ISRAEL JOURNAL OF MATHEMATICS **221** (2017), 85–115
DOI: 10.1007/s11856-017-1563-2

DESCRIBING FINITE GROUPS BY SHORT FIRST-ORDER SENTENCES

BY

André Nies

Department of Computer Science, Private Bag 92019,
The University of Auckland, Auckland, New Zealand
e-mail: andre@cs.auckland.ac.nz

AND

Katrin Tent

Mathematisches Institut, Einsteinstrasse 62,
Universität Münster, 48149 Münster, Germany
e-mail: tent@wwu.de

ABSTRACT

We say that a class of finite structures for a finite first-order signature is r-compressible for an unbounded function $r\colon \mathbb{N} \to \mathbb{N}^+$ if each structure G in the class has a first-order description of size at most $O(r(|G|))$. We show that the class of finite simple groups is log-compressible, and the class of all finite groups is $\log^3$-compressible. As a corollary we obtain that the class of all finite transitive permutation groups is $\log^3$-compressible. The results rely on the classification of finite simple groups, the bi-interpretability of the twisted Ree groups with finite difference fields, the existence of profinite presentations with few relators for finite groups, and group cohomology. We also indicate why the results are close to optimal.

1. Introduction

Let L be first-order logic in a signature consisting of finitely many relation symbols, function symbols, and constants. We say that a sentence φ in L *describes* G if G is the unique model of φ up to isomorphism. We study the

Received September 1, 2015 and in revised form April 29, 2016

compressibility of finite L-structures G up to isomorphism via such descriptions. Our main results are about compressibility of finite groups.

Note that every finite L-structure G can be described by some sentence φ: for each element of G we introduce an existentially quantified variable; we say that these are all the elements of G, and that they satisfy the atomic formulas valid for the corresponding elements of G. However, this sentence is at least as long as the size of the domain of G. We may think of a description of G which is much shorter than $|G|$ as a *compression* of G up to isomorphism.

Unless stated otherwise, descriptions of structures will be in first-order logic. For an infinite class of L-structures, we are interested in giving descriptions that are asymptotically short relative to the size of the described structure. This is embodied in the following definition. Usually the function r grows slowly.

Definition 1.1: Let $r\colon \mathbb{N} \to \mathbb{N}^+$ be an unbounded function. We say that an infinite class $\mathcal{C}$ of finite L-structures is r**-compressible** if for each structure G in $\mathcal{C}$, there is a sentence φ in L such that $|\varphi| = O(r(|G|))$ and φ describes G.

Sometimes we also want to give a short description of a structure in $\mathcal{C}$, together with a tuple of elements. We say that the class $\mathcal{C}$ is *strongly r-compressible* if for each structure G in $\mathcal{C}$, each k and each $\overline{g} \in G^k$, there is a formula $\varphi(y_1, \dots, y_k)$ in L such that $|\varphi| = O(r(|G|))$ and φ describes $(G, \overline{g})$ (where the O constant can depend on k).

In this paper, for notational convenience we will use the definition

$$\log m = \min\{r\colon\ 2^r \geq m\}.$$

The following is our first main result.

Theorem 1.2: *The class of finite simple groups is* log*-compressible.*

Finite groups can be described up to isomorphism via presentations. There is a large amount of literature on finding very short presentations for "most" finite groups G; see e.g. [1, 7, 4]. Using composition series, these presentations can be converted into first-order descriptions of G that are at most $O(\log^2 |G|)$ longer, as we will see in Proposition 5.5.

The small Ree groups ${}^2G_2(q)$ are finite simple groups that arise as subgroups of the automorphism group $G_2(q)$ of the octonion algebra over the q-element field $\mathbb{F}_q$, where q has the form 3^{2k+1} [12, Section 4.5]. They form a notorious case where short presentations are not known to exist. Nonetheless, we are able to find short first-order descriptions by using the bi-interpretability with the difference field $(\mathbb{F}_q, \sigma)$, where σ is the 3^{k+1}-th power of the Frobenius automorphism. This was proved by Ryten [11, Prop. 5.4.6(iii)]. It then suffices to give a short description of the difference field, which is not hard to obtain.

Let $\log^k$ denote the function $g(n) = (\log(n))^k$. Our second main result is the following:

Theorem 1.3: *The class of finite groups is strongly* $\log^3$*-compressible.*

We describe a general finite group G by choosing a composition series $1 = G_0 \lhd G_1 \lhd \cdots \lhd G_r = G$, where $r \leq \log|G|$. We use Theorem 1.2 to describe the factors $H_i = G_{i+1}/G_i$ of the series, which are simple by definition. We then use the method of straight line programs due to [2], and some group extension theory, to obtain short formulas describing G_{i+1} for each $i < r$ as an extension of G_i by H_i.

The proof of Theorem 1.3 assuming Theorem 1.2 is analogous to the proof of similar results for presentations, such as Babai et al. [1, Section 8] and Mann [10, Thm. 2]. However, in our case the deduction is different because we have to describe the extension of G_i by H_i in first-order logic rather than presentations. This is where we use the existence of profinite presentations with few relators for finite groups, and group cohomology (Section 6).

Recall that a permutation group is a group G together with an action of G on a set X given by a homomorphism of G into the symmetric group of X. If the action is transitive, then it is equivalent to the action of G on $H\backslash G$ by right translation, where H is the stabilizer in G of a point $x \in X$. Thus, describing the action of G on X amounts to describing G together with a distinguished subgroup H of G. Since our methods yield short descriptions of this kind, we obtain:

Corollary 1.4: *The class of finite transitive permutation groups is* $\log^3$*-compressible.*

By counting the number of non-isomorphic groups of a certain size, in Remark 7.2 we will also provide lower bounds on the length of a description, which show the near-optimality of the two main results. In particular, from the point of view of the length of first-order descriptions, simple groups are indeed simpler than general finite groups. The lower bounds apply to descriptions in any formal language, such as second-order logic. Thus, for describing finite groups, first-order logic is already optimal.

As usual in the theory of Kolmogorov complexity, we gauge how short a description of an object is by comparing it to the size of the object itself, considered as its own trivial description. Given a group of size m, each element can be encoded by $\log m$ bits. The size of the table for the operation $(a,b) \mapsto ab^{-1}$ is therefore $m^2 \log m$. This table can be seen as a trivial description. Since our bounds on the lengths of short descriptions are powers of $\log m$, up to a linear constant it does not matter whether we take m or $m^2 \log m$ as the length of the trivial description.

A Σ_r*-sentence* of L is a sentence that is in prenex normal form, starts with an existential quantifier, and has $r-1$ quantifier alternations. We say that $\mathcal{C}$ is g*-compressible using* Σ_r*-sentences* if φ in Definition 1.1 can be chosen in Σ_r form. We will provide variants of the results above where the sentences are Σ_r for a certain r. The describing sentences will be of length $O(\log^4 |G|)$.

Usually we view a formula φ of L as a string over the infinite alphabet consisting of: a finite list of logical symbols, an infinite list of variables, and the finitely many symbols of L. Sometimes we want the alphabet to be finite, which we can achieve by indexing the variables with numbers written in decimal (such as x_{901}). This increases the length of a formula by a logarithmic factor (assuming that φ always introduces new variables with the least index that is available, so that x_i occurs in φ only when $i < |\varphi|$). We then encode the resulting string by a binary string, which we call the *binary code* for φ. Its length is called the *binary length* of φ, which is $O(|\varphi| \log |\varphi|)$.

Our results are particular to the case of groups. For instance, in the case of all undirected graphs, not much compression is possible using any formal language: the length of the "brute force" descriptions given above, involving the open diagram, is close to optimal. To see this, note that there are $2^{\binom{n}{2}}$ undirected graphs on n vertices. The isomorphism class of each such graph has at most $n!$ elements. Hence the number of non-isomorphic undirected graphs with n vertices is at least $2^{\binom{n}{2}}/n! = \frac{1}{n}\Pi_{i=1}^{n-1} 2^i/i$, which for large n exceeds $\frac{1}{n}2^{n^2/6}$. For each k there are at most 2^k sentences φ with a binary code of length less than k. So for each large enough n there is an undirected graph G with n vertices such that $n^2 - 6\log n = O(|\varphi| \log |\varphi|)$ for any description φ

of G. (See [5, Cor. 2.12] for a recent proof that the lower bound $2^{\binom{n}{2}}/n!$ is asymptotically equal to the number of nonisomorphic graphs on n vertices.)

2. Short first-order formulas related to generation

This section provides short formulas related to generation in monoids and groups. They will be used later on to obtain descriptions of finite groups. Some of the results are joint work with Yuki Maehara, a former project student of Nies.

Firstly, we consider exponentiation in monoids.

LEMMA 2.1: *For each positive integer n, there is an existential formula $\theta_n(g,x)$ in the first-order language of monoids $L(e,\circ)$, of length $O(\log n)$, such that for each monoid M, $M \models \theta_n(g,x)$ if and only if $x^n = g$.*

Proof. We use a standard method from the theory of algorithms known as exponentiation via repeated squaring. Let $k = \log n$. Let $\alpha_1 \cdots \alpha_k$ be the binary expansion of n. Let $\theta_n(g,x)$ be the formula

$$(*) \qquad \exists y_1 \cdots \exists y_k [y_1 = x \ \wedge \ y_k = g \ \wedge \bigwedge_{1 \le i < k} y_{i+1} = y_i \circ y_i \circ x^{\alpha_{i+1}}],$$

where x^{α_i} is x if $\alpha_i = 1$, and x^{α_i} is e if $\alpha_i = 0$. Clearly θ_n has length $O(\log\ n)$. One verifies by induction on k that the formulas are correct. ∎

We give a sample application of Lemma 2.1 which will also be useful below. By the remark after Proposition 4.3 below, the upper bound on the length of the descriptions is close to optimal.

PROPOSITION 2.2: *The class of cyclic groups G of prime power order is* log-*compressible via sentences in the language of monoids that are in Σ_3 form.*

Proof. Suppose that $n = |G| = p^k$ where p is prime. A group H is isomorphic to G if and only if there is an element h such that $h^{p^k} = 1$, $h^{p^{k-1}} \neq 1$, and h generates H. By Lemma 2.1, the first two conditions can be expressed by formulas of length $O(\log n)$ with h as a free variable, the first existential, the second universal. For the third condition, we need a modification of Lemma 2.1, namely a formula χ_n such that for each monoid M, $M \models \chi_n(g,x)$ if and only if $x^r = g$ for some r such that $0 \le r < 2^k$, where $k = \log n$ (recall that by our definition of log, the number 2^k is the least power of two which is not less than n). We define $\chi_n(g,x)$ by

$$\exists y_0 \cdots \exists y_k [y_0 = 1 \ \wedge \ y_k = g \ \wedge \bigwedge_{0 \le i < k} (y_{i+1} = y_i \circ y_i \circ x \vee y_{i+1} = y_i \circ y_i)].$$

It is now clear that the condition that h generates the group can be expressed by a formula of length $O(\log n)$ which is in Π_2 form. ∎

For elements $x_1, \ldots, x_n$ in a group G we let $\langle x_1, \ldots, x_n \rangle$ denote the subgroup of G generated by these elements. The pigeon hole principle easily implies the following:

Lemma 2.3: *Given a generating set S of a finite group G, every element of G can be written as a product of elements of S of length at most $|G|$.*

We define a crucial collection of formulas $\alpha_k(g; x_1, \ldots, x_k)$ in the first-order language of monoids so that $\alpha_k(g; h_1, \ldots, h_k)$ expresses that g is in $\langle h_1, \ldots, h_k \rangle$. These formulas depend only on k and the size of the group G.

Lemma 2.4: *For each pair of positive integers k, v, there exists a first-order formula $\alpha_k(g; x_1, \ldots, x_k)$ in the language of monoids of length $O(k + \log v)$ such that for each group G of size at most v, $G \models \alpha_k(g; x_1, \ldots, x_k)$ if and only if $g \in \langle x_1, \ldots, x_k \rangle$.*

Proof. We use a technique that originated in computational complexity to show that the set of true quantified boolean formulas is PSPACE-complete. For $i \in \mathbb{N}$ we inductively define formulas $\delta_i(g; x_1, \ldots, x_k)$. Let

$$\delta_0(g; x_1, \ldots, x_k) \equiv \bigvee_{1 \le j \le k} [g = x_j \ \vee \ g = 1].$$

For $i > 0$ let

$$\delta_i(g; x_1, \ldots, x_k) \equiv \exists u_i \exists v_i [g = u_i v_i \ \wedge \\ \forall w_i [(w_i = u_i \vee w_i = v_i) \rightarrow \delta_{i-1}(w_i; x_1, \ldots, x_k)]].$$

Note that δ_i has length $O(k+i)$, and $G \models \delta_i(g; x_1, \ldots, x_k)$ if and only if g can be written as a product, of length at most 2^i, of x_r's.

Now let $\alpha_k(g; x_1, \ldots, x_k) \equiv \delta_p(g; x_1, \ldots, x_k)$ where $p = \log v$. Then $2^p \ge v$ by our definition of log, so α_k is a formula as required by Lemma 2.3. ∎

Remark 2.5: We note that we can optimize the formulas in Lemmas 2.1 and 2.4 so that the length bounds apply to the binary length. For instance, in Lemma 2.4 we can "reuse" the quantified variables u, v, w at each level i, so that α_k becomes a formula over an alphabet of size $k + O(1)$.

3. Straight line programs and generation

In this section we recall the Reachability Lemma from Babai and Szemerédi [2, Theorem 3.1], and the notion of a pre-processing set introduced in Babai et al. [1, Lemma 8.2] following their proof sketch. Let G be a finite group, $S \subseteq G$ and $g \in G$. A *straight line program (SLP) $\mathcal{L}$ over S* is a sequence of group elements such that each element of $\mathcal{L}$ is either in S, an inverse of an earlier element or a product of two earlier elements. We say that an SLP $\mathcal{L}$ *computes g from S* if $\mathcal{L}$ is an SLP over S containing g.

The *reduced length* of $\mathcal{L}$ is the number of elements in $\mathcal{L}$ outside S. For a set $A \subseteq G$ we say that a straight line program $\mathcal{L}$ over S *computes* A if every element of A occurs in $\mathcal{L}$. Let $\operatorname{cost}(A \mid S)$ be the shortest reduced length of a straight line program computing A from S.

For a subset S of a finite group G, Babai and Szemerédi [2] construct a set of generators A for $\langle S\rangle$ with $|A| \le \log|\langle S\rangle|$ such that every element of $\langle S\rangle = \langle A\rangle$ has length at most $2\log|G|$ as a word over A (cf. Lemma 2.3). Such *pre-processing sets* will reduce the length of the formulas in Section 7. We include the construction for convenience, and in order to adjust it for future reference to an increasing sequence of subsets of G.

Lemma 3.1 ([2, 1]): *Let G be a finite group. Suppose $T_1 \subset \cdots \subset T_k \subseteq G$ is an ascending sequence of subsets and $G_i = \langle T_i\rangle, i = 1, \ldots, k$.*

There is an ascending sequence of pre-processing sets A_i for G_i, $i = 1, \ldots, k$, with $|A_i| \le \log|G_i|, \langle A_i\rangle = \langle T_i\rangle$, $\operatorname{cost}(A_i \mid T_i) < (\log|G_i|)^2$, and $\operatorname{cost}(g \mid A_i) < 2\log|G_i|$ for every $g \in G_i$.

Proof. We first consider the case $k = 1$, i.e., a single set $S = T_1 \subseteq G$. Let s be minimal with $2^s \ge |\langle S\rangle|$ so that $s = \log|\langle S\rangle|$ according to our definition of log. For $i \le s$, we inductively define an increasing sequence of subsets $K(i) \subseteq \langle S\rangle$ of size 2^i, elements $z_i \in K(i)$ and an increasing sequence of SLPs $\mathcal{L}_i$ computing $z_1, \ldots, z_i$ from S. The set $\{z_1, \ldots, z_s\}$ will serve as our pre-processing set for $\langle S\rangle$.

To begin with, let $K(0) = \{1\}, z_0 = 1, \mathcal{L}_0 = \emptyset$. Suppose $K(i)$ and $\mathcal{L}_i$ have been defined with the required properties. If $K(i)^{-1}K(i) \ne \langle S\rangle$, there are

$v \in K(i)^{-1}K(i)$ and $x \in S$ such that $z_{i+1} := vx \notin K(i)^{-1}K(i)$. Let $K(i+1)$ be the set of products $\Pi_{l \le i+1} z_l^{\alpha_l}$ where $\alpha_l \in \{0,1\}$. By the choice of z_{i+1} we have $|K(i+1)| = 2|K(i)| = 2^{i+1}$.

One can write $z_{i+1} = v_0^{-1}v_1x$ with $v_0, v_1 \in K(i)$, and $v_0 \neq v_1$, so that we can also assume that not both v_0 and v_1 have length i. Since $\mathcal{L}_i$ computes $z_1, \ldots, z_i$, an SLP computes v_0 and v_1 which extends $\mathcal{L}_i$ by at most $2i-3$ elements (corresponding to the initial segments of the v_r of length > 1). To obtain $\mathcal{L}_{i+1}$, we append v_0^{-1}, $v_0^{-1}v_1$ and finally $z_{i+1} = v_0^{-1}v_1x$ to $\mathcal{L}_i$. In total we have appended at most $2i$ elements to $\mathcal{L}_i$. Clearly this process ends after s steps, when $K(s)^{-1}K(s) = \langle S \rangle$. Then $A = \{z_1, \ldots, z_s\}$ is a generating set for $\langle S \rangle$, and $\mathcal{L}_s$ is an SLP computing A from S of reduced length at most $2\sum_{i=1}^{s}(i-1) \le s^2$. Since $K(s)^{-1}K(s) = \langle S \rangle$, we see that any $g \in \langle S \rangle$ can be computed from $A = \{z_1, \ldots, z_s\}$ by an SLP of reduced length at most $2s - 1 < 2\log|G|$. Thus, A is the required pre-processing set.

Now suppose we have $S = T_1 \subset T_2$ and $A = A_1$ as above computed by $\mathcal{L}_{s_1} = \mathcal{L}_s$ from T_1 so that $K(s_1)^{-1}K(s_1) = \langle T_1 \rangle$. We continue the construction using elements $x \in T_2$ for $s_2 - s_1$ steps extending A_1 to a set A_2 and $\mathcal{L}_{s_1} = \mathcal{L}_s$ to an SLP $\mathcal{L}_{s_2}$. Inductively we find the required $A_i, i \le k$. ∎

Corollary 3.2: *Any finite group G has a generating set A of size at most $\log|G|$ such that any element of G has length at most $2|A|$ over A.*

We call a generating set with the latter property *swift*. Swift generating sets will be used below to give short descriptions for finite groups.

For reference we also note that the first part of the proof of Lemma 3.1 shows the following:

Corollary 3.3 (Reachability Lemma [2]): *Let $r = \log|G|$. For each set $S \subseteq G$ and any $g \in \langle S \rangle$, there is a straight line program $\mathcal{L}$ of reduced length at most $(r+1)^2$ that computes g from S.*

Proof. We build the sequences $z_1, \ldots, z_i$, $K(0), \ldots, K(i)$ and $\mathcal{L}_0, \ldots, \mathcal{L}_i$ as in the proof of Lemma 3.1 until $g \in K(s)^{-1}K(s)$. This yields an SLP computing g from S of reduced length at most $(r+1)^2$. ∎

We say that an L-formula $\varphi(\overline{x})$ in variables $\overline{x}$, possibly with parameters in a group G, *defines* an ordered tuple A (of the same length as $\overline{x}$) in G if A is the unique tuple in G such that $\varphi(A)$ holds in G (here the elements of A are substituted for the variables of φ). We can define swift generating sets for a normal series of a group G by a formula in $O(\log^2 |G|)$.

Lemma 3.4: *Let G be a finite group with a normal series*

$$1 \lhd G_0 \lhd G_1 \lhd \cdots \lhd G_r = G$$

and an ascending sequence of generating sets

$$T_0 \subset T_1 \subset \cdots \subset T_r = T$$

with $\langle T_i\rangle = G_i, |T_i| \leq \log |G_i|, 0 \leq i \leq r$.

There is a formula ψ using parameters from the set T with $|\psi| = O(\log^2 |G|)$ defining a sequence $A_1 \subseteq \cdots \subseteq A_r = A \subseteq G$ of pre-processing sets for G_i over $T_i, i \leq r$.

Proof. Note first that $r \leq \log |G|$. Let $A_1 \subseteq \cdots \subseteq A_r = A \subseteq G$ be pre-processing sets computed by SLP's $\mathcal{L}_1 \prec \cdots \prec \mathcal{L}_r = \mathcal{L}$ from $T_1 \subset \cdots \subset T_r = T$ according to Lemma 3.1. Recall that the reduced length of $\mathcal{L}_i$ is at most $\log^2 |\langle T_i\rangle|$.

The formula ψ in free variables corresponding to the elements of T and A expresses that A is a pre-processing set for G over T. We first build a formula ψ_0 in the same free variables. We start with a prenex of existential quantifiers that refers to the sequence of elements of the SLP $\mathcal{L}$ that are not in A. The formula ψ_0 expresses each member of the sequence as the product of two previous elements, or inverse of a previous element, according to $\mathcal{L}$. Then ψ_0 has length $O(\log^2 |G|)$.

To build ψ, we use the formulas $\alpha_{|T_i|}(y, T_i)$ and $\alpha_{|A_i|}(y, A_i)$ from Lemma 2.4, of length $O(\log |G_i|)$, to also express that $\langle A_i\rangle = \langle T_i\rangle, i \leq r$. Then ψ has length $O(\log^2 |G|)$. ∎

The formulas α_k in Lemma 2.4 have about $2 \log v$ quantifier alternations for $k > 0$, and use negation. Via the argument in Lemma 3.1, we can obtain existential formulas without negation symbols that are somewhat longer.

Lemma 3.5: *For each pair of positive integers k, v, there exists an existential negation-free first-order formula $\beta(g; x_1, \ldots, x_k)$ of length $O(k \log v + \log^2 v)$ such that for each group G of size at most v,*

$$G \models \beta(g; x_1, \ldots, x_k) \text{ if and only if } g \in \langle x_1, \ldots, x_k\rangle.$$

We note that in applications we will have $k \leq \log v$ so that the length is $O(\log^2 v)$.

Proof. The formula describes the generation of a SLP computing g from $x_1, \ldots, x_k$ according to the proof of Lemma 3.1, for a single set S. The existentially quantified variables $z_1, \ldots, z_s$, where $s = \log v$, correspond to the pre-processing set A, while z_{s+1} equals g. The rest of the formula expresses that each z_{t+1} for $t < s$ has the form $v_0^{-1} v_1 x$ for $v_0, v_1 \in K(t)$ and a generator x, and that $g = v_0^{-1} v_1$ for $v_0, v_1 \in K(s)$. In detail, let

$$\begin{aligned}\beta(g; x_1, \ldots, x_k) \equiv\\ &\exists z_1, \ldots, z_{s+1} \exists y_0, \ldots, y_s [g = z_{s+1} \wedge y_s = 1 \wedge (\bigwedge_{0 \le i < s} \bigvee_{1 \le r \le k} y_i = x_r) \wedge \\ &\bigwedge_{0 \le t \le s} \exists p_0, \ldots, p_t \exists q_0, \ldots, q_t [p_0 = q_0 = 1 \wedge z_{t+1} = p_t^{-1} q_t y_t \wedge \\ &\bigwedge_{0 \le j < t} [(p_{j+1} = p_j z_j \vee p_{j+1} = p_j) \wedge (q_{j+1} = q_j z_j \vee q_{j+1} = q_j)]]].\end{aligned}$$

Clearly β has length $O(k \log v + \log^2 v)$. ∎

4. Describing finite fields and finite difference fields

Recall that a finite field $\mathbb{F}$ has size $q = p^n$ where p is a prime called the characteristic of $\mathbb{F}$. For each such q there is a unique field $\mathbb{F}_q$ of size q. Let Frob_p denote the Frobenius automorphism $x \to x^p$ of $\mathbb{F}_q$. The group of automorphisms of $\mathbb{F}_q$ is cyclic of order n with Frob_p as a generator. In particular, $(\mathrm{Frob}_p)^n$ is the identity on $\mathbb{F}_q$.

A *difference field* $(\mathbb{F}, \sigma)$ is a field $\mathbb{F}$ together with a distinguished automorphism σ. Examples are the field of complex numbers with complex conjugation and finite fields of characteristic p with a fixed power of the Frobenius automorphism. We show that finite fields and finite difference fields are log-compressible in the language of rings $L(+, \times, 0, 1)$, extended by a unary function symbol σ in the second case. Besides providing another example for our main Definition 1.1, this will be used in one case of the proof of our first main result, Theorem 1.2.

Proposition 4.1:
(i) *For any finite field $\mathbb{F}_q$, there is a Σ_3-sentence φ_q of length $O(\log q)$ in $L(+, \times, 0, 1)$ describing $\mathbb{F}_q$.*
(ii) *For any finite difference field $(\mathbb{F}_q, \sigma)$ there is a Σ_3-sentence $\psi_{q,\sigma}$ of length $O(\log q)$ in $L(+, \times, 0, 1, \sigma)$ describing $\langle \mathbb{F}_q, \sigma \rangle$.*

(iii) *For any finite field $\mathbb{F}_q$, and any $c \in \mathbb{F}_q$, there is a Σ_3-formula $\varphi_c(x)$ of length $O(\log q)$ in $L(+, \times, 0, 1)$ describing the structure $\langle \mathbb{F}_q, c\rangle$.*

Proof. (i) The sentence φ_q says that the structure is a field of characteristic p such that for all elements x we have $x^{p^n} = x$ and there is some x with $x^{p^{n-1}} \neq x$. By Lemma 2.1 one can ensure that $|\varphi_q| = O(\log q)$ and the sentence φ_q is Σ_3.

(ii) Since any automorphism of $\mathbb{F}_q$ is of the form $(\mathrm{Frob}_p)^k$ for some $k \leq n$, we can use Lemma 2.1 again to find a sentence of length $O(\log q)$ expressing that $\sigma(x) = x^{p^k}$ for each x.

(iii) By (i) it suffices to give a formula that determines c within $\mathbb{F}_q$ up to an automorphism of the field. Let $q = p^n$ as above. Since $\mathbb{F}_q$ is a Galois extension of $\mathbb{F}_p$ of degree n, any $c \in \mathbb{F}_q$ is determined within $\mathbb{F}_q$ up to an automorphism of the field by being a zero of its minimal polynomial over $\mathbb{F}_p$. This polynomial has degree at most n. So we can apply Lemma 2.1 repeatedly to express that c is a zero of the polynomial by a formula of length $O(n \cdot \log p) = O(\log q)$. ∎

Corollary 4.2: *The class of finite fields is strongly* log*-compressible (as defined after Definition 1.1) via Σ_3-sentences in $L(+, \times, 0, 1)$.*

Proof. By Proposition 4.1(iii), a generator b of the multiplicative group of $\mathbb{F}_q$ can be determined within $\mathbb{F}_q$ up to automorphism by a Σ_3-formula $\varphi_q(x)$ of length $O(\log q)$. In order to determine a finite tuple of field elements up to automorphism, it thus suffices to pin down the corresponding tuple of exponents of b. Since these exponents are bounded by $q - 1$, via Lemma 2.1 this can be done with a formula of length $O(\log q)$. ∎

The following shows that the upper bound of $O(\log q)$ on the length of a sentence describing $\mathbb{F}_q$ is close to optimal for infinitely many q.

Proposition 4.3: *There is a constant $k > 0$ such that for infinitely many primes q, for any description φ for $\mathbb{F}_q$, we have*

$$\log(q) \leq k|\varphi| \log |\varphi|.$$

Proof. Let $C(n)$ denote the Kolmogorov complexity of the binary expansion of a natural number n. A sentence φ describing $\mathbb{F}_q$ also yields a description of the number q. Therefore $C(q) \leq k'|\varphi| \log |\varphi|$ for some k', where the corrective factors are needed because the string φ over an infinite alphabet has to be encoded by a binary string in order to serve as a description in the sense of Kolmogorov complexity.

Infinitely many $n \in \mathbb{N}$ are random numbers, in that $C(n) =^+ \log_2 n$ (the superscript $+$ means that the inequality holds up to a constant). Now let $q = p_n$, the n-th prime number, so that $C(q) =^+ C(n) =^+ \log_2 n$. By the prime number theorem $p_n/\ln(p_n) \leq 2n$ for large n, so that $\log(q/\ln q) \leq^+ \log_2 n$. Note that $\sqrt{q} \leq q/\ln q$ for $q \geq 3$ so that $\log q - 1 \leq \log(q/\ln q)$. Choosing $k \geq k'$ appropriately and putting the inequalities together, we obtain $\log q \leq k|\varphi| \log |\varphi|$ as required. ∎

A similar argument shows that Proposition 2.2, for descriptions of cyclic groups of prime order, is close to optimal.

5. Describing finite simple groups

The main result of this section is the following.

Theorem 1.2: *The class of finite simple groups is* log*-compressible.*

We do not know whether the class of finite simple groups is *strongly* log-compressible (cf. Lemma 5.2, but see also Propositions 5.8). For the proof of Theorem 1.2, recall that any finite simple group belongs to one of the following classes:

(1) the finite cyclic groups C_p, p a prime;
(2) the alternating groups $A_n, n \geq 5$;
(3) the finite simple groups $L_n(\mathbb{F}_q)$ of fixed Lie type L and Lie rank n, possibly twisted, over a finite field $\mathbb{F}_q$;
(4) the 26 sporadic simple groups.

See e.g. [12], Section 1.2.

5.1. Short first-order descriptions via short presentations.

Clearly for the proof of Theorem 1.2 we may disregard the finite set of sporadic simple groups. For most of the other classes we will use that there exist short presentations. Recall that a finite presentation of a group G is given by a normal subgroup N of a free group $F(x_1, \ldots, x_k)$ such that $G = F(x_1, \ldots, x_k)/N$, and N is generated as a normal subgroup by relators $r_1, \ldots, r_m$. One writes $G = \langle x_1, \ldots, x_k \mid r_1, \ldots, r_m \rangle$.

Definition 5.1: We define the **length** of a presentation

$$G = \langle x_1, \ldots, x_k \mid r_1, \ldots, r_m \rangle$$

to be $k + \sum_j |r_j|$, where $|r_j|$ denotes the length of the relator r_j expressed as a word in the generators x_i and their inverses.

Lemma 5.2: *Suppose that a finite simple group G has a presentation $\langle x_1, \ldots, x_k \mid r_1, \ldots, r_m \rangle$ of length ℓ. Let g_i be the image of x_i in G, $i = 1, \ldots, k$.*

(i) *There is a sentence ψ of length $O(\log |G| + \ell)$ describing the structure $\langle G, \overline{g} \rangle$.*

(ii) *There is a Σ_3-sentence ψ of length $O(\log^2 |G| + \ell)$ describing the structure $\langle G, \overline{g} \rangle$, provided that $k \leq \log |G|$.*

Proof. (i) Let ψ be

$$x_1 \neq 1 \wedge \bigwedge_{1 \leq i \leq m} r_i = 1 \wedge \forall y \, \alpha_k(y; x_1, \ldots, x_k),$$

where α_k is the formula from Lemma 2.4 of length $O(k + \log |G|)$ expressing that y is generated from the x_i within G. Replacing the $x_1, \ldots, x_k$ by new constant symbols, the models of the sentence thus obtained are the nontrivial quotients of G. Since G is simple, this sentence describes $\langle G, \overline{g} \rangle$.

(ii) is similar, using the formula β_k from Lemma 3.5 instead of α_k. ∎

For most classes of finite simple groups, Guralnick et al. [7] obtained a presentation for each member G that is very short compared to $|G|$.

Theorem 5.3 ([7, Thm. A]): *There is a constant C_0 such that any nonabelian finite simple group, with the possible exception of the Ree groups of type ${}^2G_2$, has a presentation with at most C_0 generators and relations and length at most $C_0(\log n + \log q)$, where n denotes the Lie rank of the group and q the order of the corresponding field.*

Note that, following Tits, they considered the alternating groups A_n as groups of Lie rank $n-1$ over the "field" $\mathbb{F}_1$ with one element. For more detail see their remark before [7, Thm. A].

Proposition 5.4: (i) *The class of finite simple groups, excluding the Ree groups of type ${}^2G_2$, is* log*-compressible.*

(ii) *The same class is* $\log^2$*-compressible using Σ_3-sentences.*

Proof. For cyclic simple groups, this follows from Proposition 2.2. Now consider a finite simple group $G = L_n(\mathbb{F}_q)$, that is, G is of Lie rank n with corresponding field $\mathbb{F}_q$. Suppose G is not a Ree group of type ${}^2G_2$. We have $\log n + \log q \leq \log |G|$: This is clear for the alternating groups A_n because $q = 1$ and $|A_n| = n!/2$. Otherwise, the calculations of sizes of finite simple groups in e.g. http://en.wikipedia.org/wiki/List_of_finite_simple_groups (August 2014) or Wilson [12] show that $|G|$ is at least q^n.

Now by the foregoing theorem, together with Lemma 5.2 (i) replacing the constants by variables x_i, we obtain a formula $\psi(x_1, \dots, x_{C_0})$ of length $O(\log |G|)$. Then the sentence $\varphi \equiv \exists x_1 \cdots \exists x_{C_0}\, \psi$ is as required for (i). For (ii) we use Lemma 5.2 (ii) instead. ∎

We also note the following:

Proposition 5.5: *For any function $f : \mathbb{N} \to \mathbb{N}$, the class of finite groups G with a presentation of total length $f(|G|)$ is strongly $(f + \log^2)$-compressible.*

Proof. Suppose G has a presentation $G = \langle x_1, \dots, x_k \mid r_1, \dots, r_m \rangle$ of length $f(|G|)$. Fix a composition series

$$1 \lhd G_1 \lhd \cdots \lhd G_r = G$$

and an ascending sequence of swift generating sets (see Corollary 3.2)

$$A_0 \subset A_1 \subset \cdots \subset A_r = A$$

with $\langle A_i \rangle = G_i, |A_i| \leq \log |G_i|, 0 \leq i \leq r$. Note that $r \leq \log |G|$.

We start with a prenex of existential quantifiers referring to the elements of A and then express that for each i the subgroup generated by A_i is a proper normal subgroup of the subgroup generated by A_{i+1}, using the α_k from Lemma 2.4 for $k = |A_i|$. This takes length $O(\log^2 |G|)$.

We next express the $x_1, \dots, x_k$ as words over the preprocessing set A. This takes a length of $2|A| \cdot k$. We note that the formula

$$\bigwedge_{1 \leq i \leq m} r_i = 1 \wedge \forall y\, \alpha_k(y; x_1, \dots, x_k),$$

holds in a group $(H, \overline{h})$ if and only if $(H, \overline{h})$ is a quotient of $(G, \overline{g})$ where $\overline{h}, \overline{g}$ are the images of $\overline{x}$ in H and G, respectively. Since a composition series of a proper quotient of G is shorter than r, we see that the conjunction of these three formulas describes $(G, \overline{g})$ with a length of $O((f + \log^2)|G|)$. For strong compressibility, note that any tuple of elements from G can be written as a word of length $2|A|$ over A. ∎

It was shown in [1] that any finite group G without a composition factor of type ${}^2G_2$ has a presentation of length $O(\log^3|G|)$. Hence we obtain:

Corollary 5.6: *The class of finite solvable groups, and more generally of groups without a composition factor of type ${}^2G_2$, is strongly* $\log^3$*-compressible.*

While this also follows from our main result Theorem 1.3 proved below, it is interesting to note that this restricted form can be obtained already at this stage.

Remark 5.7: Using the argument of Proposition 5.5, one can see that if a class of finite groups is f-compressible for some function $f : \mathbb{N} \to \mathbb{N}$, then this class is strongly $(f + \log^2)$-compressible. However, we do not know whether the class of finite simple groups is strongly log-compressible.

5.2. Short first-order descriptions via interpretations. It remains to treat the class of Ree groups of type ${}^2G_2$. While the Chevalley groups of type G_2 exist over any field $\mathbb{F}$ as the automorphism group of the octonion algebra over $\mathbb{F}$, the (twisted) groups ${}^2G_2$ exist only over fields of characteristic 3 which have an automorphism σ with square the Frobenius automorphism. For a finite field $\mathbb{F}_q$, this happens if and only if $q = 3^{2k+1}$. The untwisted group can be presented as a matrix group over such a field. The twisted group can be seen as the group of fixed points under a certain automorphism of G_2 arising from the symmetry in the corresponding Dynkin diagram, which induces σ on the entries of the matrix (see e.g. [6, Section 13.4]).

Strong r-compressibility was introduced after Definition 1.1.

Proposition 5.8: *The class of Ree groups of type ${}^2G_2$ is strongly* log*-compressible via Σ_d-sentences for some constant d.*

No short presentations are known for these Ree groups. Instead, we use first-order interpretations between groups and finite difference fields in order to derive the proposition from Lemma 4.1.

Suppose that L, K are languages in finite signatures. Interpretations via first-order formulas of L-structures in K-structures are formally defined, for instance, in [8, Section 5.3]. Informally, an L-structure G is interpretable in a K-structure

F if the elements of G can be represented by tuples in a definable k-ary relation D on F, in such a way that equality of G becomes an F-definable equivalence relation $\approx$ on D, and the other atomic relations on F are also definable.

A simple example is the field of fractions of a given intergral domain, which can be interpreted in the domain. For an example more relevant to this paper, fix $n \geq 1$. For any field $\mathbb{F}$, the linear group $SL_n(\mathbb{F})$ can be interpreted in $\mathbb{F}$. A matrix B is represented by a tuple of length $k = n^2$, D is given by the first-order condition that $\det(B) = 1$, and $\approx$ is equality of tuples. The group operation of $SL_n(\mathbb{F})$ is then given by matrix multiplication, and can be expressed in a first-order way using the field operations.

We think of the interpretation of F in G as a decoding function Δ. It decodes F from G using first-order formulas, so that $F = \Delta(G)$ is an L-structure.

Definition 5.9: Suppose that L, K are languages in a finite signature, and that classes $\mathcal{C} \subseteq M(L), \mathcal{D} \subseteq M(K)$ are given. We say that a function Δ as above is a **uniform interpretation of $\mathcal{C}$ in $\mathcal{D}$** if for each $G \in \mathcal{C}$, there is $F \in \mathcal{D}$ such that $G = \Delta(F)$.

Note that if Δ is a uniform interpretation of $\mathcal{C}$ in $\mathcal{D}$, then there is some $k \in \mathbb{N}$, namely the arity of the relation D, such that for $G = \Delta(F)$ we have $|G| \leq |F|^k$.

For example, the class of special linear groups $SL_2(\mathbb{F})$ over finite fields $\mathbb{F}$ is uniformly interpretable in the class of finite fields via the decoding function Δ given by the formulas above.

Suppose K' is the signature K extended by a finite number of constant symbols. Let $\mathcal{D}'$ be the class of K'-structures, i.e. K-structures giving values to these constant symbols. We say that a function Δ based on first-order formulas in K' is a *uniform interpretation of $\mathcal{C}$ in $\mathcal{D}$ with parameters* if Δ is a uniform interpretation of $\mathcal{C}$ in $\mathcal{D}'$.

We will apply the following proposition to the class $\mathcal{C}$ of finite Ree groups of type ${}^2G_2(q)$, and the class $\mathcal{D}$ of finite difference fields for which these Ree groups exist.

PROPOSITION 5.10: *Suppose that L, K are languages in a finite signature, and that classes $\mathcal{C} \subseteq M(L), \mathcal{D} \subseteq M(K)$ are given. Suppose furthermore that*

(1) *there is a uniform interpretation Δ without parameters of $\mathcal{C}$ in $\mathcal{D}$,*

(2) *there is a uniform interpretation Γ with parameters of $\mathcal{D}$ in $\mathcal{C}$, and*

(3) *there is an* L*-formula* η *involving parameters such that for each* $G \in \mathcal{C}$ *there is a list of parameters* $\overline{p}$ *in* G *so that* η *defines an isomorphism between* G *and* $\Delta(\Gamma(G, \overline{p}))$*. The following hold.*

(i) *If* $\mathcal{D}$ *is* log*-compressible, then so is* $\mathcal{C}$*.*
(ii) *If* $\mathcal{D}$ *is strongly* log*-compressible, then so is* $\mathcal{C}$*.*

Proof. Let $G \in \mathcal{C}$, so that $G = \Delta(F)$ for some $F \in \mathcal{D}$. Let φ be a sentence of length $O(\log |F|)$ describing F. The sentence ψ expresses the following about an L-structure H:

there are parameters $\overline{q}$ *in* H *such that* $\Gamma(H, \overline{q}) \models \varphi$ *and*
η *describes an isomorphism* $H \cong \Delta(\Gamma(H, \overline{q}))$*.*

We claim that ψ describes G. To see this, note that certainly $G \models \psi$ via $\overline{p}$. If $\widetilde{G}$ is an L-structure satisfying ψ via a list of parameters $\overline{q}$, then $\Gamma(\widetilde{G}, \overline{q}) \models \varphi$ implies that $\Gamma(\widetilde{G}, \overline{q}) \cong F$, so that $\widetilde{G} \cong \Delta(F) \cong G$.

To see that $|\psi| = O(\log(|G|))$, recall that the uniform interpretations are by definition based on fixed sets of formulas. Therefore $|\psi| = O(|\varphi|)$. Since $\log |G| = O(\log |F|)$ by the remark after Definition 5.9, we have $|\psi| = O(\log |G|)$. This shows (i).

To prove (ii) suppose that $G \in \mathcal{C}, G = \Delta(F)$ as above. Suppose g is a tuple in G; for notational simplicity assume its length is 1. Then g is given by a k-tuple u in F for fixed k; we denote this by $(G, g) = \Delta(F, u)$. This tuple in turn is given by a $k \cdot l$-tuple w in G when an appropriate list $\overline{q}$ of parameters is fixed; we write $(F, u) = \Gamma(G, \overline{q}, w)$.

Now by hypothesis on $\mathcal{D}$ there is a formula $\theta(x_1, \ldots, x_k)$ of length $O(\log(|F|)$ describing (F, u). Obtain a formula $\chi(y)$ by adding to the expression for ψ above the condition on y that there is a $k \cdot l$ tuple w of elements of H such that $\Gamma(G, \overline{q}, w)$ satisfies θ, and $\Delta(\Gamma(G, \overline{q}, w)) = (H, y)$. Then $|\chi| = O(\log |G|)$ and χ describes (G, g). ∎

Note that if φ is a Σ_k sentence, then ψ is a Σ_{k+c} sentence for a constant c depending only on the interpretations and the formula η. Thus, if $\mathcal{D}$ is log-compressible using Σ_k sentences, then $\mathcal{C}$ is log-compressible using Σ_{k+c} sentences.

The previous proposition allows us to deal with the class of Ree groups of type ${}^2G_2$ using a result of Ryten. Note that the class of difference fields

$(\mathbb{F}_{3^{2k+1}}, \mathrm{Frob}_3^{k+1})$, $k \in \mathbb{N}$, is denoted $\mathcal{C}_{(1,2,3)}$ there. The following is a special case of the more general result of Ryten.

Theorem 5.11 (by [11], Prop. 5.4.6(iii)): *Let $\mathcal{C}$ be the class of finite groups ${}^2G_2(q)$, $q = 3^{2k+1}$, and let $\mathcal{D}$ be the class of finite difference fields $(\mathbb{F}_{3^{2k+1}}, \mathrm{Frob}_3^{k+1})$. The hypotheses of Proposition 5.10 can be satisfied via uniform interpretations Δ, Γ and a formula η in the language of groups.*

The details of the proof are contained in Ch. 5 of [11]. Since they require quite a bit of background on simple groups of Lie type, we merely indicate how to obtain the required formulas. The group ${}^2G_2(\mathbb{F})$ has Lie rank 1, and hence behaves similarly to the group $SL_2(\mathbb{F})$, which also has Lie rank 1. The formulas required for Proposition 5.10 are essentially the same in both cases. Since most readers will be more familiar with $SL_2(\mathbb{F})$, we use this group rather than ${}^2G_2(\mathbb{F})$ to make the required subgroups more explicit.

The uniform interpretation Δ of $\mathcal{C}$ in $\mathcal{D}$ is essentially the same as in the case of the interpretation of $SL_2(\mathbb{F})$ in $\mathbb{F}$ described above using the fact that $G_2(\mathbb{F})$ — and hence its subgroup ${}^2G_2(\mathbb{F})$ — has a linear representation as a group of matrices. The groups $G_2(\mathbb{F})$ are uniformly definable in $\mathbb{F}$ (as matrix groups which preserve the octonian algebra on $\mathbb{F}$). The subgroups ${}^2G_2(\mathbb{F})$ of $G_2(\mathbb{F})$ are then uniformly defined in the language of difference fields by expressing that its elements induce linear transformations (of the affine group $G_2(\mathbb{F})$) that commute with the field automorphism σ.

The uniform interpretation with parameters Γ of $\mathcal{D}$ in $\mathcal{C}$ can be given roughly as follows: for the group ${}^2G_2(\mathbb{F})$, the *torus* T and the *root subgroups* U_+, U_- of ${}^2G_2(\mathbb{F})$ are uniformly definable subgroups (in the language of groups) using parameters from the group.

In the case of the group $SL_2(\mathbb{F})$, the torus is (conjugate to) the group T of diagonal matrices in $SL_2(\mathbb{F})$ which can be defined uniformly as the centralizer of a nontrivial element h in T. (The same holds for the group ${}^2G_2(\mathbb{F})$.)

The root group U_+ of $SL_2(\mathbb{F})$ can be described as the upper triangular matrices with 1's on the diagonal, similarly U_- are the strict lower triangular matrices. The groups U_+, U_- are isomorphic to the additive group of the field $\mathbb{F}$ (this is easy to see in the case of $SL_2(\mathbb{F})$). The torus T acts by conjugation on U_+, U_- as multiplication by the squares in $\mathbb{F}$. As the characteristic of $\mathbb{F}$ is 3, any element of $\mathbb{F}$ is the difference of two squares. Thus the groups U_+, U_- can be defined uniformly by picking a nontrivial element u in U_+, U_-, respectively and considering the orbit $\{u^h : h \in T\}$ of u under the conjugation by elements from T. Writing the group operation on U_+, U_- additively, the set of differences $\{u^h - u^{h'} : h, h' \in T\}$ is uniformly definable and defines the root groups. This also shows that from $U_+ \rtimes T$ we definably obtain the field $\mathbb{F}$. Again, for ${}^2G_2(\mathbb{F})$ this is essentially the same.

It remains to find a formula describing the isomorphism $\eta \colon H \cong \Delta(\Gamma(H, \overline{q}))$ for a group $H \in \mathcal{C}$ and an appropriate list of parameters including the ones given above. For this we need the fact that by the Bruhat decomposition (see [6], Ch. 8, in particular 8.2.2) we have ${}^2G_2 = BNB = B \cup BsB$, where in this case $B = U_+T$, N is the normalizer of T and s is (the lift of) an involution generating the *Weyl group* N/T of ${}^2G_2$. Thus any element of ${}^2G_2$ (or in fact of any group of Lie type of Lie rank 1) can be written uniquely either as a product of the form u_1h or of the form u_1hsu_2 where $u_1, u_2 \in U_+, h \in T$ and s is a fixed generator of the Weyl group of ${}^2G_2$, i.e. $s \notin T$ normalizes T and $s^2 \in T$. This yields the required isomorphism η. ∎

Proof of Proposition 5.8. By Theorem 5.11 the class $\mathcal{C}$ of Ree groups of type ${}^2G_2$ is uniformly parameter interpretable in the class $\mathcal{D}$ of finite difference fields $(\mathbb{F}_{3^{2k+1}}, \mathrm{Frob}_3^{k+1})$. By Corollary 4.2, the class $\mathcal{D}$ is strongly log-compressible using Σ_3 sentences. By Proposition 5.10 (and the remark after its proof), this implies that the class $\mathcal{C}$ is strongly log-compressible via Σ_d sentences for some constant d. (We estimate that $d \leq 10$.) ∎

Remark 5.12: In fact, Ryten proves that for fixed Lie type $\mathbb{L}$ and rank n, the class of finite simple groups $\mathbb{L}_n$ is uniformly parameter bi-interpretable with the corresponding class of finite fields or difference fields. This means that in addition to the properties given in Proposition 5.10 there is a formula δ in the first-order language for K that defines for each $F \in \mathcal{D}$ an isomorphism between F and $\Gamma(\Delta(F), \overline{p})$. Via Proposition 5.10 this yields a proof that each class of finite simple groups is log-compressible. However, since there are infinitely many such classes, further effort would be needed in order to show that there is a single O-constant which works for all classes. We have circumvented the problem by using the results of Guralnick et al. [7].

Remark 5.13: By Remark 2.5 and the proofs above, each finite simple group G actually has a description of binary length $O(\log(|G|))$.

Based on the methods above we can somewhat strengthen Theorem 1.2.

Proposition 5.14: *The class of characteristically simple finite groups G is* log-*compressible.*

Proof. Any nontrivial characteristically simple finite group G is isomorphic to a direct power S^k, $k \geq 1$, where S is a simple group (see e.g. Wilson [12, Lemma 2.8]). Firstly we consider the case that S is abelian, and so cyclic of order p. The sentence describing G expresses that there are $x_1, \ldots, x_k$ of order p such that $x_1, \ldots, x_k$ generate the group (using the formulas α_k for $v = |G|$ from Lemma 2.4); we can say within length $O(\log|G|)$ that the x_r commute pairwise by expressing with two disjunctions of length $O(k)$ that

$$\forall z_1 \in \{x_1, \ldots, x_k\} \forall z_2 \in \{x_1, \ldots, x_k\} \, [z_1, z_2] = 1.$$

Now suppose that S is nonabelian. It is well-known that S can be generated by just two elements g, h. In the following let r range over $1, \ldots, k$. The sentence φ describing G starts with a block of existentially quantified variables $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$; we think of x_r, y_r as the generating set g, h in the r-th copy of S. Firstly, we require, using the α_{2k} for the size $|G|$, that the set $\{x_1, \ldots, x_k, y_1, \ldots, y_k\}$ generates the group H under consideration, that $[x_r, y_r] \neq 1$ for each r, and that

$$\forall z \in \{x_1, \ldots, x_k\} \exists^{\leq 1} w \in \{y_1, \ldots, y_k\} \, [z, w] \neq 1.$$

This ensures that the subgroup U_r generated by x_r, y_r is normal in the group H; hence so is its centraliser $C(x_r, y_r)$.

Secondly, let φ_S be a description of S with $|\varphi_S| = O(\log|S|)$ according to Theorem 1.2. We require that the center of U_r is trivial (this is possible using the formula α_2 for size $|G|$), and that $H/C(z, w) \models \varphi_S$ for each $z \in \{x_1, \ldots, x_k\}$ and $w \in \{y_1, \ldots, y_k\}$ such that $[z, w] \neq 1$. This can be done within the required length bound since $C(z, w)$ is defined by a formula of fixed length. Since $H \models \varphi$ implies $U_r \cong H/C(U_r)$ for each r, the sentence describes G. ∎

6. Background on group extensions

In this section we provide the tools needed for obtaining short first-order descriptions of general finite groups in Section 7. To obtain such descriptions, we will use a composition series of the group in question. Besides describing the simple quotients, we will also need to describe the extension of a group N by a group H. Such an extension can be understood via the second cohomology groups of certain associated modules. Here we give a more elementary account of the relevant part of the theory of group extensions, an account which we can translate into a first-order description of the extension. We consider a group extension E containing N as a normal subgroup such that $E/N \cong H$. (While all of this is in principle well-known, we include it to keep the paper self-contained in this regard.)

In contrast to the presentations of Section 5, we will use *profinite* presentations for the group H because in this setting it is known that a small number of relators suffices. So we consider a presentation

$$H \cong F/R,$$

where $F = \widehat{F}(s_1, \ldots, s_k)$ is the profinite completion of the free group of rank k on generators $s_1, \ldots, s_k$ and R is the closed normal subgroup of F topologically generated (as a normal subgroup) by $r_1, \ldots, r_m$. For details see e.g. Lubotzky and Segal [9, p. 47].

We will show that any group extension E of N by H is determined by the action of F on N, and an F-homomorphism from R into N. Such a homomorphism is determined by the generators of R as a normal subgroup, i.e., the relators for the profinite presentation, which is why we want the presentation to have as few relators as possible.

Let E be an extension of N by $H = \langle s_1, \ldots, s_k \rangle$. Let $s_1', \ldots, s_k' \in E$ be *lifts* of $s_1, \ldots, s_k \in H$, i.e. $\pi_H(s_i') = s_i, i = 1, \ldots, k$. Then the $s_i', i = 1, \ldots, k$, act on N by conjugation and hence any *word* $w(\overline{s}) = w(s_1, \ldots, s_k)$ in the profinite free group F with generators $s_1, \ldots, s_k$ acts on N (as an automorphism of N) via the natural action of $w(s_1', \ldots, s_k') \in E$. By density this extends to a unique continuous action of all of F on N. Hence any group extension E of N by a k-generated group $H = \langle s_1, \ldots, s_k \rangle$ determines an action of $F = \widehat{F}(s_1, \ldots, s_k)$ on N, where the s_i are now seen as generators of F, rather than as elements of H. In order to describe E we will have to express this action of F on N.

Define

$$\varphi_E : R \longrightarrow N \text{ by } w(s_1, \ldots, s_k) \mapsto w(s_1', \ldots, s_k')$$

and extend the definition to the unique continuous function defined on all of R. Then $\varphi_E \in \mathrm{Hom}_F(R, N)$. The next lemma states that the group E is determined — up to an isomorphism over N — by the action of F on N and the homomorphism φ_E.

LEMMA 6.1: *Using the previous notation, suppose that E^1, E^2 are groups with a common normal subgroup N and let $s_i^j \in E^j, j = 1, 2, i = 1, \ldots, k$ be lifts of $s_1, \ldots, s_k$, respectively, such that $(E^j/N, \overline{s}^j) \cong (H, \overline{s}), j = 1, 2$.*

Suppose that the induced F-actions agree, i.e. for all $a \in N$ we have

$$(*) \qquad a^{s_i^1} = a^{s_i^2}, \quad i = 1, \ldots, k.$$

Then E^1 and E^2 are isomorphic over N via an isomorphism taking s_i^1 to s_i^2, $i = 1, \ldots, k$, if and only if $\varphi_{E^1} = \varphi_{E^2}$.

Proof. First suppose that $\varphi_{E^1} = \varphi_{E^2}$. Define for $a \in N$

$$f : E^1 \longrightarrow E^2,\ aw(\overline{s}^1) \mapsto\ aw(\overline{s}^2).$$

Note that

$$aw(\overline{s}^1) = a'w'(\overline{s}^1) \Leftrightarrow w(\overline{s}^1)(w'(\overline{s}^1))^{-1} \in N \Leftrightarrow w(\overline{s})(w'(\overline{s}))^{-1} \in R.$$

Since $\varphi_{E^1} = \varphi_{E^2}$, we see that indeed f is well-defined. Exchanging the roles of E^1 and E^2 shows that f is injective.

Note that f is a homomorphism because the F-actions on N agree: let $a_0, a_1 \in N$, and let w_0, w_1 be group words in variables $\overline{s} = s_1, \ldots, s_k$. Then

$$\begin{aligned} f(a_0 w_0(\overline{s}^1) a_1 w_1(\overline{s}^1)) &= f(a_0 a_1^{w_0^{-1}(\overline{s}^1)} w_0(\overline{s}^1) w_1(\overline{s}^1)) \\ &= a_0 a_1^{w_0^{-1}(\overline{s}^2)} w_0(\overline{s}^2) w_1(\overline{s}^2) && \text{by } (*) \\ &= a_0 w_0(\overline{s}^2) a_1 w_1(\overline{s}^2) \\ &= f(a_0 w_0(\overline{s}^1)) f(a_1 w_1(\overline{s}^1)). \end{aligned}$$

Since E^j is generated by N and $\overline{s}^j, j = 1, 2$, this now implies that f is surjective and hence an isomorphism fixing N pointwise.

For the converse implication, suppose that $g : E^1 \longrightarrow E^2$ is an isomorphism fixing N pointwise and taking s_i^1 to $s_i^2, i = 1, \ldots, k,$. For any word w with $w(\overline{s}) \in R$ we have

$$g(w(\overline{s}^1)) = w(\overline{s}^1) = \varphi_{E^1}(w(\overline{s})).$$

Also

$$g(w(\overline{s}^1)) = w(\overline{s}^2) = \varphi_{E^2}(w(\overline{s})),$$

proving the lemma. ∎

A close inspection of the proof of Lemma 6.1 yields the following variant, which will be used in Section 7 for the first-order description of group extensions.

Lemma 6.2: *Suppose that in the situation of Lemma 6.1 every element of H has length at most m with respect to $\overline{s}$. Then E^1 and E^2 are isomorphic over N provided that $\varphi_{E^1}^{3m} = \varphi_{E^2}^{3m}$, where $\varphi_{E^j}^{3m}, j = 1, 2$, denotes the restriction of φ_{E^j} to the elements of R of word length at most $3m$ over $\overline{s}$.*

Proof. Define $f : E^1 \longrightarrow E^2$ by

$$aw(\overline{s}^1) \mapsto aw(\overline{s}^2),$$

where $a \in N$ and $w \in F(\overline{s})$ is a group word such that $|w| \leq m$. By assumption, f is defined on all of E^1. We verify as in the proof of Lemma 6.1 that f is well-defined and injective, noting that only words of length $\leq 2m$ are relevant now.

To check that f is a homomorphism, let $a_0, a_1 \in N$, and let w_0, w_1 be group words in variables $s_1, \dots, s_k$ of length at most m. By assumption there are $a \in N$ and a word w_2 of length at most m such that

$$w_0(\overline{s}^1)w_1(\overline{s}^1) = aw_2(\overline{s}^1).$$

Since $\varphi_{E^1}^{3m} = \varphi_{E^2}^{3m}$ we have

$$w_0(\overline{s}^2)w_1(\overline{s}^2) = aw_2(\overline{s}^2).$$

Hence as in the proof of Lemma 6.1 we have

$$\begin{aligned}
f(a_0 w_0(\overline{s}^1) a_1 w_1(\overline{s}^1)) &= f(a_0 a_1^{w_0^{-1}(\overline{s}^1)} w_0(\overline{s}^1) w_1(\overline{s}^1)) \\
&= f(a_0 a_1^{w_0^{-1}(\overline{s}^1)} a w_2(\overline{s}^1)) \\
&= a_0 a_1^{w_0^{-1}(\overline{s}^2)} a w_2(\overline{s}^2) && \text{by } (*) \\
&= a_0 w_0(\overline{s}^2) a_1 w_1(\overline{s}^2) \\
&= f(a_0 w_0(\overline{s}^1)) f(a_1 w_1(\overline{s}^1)).
\end{aligned}$$

Since $E^j, j = 1, 2$, is generated by N and $\overline{s}^j$, this now implies that f is surjective and hence an isomorphism fixing N pointwise. ∎

Recall that a group action is called *regular* if it is transitive and point stabilizers are trivial.

Lemma 6.3: *Let $Z = Z(N)$. The group* $\mathrm{Hom}_F(R, Z)$ *acts regularly on the set*

$$X = \{\varphi_E\colon E \text{ is extension of } N \text{ by } H \text{ with prescribed } F\text{-action on } N\}$$

via $\varphi_E^\psi(w(\overline{s})) = \varphi_E(w(\overline{s}))\psi(w(\overline{s}))$ *for* $\psi \in \mathrm{Hom}_F(R, Z)$ *and* $\varphi_E \in X$

Proof. To see that the action is transitive, just notice that for extensions E_1, E_2 of N by H with the given F-action on N, and lifts $s_i^j, j = 1, 2, i = 1, \ldots, k$ as before we have for all $n \in N$

$$n^{\varphi_{E_1}(w(\overline{s}))} = n^{w(\overline{s}^1)} = n^{w(\overline{s})} = n^{w(\overline{s}^2)} = n^{\varphi_{E_2}(w(\overline{s}))}$$

and hence $\varphi_{E_1}(w(\overline{s}))(\varphi_{E_2}(w(\overline{s})))^{-1} \in Z$. By continuity, φ_{E_1} and φ_{E_2} differ by an element in $\mathrm{Hom}_F(R, Z)$.

Let $\psi \in \mathrm{Hom}_F(R, Z)$. To see that $\varphi_E^\psi = \varphi_{E^1}$ for some extension E^1 with prescribed F-action on N, define E^1 by choosing a transversal T for F/R so that any element $w(\overline{s}) \in F$ can be written uniquely as

$$w(\overline{s}) = v(\overline{s})r(\overline{s})$$

where $v(\overline{s}) \in T, r(\overline{s}) \in R$.

Let $s_i^0, i = 1, \ldots, k$ be the lifts of s_i to E. We now define an extension E^1 with lifts $s_i^1, i = 1, \ldots, k$, by letting the elements of E^1 be

$$nw(\overline{s}^1) = nv(\overline{s}^0)\varphi_E(r(\overline{s}))\psi(r(\overline{s}))$$

with the induced multiplication. Then E^1 is an extension with prescribed $F = \widehat{F}(\overline{s})$ action and $\varphi_{E^1} = \varphi_E^\psi$. ∎

The *rank* of an abelian group A, denoted $\mathrm{rk}\, A$, is the minimal size of a set of generators, or, in other words, the least k such that there is an onto map $\mathbb{Z}^k \to A$. Clearly $B \leq A$ implies $\mathrm{rk}\, B \leq \mathrm{rk}\, A$. Letting λn denote the number of prime factors of n with multiplicity, we have $\mathrm{rk}\, A \leq \lambda|A| \leq \log|A|$.

Remark 6.4: Lemma 6.3 implies that the number of extensions of N by H is at most $|Z|^r$, where $Z = Z(N)$ and r is the minimum number of generators of R as a closed normal subgroup of F. The *rank* of $\mathrm{Hom}_F(R, Z)$ is at most $r \cdot \lambda|Z|$ since each $\varphi \in \mathrm{Hom}_F(R, Z)$ is determined by its values on the r generators of R.

Corollary 6.5: *If $Z(N) = 1$, then an extension E of N by H is determined up to isomorphism over N by the F-action on N.*

Lemma 6.2 states that the restriction φ_E^{3m} of φ_E to words of length at most $3m$ is sufficient for describing an extension E. To give a short description of φ_E^{3m}, we rely heavily on the following lemma originally suggested by Alex Lubotzky.

Lemma 6.6: *Let A be a finite abelian group, X a set and let $V \leq A^X$ be a subgroup of rank d. There exists a set $Y \subseteq X$ of size at most $d \cdot \lambda(|A|)$ such that for all $g \in V$, $g \restriction Y = 0$ implies $g = 0$.*

Proof. Decompose A into its p-primary components $A = \bigoplus_p A_p$. Since the number of different primes dividing the order of A is at most $\lambda(|A|)$, the lemma follows from applying Lemma 6.7 below to each of the A_p separately. ∎

Lemma 6.7: *Let A be a finite abelian p-group, X a set and let $V \leq A^X$ be a subgroup of rank d. There exists a set $Y \subseteq X$ of size at most d such that for all $g \in V$, $g \restriction Y = 0$ implies $g = 0$.*

Proof. Since A is a direct product of k cyclic p-groups for some k, we may consider $V \leq A^X \leq (C_q^k)^X \cong (C_q)^{k|X|}$ where q is the exponent of A. Then, replacing each element x of X by k new elements $\langle x, 1\rangle, \ldots, \langle x, k\rangle$, we may assume $A = C_q$ and $V \leq C_q^{X'}$ where $X' = X \times \{1, \ldots, k\}$. Once we have found a subset Y' of X' of size at most d with the required property, we obtain $Y \subseteq X, |Y| \leq d$, by replacing each element $\langle y, i\rangle \in Y'$ by y. If $g \in V$ and $g(y) = 0$ then $g(\langle y, i\rangle) = 0$ for each i when g is viewed as a function on X' with values in C_q.

Without loss of generality we may thus assume that $A = C_q$. For $x \in X$ let $g_x\colon V \to A$ denote the coordinate function mapping $p \in V$ to $p(x)$. There is nothing to show if $d = 0$, so suppose $d > 0$.

Let $x_1 \in X, v_1 \in V$ such that $g_{x_1}(v_1)$ has maximal order. Then $g_{x_1}(V) \leq \mathbb{Z}g_{x_1}(v_1)$. We claim that V decomposes as $V = \mathbb{Z}v_1 \oplus \ker(g_{x_1})$. First note that clearly $\mathbb{Z}v_1 \cap \ker(g_{x_1}) = 0$. Next, given arbitrary $w \in V$, choose $r \in \mathbb{Z}$ so that $g_{x_1}(w) = r \cdot g_{x_1}(v)$. Then $g_{x_1}(w - r \cdot v) = 0$, so that $w \in \mathbb{Z}v_1 + \ker(g_{x_1})$.

Clearly $\operatorname{rk}\ker(g_{x_1}) \leq d - 1$ and we may consider $\ker(g_{x_1})$ as a subgroup of $A^{X \setminus \{x_1\}}$. Inductively we find $x_2, \ldots, x_d \in X \setminus \{x_1\}$ with corresponding elements $v_2, \ldots, v_d$ such that $V = \bigoplus_{i \leq d} \mathbb{Z}v_i$ and $\bigcap_{i \leq d} \ker(g_{x_i}) = 0$. Hence $Y = \{x_1, \ldots, x_d\}$ is as required. ∎

Remark 6.8: The bound given in Lemma 6.6 is optimal. For suppose A is a product of n cyclic groups of different prime orders $p_1, \ldots, p_n$ with generators

$g_1, \dots, g_n$. Let $X = \{x_1, \dots, x_n\}$, and let $f(x_i) = g_i$. For each i, $\langle f \rangle$ contains an element that only differs from 0 at the i-th component.

We summarize and assemble all the pieces of this section in the following proposition, which we will use in the next section to describe arbitrary finite groups.

Proposition 6.9: *Suppose $H = F/R$ where $F = \widehat{F}(s_1, \dots, s_k)$ and R is generated as a closed normal subgroup of F by r elements. Let $\overline{s} \subset H$ be the image of $(s_1, \dots, s_k)$, so $\langle \overline{s} \rangle = H$, and suppose that any element of H has length at most m over $\overline{s}$. Let N be a finite group, and let $Z = Z(N)$.*

There are words

$$w_1, \dots, w_d \in R \text{ of length at most } 3m, \text{ where } d = r \cdot \lambda|Z|,$$

such that group extensions E^j, $j = 1, 2$, of N by H are isomorphic over N under an isomorphism taking a lift $\overline{s}^1 \in E^1$ of $\overline{s}$ to a lift $\overline{s}^2 \in E^2$, provided the following conditions hold:

(a) $a^{s_i^1} = a^{s_i^2}, i = 1, \dots, k$ *for all* $a \in N$;
(b) $(E^1/N, \overline{s}^1) \cong (E^2/N, \overline{s}^2)$;
(c) $w_i(\overline{s}^1) = w_i(\overline{s}^2) \in N$ *for* $i = 1, \dots, d$.

Proof. By Lemmas 6.2 and 6.3, the abelian group $\mathrm{Hom}_F(R, Z)$ can be seen as a subgroup of Z^X where X is the set of group words in $s_1, \dots, s_k$ of length $\leq 3m$. Note that we have $\mathrm{Hom}_F(R, Z) = \bigoplus_p \mathrm{Hom}_F(R, Z_p)$ where Z_p are the p-primary components of Z. Now for each prime p, the group $\mathrm{Hom}_F(R, Z_p)$ can be seen as a subgroup of Z_p^X and, by Remark 6.4, $\mathrm{rk}\,\mathrm{Hom}_F(R, Z_p) \leq r \cdot \lambda|Z_p|$. Since $\sum_p r \cdot \lambda|Z_p| = r \cdot \lambda|Z| = d$, we can use Lemma 6.6 to find the required $w_1, \dots, w_d \in X$. ∎

We will apply the previous proposition in the situation where H is a finite simple group, $\{s_1, \dots, s_k\} \subseteq H$ is a swift generating set of H of size at most $\log|H|$ and H has a profinite presentation $H \cong F/R$, where R is generated as a closed normal subgroup of F by $O(\log|H|)$ elements. The existence of such a profinite presentation is guaranteed by results in [9] and [7]:

Theorem 6.10 ([9]): *There is a constant C such that any finite simple group generated by d elements has a profinite presentation with d generators and $C+d$ relations.*

Proof. This follows from Theorem B of [7] and the proof of Theorem 2.3.3 of [9] in the case of simple groups. For the latter, we refer to the proof that Conjecture B implies Conjecture A in [9]. This proof also works for profinite presentations as verified by the authors after Thm. 3.3. ∎

7. Describing general finite groups

We are now in a position to give short descriptions of arbitrary finite groups.

THEOREM 1.3: *The class of finite groups is strongly* $\log^3$*-compressible.*

Proof. Let G be a finite group. We fix a subnormal series

$$1 = G_0 \lhd G_1 \lhd \cdots \lhd G_r = G$$

with simple factors $H_i := G_i/G_{i-1}$, $i = 1, \ldots, r$.

Note that the length r is bounded by $\log |G|$.

Choose an ascending sequence of sets

$$\emptyset = T_0 \subset T_1 \subset \cdots \subset T_r = T$$

with $\langle T_i \rangle = G_i, 0 \leq i \leq r \leq \log |G|$, as follows.

(1) If H_i is a finite simple group not of type ${}^2G_2$, we let $T_i = T_{i-1} \cup \{s_1, \ldots, s_C\}$ for any elements $s_1, \ldots, s_C \in G_i$ such that $s_1G_{i-1}, \ldots, s_CG_{i-1}$ are generators for $G_i/G_{i-1} = H_i$ according to Theorem 5.3. Then $(H_i, s_1G_{i-1}, \ldots, s_CG_{i-1})$ can be described by a sentence φ_i of length $O(\log |H_i|)$ by Proposition 5.4(i) and its proof.

(2) If H_i is a group of type ${}^2G_2$, we let $C = 2$ and $T_i = T_{i-1} \cup \{s_1, s_C\}$ for any elements $s_1, s_2 \in G_i$ such that s_1G_{i-1}, s_2G_{i-1} generate $G_i/G_{i-1} \cong H_i$. Then $(H_i, s_1G_{i-1}, s_2G_{i-1})$ can be described by a sentence φ_i of length $O(\log |H_i|)$ by Proposition 5.8.

Note that for each $i = 1, \ldots, r$ we have $|T_i \setminus T_{i-1}| \leq C_0$ for the constant C_0 given in Theorem 5.3. The sentence φ describing G needs to express conditions (a), (b) and (c) of Proposition 6.9 for each group G_i with normal subgroup G_{i-1}.

We start with a prenex of existential quantifiers referring to the elements of T.

1. Obtain a pre-processing set A for G over T:

To give short descriptions for the conditions of Proposition 6.9, at each step $i = 1, \ldots, r$ we first obtain pre-processing sets from T using formula ψ from Lemma 3.4 where we replace the parameters from T by the corresponding variables. The formula ψ has length $O(\log^2 |G|)$.

Since the pre-processing set A will be used in each part of the sentence φ, the scope of the existential quantifiers referring to A extends over all of φ.

2. Express $(G_i/G_{i-1}, T_i \setminus T_{i-1}) \cong (H_i, s_1 G_{i-1}, \ldots, s_C G_{i-1})$:

   We let the formula χ_i, $i = 1, \ldots, r$, express that

$$(G_i/G_{i-1}, T_i \setminus T_{i-1}) \models \varphi_i.$$

We can use the $\alpha_{|T_i|}$ to express that G_{i-1} is a normal subgroup of G_i using a length of $O(\log |G_i|)$. We now restrict the quantifiers in φ_i to G_i using $\alpha_{|T_i|}$ and replace each occurrence of "$u = v$" in φ_i by

$$\text{“}uv^{-1} \in G_{i-1}\text{”}.$$

Since we replace the equality symbols in φ_i by strings of length $O(\log |G_{i-1}|)$, the resulting formula χ_i has length $O(\log |H_i| \log |G_{i-1}|)$. Then the conjunction χ of the formulas χ_i has length $O(\log^2 |G|)$.

3. Conjugation action of G_i on G_{i-1}:

   For each $i = 2, \ldots, r$, let κ_i describe the action of $g \in T_i \setminus T_{i-1}$ on G_{i-1} by conjugation. Since T_{i-1} generates G_{i-1}, it suffices to determine $g^{-1}wg$ for each $w \in T_{i-1}$ and $g \in T_i \setminus T_{i-1}$ as an element $h_{w,g} \in G_{i-1}$. Since $h_{w,g}$ has length at most $2 \log |G_{i-1}|$ over A_{i-1} and there are at most $C_0 \cdot \log |G_{i-1}|$ such pairs, κ_i has length in $O(\log^2 |G_{i-1}|)$. The conjunction κ of the κ_i has length $O(\log^3 |G|)$.

4. Describing the extension of G_{i-1} by H_i:

   We use Theorem 6.10 to obtain a profinite presentation for H_i with a swift generating set (Corollary 3.2) of size $k \leq \log |H_i|$ corresponding to the elements of $A_i \setminus A_{i-1}$ and with $r \leq C + \log |H_i|$ relations.

   By Proposition 6.9 there is $d \leq \log |Z(G_{i-1})|(C + \log |H_i|)$, and there are words $w_1, \ldots, w_d$ in $\overline{a}_i = A_i \setminus A_{i-1}$ of length at most $3 \log |H_i|$ such that $w_j(\overline{a}_i) = h_j \in G_{i-1}, j = 1, \ldots, d$, determine G_i. Since any element of G_{i-1} has length at most $2 \log |G_{i-1}|$ over A_{i-1}, we obtain a formula ρ_i of length $O(\log |Z(G_{i-1})| \log |H_i| \log |G|)$. Since $\sum_i \log |H_i| \leq \log |G|$, the conjunction ρ of the ρ_i yields a formula of length $O(\log |G| \log^2 |G|)$.

We now let φ be the sentence consisting of the prenex of existential quantifiers referring to T followed by the conjunction of ψ, κ, χ, and ρ. By repeated application of Proposition 6.9 one verifies that φ describes G. The strong $\log^3$-compressibility of the class of finite groups follows since any element of G has length at most $2\log|G|$ over the pre-processing set A. ∎

The strong $\log^3$-compressibility of the class of finite groups allows us to also describe finite transitive permutation groups (as explained in the introduction), and finite groups with a distinguished automorphism.

Corollary 7.1: (i) *The class of finite groups with a distinguished subgroup is* $\log^3$*-compressible in the language of groups with an additional unary predicate.*

(ii) *The class of finite groups with a distinguished automorphism is* $\log^3$*-compressible in the language of groups with an additional unary function.*

Proof. (i) Given a finite group G and a subgroup $U \leq G$, choose a string $\overline{g}$ of generators for U of length $k \leq \log|G|$. Let φ be the description of $(G, \overline{g})$ obtained above. Then $|\varphi| = O(\log^3|G|)$. Use the formula α_k from Lemma 2.4 of length $O(\log|G|)$ to express that $U = \langle\overline{g}\rangle$ in G. (ii) is similar. ∎

Remark 7.2: The exponent 3 in Theorem 1.3 is optimal even for p-groups of nilpotency class 2 by a result of Higman, which states that there are at least $p^{\frac{2}{27}n^2(n-6)}$ non-isomorphic such groups of order p^n (see e.g. [3, Thm. 4.5]). This result is applied in a way similar to the proof of [1, Prop. 8.6].

We provide an upper bound on the length of descriptions when only a bounded number of quantifier alternations is allowed.

Theorem 7.3: *For some m, the class of finite groups G is* $\log^4$*-compressible via Σ_m sentences.*

Proof. We only note the necessary modifications to the previous arguments. Throughout, instead of the α_k we use the existential generation formulas β_k from Lemma 3.5, which have length $O(\log^2|G|)$ because $k \leq \log|G|$ throughout.

The new version of ψ in Lemma 3.4 now has length $O(\log^3|G|)$. For some small d we can choose Σ_d-descriptions φ_i of H_i of length $O(\log^2|H_i|)$ via Propositions 5.4 and 5.8. Since we now replace the equality symbols in φ_i by strings of length $O(\log^2|G_{i-1}|)$, the resulting new version of the formula χ_i has length

$O(\log|H_i|\log^2|G_{i-1}|)$, and their conjunction has length $O(\log^3|G|)$. No generation formulas are used elsewhere in the proof of Theorem 1.3, so we conclude the argument as before. It is clear that the number of quantifer alternations is now bounded. ∎

Remark 7.4:
1. We don't know whether the exponent can be improved to 3 in Theorem 7.3.
2. Reviewing the proof of Theorem 1.3, it would be interesting to show a stronger compressibility result for the class of finite groups without a nontrivial abelian normal subgroup. In this case, we have $Z(G_i) = 1$ for each i, so that Step 4 is not needed.

Remark 7.5: Theorem 1.2 leaves open some questions. It would be interesting to show log-compressibility for classes of finite groups G that are in some sense close to simple, similar to Proposition 5.14. These include the central extensions of simple groups, and the almost simple groups (that is, $S \le G \le \mathrm{Aut}(S)$ for some simple group S). Also, it would be desirable to show the strong log-compressibility for the class of finite simple groups.

Acknowledgements. The authors wish to thank Alex Lubotzky, Bill Kantor, and Martin Ziegler for helpful discussions. This research was partially supported by the Marsden fund of New Zealand, a Hood fellowship by the Lions foundation of New Zealand, and the DFG through SFB 878.

References

[1] L. Babai, A. J. Goodman, W. M. Kantor, E. M. Luks and P. P. Pálfy, *Short presentations for finite groups*, J. Algebra **194** (1997), 79–112.

[2] L. Babai and E. Szemeredi, *On the complexity of matrix group problems I*, in *Proceedings of the 25th Annual Symposium on Foundations of Computer Science, 1984*, IEEE Computer Society, Washington, DC, 1984, pp. 229–240.

[3] S. R. Blackburn, P. M. Neumann and G. Venkataraman, *Enumeration of Finite Groups*, Cambridge Tracts in Mathematics, Vol. 173, Cambridge University Press, Cambridge, 2007.

[4] J. N. Bray, M. D. E. Conder, C. R. Leedham-Green and E. A. O'Brien, *Short presentations for alternating and symmetric groups*, Trans. Amer. Math. Soc. **363** (2011), 3277–3285.

[5] H. Buhrman, M. Li, J. Tromp and P. Vitányi, *Kolmogorov random graphs and the incompressibility method*, SIAM J. Comput. **29** (1999), 590–599.

[6] R. W. Carter, *Simple groups of Lie type*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1989, Reprint of the 1972 original, A Wiley-Interscience Publication.

[7] R. M. Guralnick, W. M. Kantor, M. Kassabov and A. Lubotzky, *Presentations of finite simple groups: a quantitative approach*, J. Amer. Math. Soc. **21** (2008), 711–774.

[8] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, Vol. 42, Cambridge University Press, Cambridge, 1993.

[9] A. Lubotzky and D. Segal, *Subgroup growth*, Progress in Mathematics, Vol. 212, Birkhäuser Verlag, Basel, 2003.

[10] A. Mann, *Enumerating finite groups and their defining relations*, J. Group Theory **1** (1998), 59–64.

[11] M. J. Ryten, *Model theory of finite difference fields and simple groups*, PhD thesis, University of Leeds (2007), Available at http://www.logique.jussieu.fr/modnet/ as preprint no. 123.

[12] R. A. Wilson, *The finite simple groups*, Graduate Texts in Mathematics, Vol. 251, Springer-Verlag London, Ltd., London, 2009.