Cayley length function, 15
center, 8
centralizer, 8, 13
characteristic of a field, 10
codimension, 9
commutator, 8
commutator subgroup, 8
compactification of a group, 50
conjugacy class, 7
conjugate, 7
conjugate-symmetric sesquilinear form, 10
continuous geometry, 97
convergence type, 26
cycle type of a permutation, 8
cyclic group, 7
cyclotomic polynomial, 13

degree of a polynomial, 13
derived series, 8
determinant, 9
dimension of a vector space, 8
direct product of groups, 7
direct sum, 9
discrete length function, 15
double centralizer, 103
double conformal centralizer, 103

exponent of a group, 7
extensions of groups, 8

field trace, 13
finite field, 10
fixed ideal, 13
fixed point set of a permutation, 8
fixed ring, 13
form of minus type, 11
form of plus type, 11
free group, 7
Frobenius block, 9
full isometry group, 11

Galois extension, 13
Galois group, 13
general linear group, 10
general orthogonal group, 11
general unitary group, 11
generalized Jordan normal form, 9
generated normal subgroup, 7
generated subgroup, 7
generated subspace, 9
generator, 8
group action, 8
group algebra, 13
group homomorphism, 7

group presentation, 8

identity, 10
identity permutation, 8
image of a group homomorphism, 7
image of a linear map, 9
injective homomorphism, 7
intersection of subgroups, 7
intersection of subspaces, 9
invariant length function, 14
inverse element, 7
inverse vector, 8
isomorphic groups, 7
isomorphic vector spaces, 9
isotropic vector, 10
isotypic linear map, 10
isotypic permutation, 8

Jordan block, 9

kernel of a group homomorphism, 7
kernel of a linear map, 9

length function, 14
linear map, 9
linearly equivalent forms, 11
Lipschitz continuous, 16
Lipschitz equivalent, 16
Loeb space, 97
lower central series, 8

matrix ring, 9
metric ultraproduct, 14
monic polynomial, 9
multiplicative inverse, 13

natural homomorphism, 7
natural module, 12
neutral element, 7
non-degenerate quadratic form, 11
non-singular quadratic form, 11
non-singular sesquilinear form, 11
non-singular subspace, 11
non-trivial word, 7
norm, 14
normal subgroup, 7
normal subset, 7
normalized conjugacy length function, 15
normalized Hamming length function, 15
normalized projective length function, 15
normalized rank length function, 15
normed group, 14

orbit, 8
order of a group element, 7
orthogonal direct sum, 10
orthogonally indecomposable, 77

perpendicular space, 11
perpendicular subspaces, 10
perpendicular vectors, 10
polynomial ring, 9, 13
power word, 55
primary ideal, 9
primary polynomial, 9
primary rational canonical form, 9
principal ideal, 13
pro-$\mathcal{C}$ topology, 37
product of group elements, 7
product of ring elements, 13
projective general linear group, 10
projective general orthogonal group, 12
projective general unitary group, 12
projective special linear group, 10
projective special orthogonal group, 12
projective special unitary group, 12
projective symplectic group, 12
pseudofinite group, 50

quadratic form, 10
quotient group, 7
quotient vector space, 9

radical of a form, 11
rank of a free group, 7
rank of a linear map, 9

relation, 8
right ideal, 13
ring automorphism, 8

scalar multiple, 8
Schur covering group, 8
self-dual polynomial, 79
semilinear functional, 78
semisimple element, 99
special linear group, 10
special orthogonal group, 12
special unitary group, 12
stabilizer, 8
subgroup, 7
subspace, 8
sum of ring elements, 13
sum of vectors, 8
support of a permutation, 8
surjective homomorphism, 7
symmetric bilinear form, 10
symmetric group, 8

totally isotropic subspace, 11
trace, 9
trivial group, 7
Turing-approximable group, 50

ultralimit, 14
units of an algebra, 10

Witt subspace, 11
word, 7
word map, 53
wreath product, 8

zero vector, 8

Bibliography

[1] Goulnara Arzhantseva and Liviu Păunescu. Linear sofic groups and algebras. *Transactions of the American Mathematical Society*, 369(4):2285–2310, 2017.

[2] Nir Avni, Tsachik Gelander, Martin Kassabov, and Aner Shalev. Word values in p-adic and adelic groups. *Bulletin of the London Mathematical Society*, 45(6):1323–1330, 2013.

[3] Tatiana Bandman, Shelly Garion, and Fritz Grunewald. On the surjectivity of Engel words on $\mathrm{PSL}(2,q)$. *Groups, Geometry, and Dynamics*, 6(3):409–439, 2012.

[4] Armand Borel. On free subgroups of semisimple groups. *L'Enseignement Mathématique*, 29(1–2):151–164, 1983.

[5] Sarvadaman Chowla. On the least prime in an arithmetic progression. *Journal of the Indian Mathematical Society*, 1(2):1–3, 1934.

[6] Yves Cornulier. A sofic group away from amenable groups. *Mathematische Annalen*, 350(2):269–275, 2011.

[7] Marcus De Chiffre, Lev Glebsky, Alexander Lubotzky, and Andreas Thom. Stability, cohomology vanishing, and non-approximable groups. *arXiv:1711.10238*, 2017.

[8] Pierre Deligne. Extensions centrales non résiduellement finies de groupes arithmétiques. *Comptes rendus de l'Académie des Sciences*, 287(4):203–208, 1978.

[9] Warren Dicks and Peter Linnell. L^2-Betti numbers of one-relator groups. *Mathematische Annalen*, 337(4):855–874, 2007.

[10] Michal Doucha. Metric topological groups: Their metric approximation and metric ultraproducts. *arXiv:1601.07449*, 2016.

[11] Randall Dougherty and Jan Mycielski. Representations of infinite permutations by words (II). *Proceedings of the American Mathematical Society*, 127(8):2233–2243, 1999.

[12] Philip Dowerk and Andreas Thom. Bounded normal generation and invariant automatic continuity. *arXiv:1506.08549*, 2015.

[13] Gábor Elek and Endre Szabó. Sofic groups and direct finiteness. *Journal of Algebra*, 280(2):426–434, 2004.

[14] Gábor Elek and Endre Szabó. Hyperlinearity, essentially free actions and L^2-invariants. The sofic property. *Mathematische Annalen*, 332(2):421–441, 2005.

[15] Gábor Elek and Endre Szabó. On sofic groups. *Journal of Group Theory*, 9(2):161–171, 2006.

[16] Abdelrhman Elkasapy and Andreas Thom. About Gotô's method showing surjectivity of word maps. *Indiana University Mathematics Journal*, 63(5):1553–1565, 2014.

[17] Paul Ellis, Sherwood Hachtman, Scott Schneider, and Simon Thomas. Ultraproducts of finite alternating groups (combinatorial and descriptive set theory). 2008.

[18] Jan-Hendrik Evertse. Linear forms in logarithms. `http://www.math.leidenuniv.nl/~evertse/dio2011-linforms.pdf`, April 2019.

[19] Ralph Fox. Free differential calculus. I: Derivation in the free group ring. *Annals of Mathematics*, 57(3):547–560, 1953.

[20] Robert Fricke and Felix Klein. *Vorlesungen über die Theorie der automorphen Funktionen. Band 1: Die gruppentheoretischen Grundlagen. Band II: Die funktionentheoretischen Ausführungen und die Anwendungen*, volume 3 of *Bibliotheca Mathematica Teubneriana*. B. G. Teubner Verlag, 1965.

[21] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer Science & Business Media, 1991.

[22] Shelly Garion and Aner Shalev. Commutator maps, measure preservation, and T-systems. *Transactions of the American Mathematical Society*, 361(9):4631–4651, 2009.

[23] Tsachik Gelander. Limits of finite homogeneous metric spaces. *arXiv:1205.6553*, 2012.

[24] Lev Glebsky. Approximations of groups, characterizations of sofic groups, and equations over groups. *Journal of Algebra*, 477:147–162, 2017.

[25] Samuel Gonshaw, Martin Liebeck, and Eamon O'Brien. Unipotent class representatives for finite classical groups. *Journal of Group Theory*, 20(3):505–525, 2017.

[26] Nikolai Gordeev, Boris Kunyavskiĭ, and Eugene Plotkin. Word maps and word maps with constants of simple algebraic groups. *Doklady Mathematics*, 94(3):632–634, 2016.

[27] Nikolai Gordeev, Boris Kunyavskiĭ, and Eugene Plotkin. Word maps on perfect algebraic groups. *International Journal of Algebra and Computation*, 28(8):1487–1515, 2018.

[28] Edouard Goursat. Sur les substitutions orthogonales et les divisions régulières de l'espace. In *Annales scientifiques de l'École Normale Supérieure*, volume 6, pages 9–102. Elsevier, 1889.

[29] Mikhael Gromov. Endomorphisms of symbolic algebraic varieties. *Journal of the European Mathematical Society*, 1(2):109–197, 1999.

[30] Funda Gul and Armin Weiß. On the dimension of matrix embeddings of torsion-free nilpotent groups. *Journal of Algebra*, 477:516–539, 2017.

[31] Robert Guralnick, Martin Liebeck, Eamon O'Brien, Aner Shalev, and Pham Huu Tiep. Surjective word maps and Burnside's $p^a q^b$ theorem. *Inventiones mathematicae*, 213(2):589–695, 2018.

[32] Robert Guralnick and Pham Huu Tiep. Effective results on the Waring problem for finite simple groups. *American Journal of Mathematics*, 137(5):1401–1430, 2015.

[33] Wim Hesselink. Nilpotency in classical groups over a field of characteristic 2. *Mathematische Zeitschrift*, 166(2):165–181, 1979.

[34] Derek Holt and Sarah Rees. Some closure results for $\mathcal{C}$-approximable groups. *Pacific Journal of Mathematics*, 287(2):393–409, 2017.

[35] James Howie. The p-adic topology on a free group: A counterexample. *Mathematische Zeitschrift*, 187(1):25–27, 1984.

[36] Chun Yin Hui, Michael Larsen, and Aner Shalev. The Waring problem for Lie groups and Chevalley groups. *Israel Journal of Mathematics*, 210(1):81–100, 2015.

[37] Bertram Huppert. *Endliche Gruppen I*, volume 134. Springer Science & Business Media, 2013.

[38] A. Ivanov. Sofic metric groups and continuous logic. *arXiv:1604.08446*, 2016.

[39] Sebastian Jambor, Martin Liebeck, and Eamon O'Brien. Some word maps that are non-surjective on infinitely many finite simple groups. *Bulletin of the London Mathematical Society*, 45(5):907–910, 2013.

[40] Aditi Kar and Nikolay Nikolov. A non-LEA sofic group. *Proceedings – Mathematical Sciences*, 127(2):289–293, 2017.

[41] Anton Klyachko and Andreas Thom. New topological methods to solve equations over groups. *Algebraic & Geometric Topology*, 17(1):331–353, 2017.

[42] Michael Larsen. How often is a permutation an nth power? *math/9712223*, 1997.

[43] Michael Larsen. Word maps have large image. *Israel Journal of Mathematics*, 139(1):149–156, 2004.

[44] Michael Larsen and Aner Shalev. Word maps and Waring type problems. *Journal of the American Mathematical Society*, 22(2):437–466, 2009.

[45] Michael Larsen and Aner Shalev. On the distribution of values of certain word maps. *Transactions of the American Mathematical Society*, 368(3):1647–1661, 2016.

[46] Michael Larsen and Aner Shalev. Words, Hausdorff dimension and randomly free groups. *Mathematische Annalen*, 371(3–4):1409–1427, 2018.

[47] Michael Larsen, Aner Shalev, and Pham Huu Tiep. Waring problem for finite quasisimple groups. *International Mathematics Research Notices*, 2013(10):2323–2348, 2012.

[48] Martin Liebeck. Width questions for finite simple groups. *Groups St. Andrews 2013*, 422:51, 2015.

[49] Martin Liebeck and Aner Shalev. Diameters of finite simple groups: Sharp bounds and applications. *Annals of mathematics*, pages 383–406, 2001.

[50] Yuri Vladimirovich Linnik. On the least prime in an arithmetic progression. I. The basic theorem. *Recueil Mathématique. Nouvelle Série*, 15(2):139–178, 1944.

[51] Alexander Lubotzky. Images of word maps in finite simple groups. *Glasgow Mathematical Journal*, 56(2):465–469, 2014.

[52] Wolfgang Lück. *L^2-invariants: Theory and applications to geometry and K-theory*, volume 44. Springer Science & Business Media, 2013.

[53] Dermot McLain. A characteristically simple group. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 50, pages 641–642. Cambridge University Press, 1954.

[54] Werner Nickel. Matrix representations for torsion-free nilpotent groups by Deep Thought. *Journal of Algebra*, 300(1):376–383, 2006.

[55] Nikolay Nikolov, Jakob Schneider, and Andreas Thom. Some remarks on finitarily approximable groups. *Journal de l'École polytechnique – Mathématiques*, 5, 2017.

[56] Martin Nitsche and Andreas Thom. Universal solvability of group equations. *arXiv:1811.07737*, 2018.

[57] Abderezak Ould Houcine and Françoise Point. Alternatives for pseudofinite groups. *Journal of Group Theory*, 16(4):461–495, 2013.

[58] Vladimir Pestov. Hyperlinear and sofic groups: A brief guide. *Bulletin of Symbolic Logic*, 14(4):449–480, 2008.

[59] Anand Pillay. Remarks on compactifications of pseudofinite groups. *Fundamenta Mathematicae*, 236(2):193–200, 2017.

[60] Derek Robinson. *Finiteness conditions and generalized soluble groups. Part 1+2.* Springer Science & Business Media, 1972.

[61] Jakob Schneider and Andreas Thom. A note on the normal subgroup lattice of ultraproducts of finite quasisimple groups. *arXiv:1709.06286*, 2017.

[62] Jakob Schneider and Andreas Thom. Word images in symmetric and unitary groups are dense. *arXiv:1802.09289*, 2018.

[63] Dan Segal. Closed subgroups of profinite groups. *Proceedings of the London Mathematical Society*, 81(1):29–54, 2000.

[64] Dan Segal. *Polycyclic groups.* Number 82. Cambridge University Press, 2005.

[65] Dan Segal. *Words: Notes on verbal width in groups*, volume 361. Cambridge University Press, 2009.

[66] Saharon Shelah. On the cardinality of ultraproduct of finite sets. *The Journal of Symbolic Logic*, 35(1):83–84, 1970.

[67] Abel Stolz and Andreas Thom. On the lattice of normal subgroups in ultraproducts of compact simple groups. *Proceedings of the London Mathematical Society*, 108(1):73–102, 2014.

[68] Ravindranathan Thangadurai and Akshaa Vatwani. The least prime congruent to one modulo n. *American Mathematical Monthly*, 118(8):737–742, 2011.

[69] Andreas Thom. Sofic groups and Diophantine approximation. *Communications on Pure and Applied Mathematics*, 61(8):1155–1171, 2008.

[70] Andreas Thom. Examples of hyperlinear groups without factorization property. *Groups, Geometry, and Dynamics*, 4(1):195–208, 2010.

[71] Andreas Thom. About the metric approximation of Higman's group. *Journal of Group Theory*, 15(2):301–310, 2012.

[72] Andreas Thom. Convergent sequences in discrete groups. *Canadian Mathematical Bulletin*, 56(2):424–433, 2013.

[73] Andreas Thom and John Wilson. Metric ultraproducts of finite simple groups. *Comptes Rendus Mathématique*, 352(6):463–466, 2014.

[74] Alan Turing. Finite approximations to Lie groups. *Annals of Mathematics*, pages 105–111, 1938.

[75] G. E. Wall. On the conjugacy classes in the unitary, symplectic and orthogonal groups. *Journal of the Australian Mathematical Society*, 3(1):1–62, 1963.

[76] Benjamin Weiss. Sofic groups and dynamical systems. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 350–359, 2000.

[77] John Wilson. *Profinite groups*, volume 19. Clarendon Press, 1998.

[78] Robert Wilson. *The finite simple groups*, volume 251. Springer Science & Business Media, 2009.

[79] Boris Zilber. Perfect infinities and finite approximation. *Infinity and Truth. IMS Lecture Notes Series*, 25, 2014.