DESCRIBING FINITE GROUPS BY SHORT FIRST-ORDER SENTENCES

ANDRÉ NIES AND KATRIN TENT

Abstract. We say that a class of finite structures for a finite signature is \( r \)-compressible 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. The first result relies on the classification of finite simple groups, and the bi-interpretability of the small Ree groups with finite difference fields. We also indicate why these results are close to optimal.

1. Introduction

The expressiveness of first-order logic in finite structures is a recurrent theme of theoretical computer science. Let \( L \) be first-order logic in a signature consisting of finitely many relation symbols, function symbols, and constants. We study the compressibility of a finite \( L \)-structure \( G \) up to isomorphism via descriptions that are sentences from \( L \). We say that a sentence \( \varphi \) in \( L \) describes \( G \) if \( G \) is the unique model of \( \varphi \) up to isomorphism.

Note that every finite \( L \)-structure \( G \) can be described by some sentence \( \varphi \): 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 \( G \). On the other hand, a description of \( G \) that is much shorter than \( |G| \) can be thought of as a compression of the isomorphism type of \( G \).

For an infinite class of \( S \)-structures, we are interested in giving first-order descriptions that are asymptotically short relative to the size of the described structure. This is embodied in the following definition, where we think of the function \( r \) as a slowly-growing.

Definition 1.1. Let \( r: \mathbb{N} \to \mathbb{N}^+ \) 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 \( \varphi \) in \( L \) such that \( |\varphi| = O(r(|G|)) \) and \( \varphi \) describes \( G \).

Note that one can verify using space polynomial in \( |G| + |\varphi| \) whether a finite structure \( G \) satisfies a sentence of \( L \). Hence, given \( G \) one can compute a shortest description \( \varphi \) of \( G \) within space polynomial in \( |G| \), where the polynomial only depends on the signature.

In this paper we will use the definition \( \log k = \max \{ r: 2^r \leq k \} \) \( (k > 0) \). As our first main result, we obtain a logarithmic rate of compression for the class of finite simple groups.
Theorem 1.2. The class of finite simple groups is $O(\log)$ compressible.

Finitely generated groups can be described up to isomorphism via presentations. There is a large amount of literature on finding very short presentations for finite simple groups, such as [1, 7, 3]. In particular the presentations are of length $O(\log |G|)$. Using simplicity, it is possible to convert such a presentation into a first-order description of $G$ of length $O(\log |G|)$ (Subsection 4.1).

The small Ree groups $2G_2(q)$ 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 $3^{2k+1}$ ([10], 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 $\sigma$ is the $3^{k+1}$-th power of the Frobenius automorphism. This was proved by Ryten [9], Prop. 5.4.6(iii)). It now suffices to give a short description of the difference field, which is not hard to obtain.

Let $\log^3$ denote the function $g(n) = (\log(n))^3$. Our second main result is the following:

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

We describe a general finite group $G$ by choosing a decomposition series $1 = G_0 \triangleleft G_1 \triangleleft \ldots \triangleleft G_r = G$, where $r \leq \log |G|$. We use Theorem 1.2 to describe the factors $G_{i+1}/G_i$ of the series, which are simple by definition. We then use the method of straight line programs due to [2] to obtain short formulas describing the action of $G_{i+1}$ on $G_i$ via conjugation.

By counting the number of non-isomorphic groups of a certain size, 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.

A $\Sigma_k$-sentence of $L$ is a sentence that is in prenex normal form, starts with an existential quantifier, and has $k - 1$ quantifier alternations. We say that $C$ is $g$-compressible using $\Sigma_k$-sentences if $\varphi$ in Definition 1.1 can be chosen in $\Sigma_k$ form. We will also provide variants of the results above where the sentences are $\Sigma_k$ for a certain $k$. The describing sentences will be of length poly-logarithmic in the size of $G$.

Usually we view a formula $\varphi$ 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 $\varphi$ always introduces new variables with the least index that is available, so that $x_i$ occurs in $\varphi$ only when $i < |\varphi|$). We then encode the resulting string by a binary string, which we call the binary code for $\varphi$. Its length is called the binary length of $\varphi$, 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 logic: the length of the “brute force” descriptions given above, involving the open
diagram, is close to optimal in the general case. To see this, note that there are \(2^{n(n-1)/2}\) 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 \(2^{n(n-1)/2}/n!\), which for large \(n\) exceeds \(\frac{1}{n} \prod_{i=1}^{n-1} 2^i\). For each \(k\) there are at most \(2^k\) sentences \(\varphi\) with length of binary code < \(k\). 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 \(\varphi\) of \(G\). (See [4], Cor. 2.12 for a recent proof that the lower bound \(2^{n(n-1)/2}/n!\) 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 formulas \(\theta_n(g,x)\) 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 \(\alpha_1 \ldots \alpha_k\) be the binary expansion of \(n\). Let \(\theta_n(g,x)\) be the formula

\[
\exists y_1 \ldots \exists y_k [y_1 = x \land y_k = g \land \bigwedge_{1 \leq i < k} y_{i+1} = y_i \circ y_i \circ x^{\alpha_{i+1}}]
\]

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

We give a sample application of Lemma 2.1 which will also be useful below. By the remark after Prop. 3.2 below, the upper bound on the length of descriptions is close to optimal.

**Proposition 2.2.** The class of cyclic groups \(G\) of prime power order is log-compressible via \(\Sigma_3\)-sentences.

**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 \(h\) generating \(H\) such that \(h^{p^k} = 1\), \(h^{p^{k-1}} \neq 1\). By Lemma 2.1, this can be expressed by a \(\Sigma_3\)-sentence of length \(O(\log n)\). \(\square\)

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 length at most \(|G|\) of elements of \(S\).
We next define formulas $\alpha_k(g, x_1, \ldots, x_k)$ in the first-order language of groups $L(e, o, ^{-1})$ expressing that $g$ is in $\langle x_1, \ldots, x_k \rangle$. These formulas depend on $k$ and the size of the group $G$; their length is logarithmic in $k + |G|$. 

**Lemma 2.4.** Let $G$ be a finite group. For each positive integer $k$, there exists a first-order formula $\alpha_k(g; x_1, \ldots, x_k)$ of length $O(k + \log |G|)$ such that

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

*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

$$\delta_0(g; x_1, \ldots, x_k) \equiv \bigvee_{1 \leq j \leq k} g = x_j \lor g = x_j^{-1} \lor g = 1.$$  

For $i > 0$ let

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

Note that $\delta_i$ 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 and their inverses.

Now let $\alpha_k(g; x_1, \ldots, x_k) \equiv \delta_p(g; x_1, \ldots, x_k)$ where $p = 1 + \log |G|$. Then $2^p \geq |G|$, so $\alpha_k$ is a formula as required by Lemma 2.3. \hfill $\Box$

**Remark 2.5.** We note that we can optimize the formulas in Lemmas 2.1 and Lemma 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 $\alpha_k$ becomes a formula over an alphabet of size $k + O(1)$.

The formulas $\alpha_k$ in Lemma 2.4 have about $2 \log k$ quantifier alternations for $k > 0$. We can obtain existential formulas if we allow them to be somewhat longer compared to the size of the group.

**Lemma 2.6.** Let $G$ be a finite group. For each positive integer $k$, there is an existential first-order formula $\beta_k(g; x_1, \ldots, x_k)$ of length $O(k \log^2 |G| + \log^3 |G|)$ such that

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

*Proof.* Let $r = \log |G|$. We use the Reachability Lemma in Babai and Szemerédi [2]: for each set $S = \{x_1, \ldots, x_k\}$ and each $g \in \langle S \rangle$, there is a so-called straight line program $\mathcal{L}$ of reduced length at most $(1 + r)^2$ that “computes” $g$ from $S$. Such a program is a sequence that contains $g$. Each term is either in $S$, or a product or inverse of earlier terms. The sequence has at most $(1 + r)^2$ entries outside $S$. Since the elements of $S$ are not counted, they use the term “reduced length”.


The formula $\beta_k$ formalises the existence of such a program:

$$\beta_k(g; x_1, \ldots, x_k) \equiv \exists w_0, \ldots, w_{r^2} \wedge$$

$$\bigvee_{\ell \leq r^2} w_i = x_{\ell} \lor$$

$$\bigvee_{p, q < i} w_i = w_p w_q \lor$$

$$\bigvee_{p < i} w_i = w_p^{-1}$$

Clearly $|\beta_k| = O(k r^2 + r^4)$.

3. Describing finite fields and finite difference fields

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

A difference field $(F, \sigma)$ is a field $F$ together with a distinguished automorphism $\sigma$. Examples are the field of complex numbers with complex conjugation or a finite field of characteristic $p$ with a fixed power of the Frobenius automorphism. We show that finite fields and finite difference fields are log-compressible using the language of rings $L(\times, 0, 1)$. 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.

Lemma 3.1.

(i) For any finite field $F_q$, there is a $\Sigma_3$-sentence $\varphi_q$ in $L(\times, 0, 1)$ of length $O(\log q)$ describing $F_q$.

(ii) For any finite difference field $(F_q, \sigma)$ there is a $\Sigma_3$-sentence $\psi_{q, \sigma}$ in $L(\times, 0, 1, \sigma)$ of length $O(\log q)$ describing $(F_q, \sigma)$.

Proof. (i). The sentence $\varphi_q$ 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 $\varphi_q$ is $\Sigma_3$.

(ii). Since any automorphism of $F_q$ is of the form $(Frob_p)^k$ 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 $x$.

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

Proposition 3.2. There is a constant $k > 0$ such that for infinitely many primes $q$, for any description $\varphi$ for $F_q$, we have

$$\log(q) \leq k|\varphi| \log |\varphi|.$$
for some $k'$, where the correctional factors are needed because the string $\varphi$ 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 $\log(q/\ln q) \leq +\log n$. Note that $\sqrt{q} \leq q/\ln q$ for $q \geq 3$ so that $\log(q/\ln q) \leq +\log n$.

Choosing $k \geq k'$ appropriately and putting the inequalities together, we obtain $\log q \leq k|\varphi| \log |\varphi|$ as required.

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

4. Describing finite simple groups

**Theorem 1.2.** The class of finite simple groups is $O(\log)$ compressible.

For the proof, 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 $L_n(F_q)$ of fixed Lie type $L$ and Lie rank $n$, possibly twisted, over a finite field $F_q$;
4. the 26 sporadic simple groups.

See e.g. [10], Section 1.2.

4.1. Short first-order descriptions via short presentations. Clearly for the proof of Theorem 1.2 we may ignore the finite set of sporadic simple groups. For most of the other classes we will be using the existence of short presentations:

**Definition 4.1.** We define the length of a presentation $G = \langle x_1, \ldots, x_k \mid r_1, \ldots, r_m \rangle$ as $k + \sum_i |r_i|$, where $|r_i|$ denotes the length of the element $r_i$ expressed as a word in the generators $x_j$ and their inverses.

**Lemma 4.2.** Suppose that a finite simple group $G$ has a presentation $\langle x_1, \ldots, x_k \mid r_1, \ldots, r_m \rangle$.

(i) There is a sentence $\varphi$ describing $G$ of length $O(\log |G| + k + \sum_i |r_i|)$.

(ii) There is a $\Sigma_3$-sentence $\varphi$ describing $G$ of length $O(k \cdot \log^2 |G| + \log^4 |G| + \sum_i |r_i|)$.

**Proof.** (i). The sentence is $\exists x_1, \ldots, x_k [x_1 \neq 1 \land \Lambda_{1 \leq i \leq m} r_i = 1 \land \forall g \alpha_k(g, x_1, \ldots, x_k)$, where $\alpha_k$ is the sentence of length $O(k+\log |G|)$ from Lemma 2.4, expressing that $g$ is generated from the $x_i$ within $G$. This sentence describes $G$ because $G$ is simple.

(ii) is similar using the formula $\beta_k$ from Lemma 2.6 instead of $\alpha_k$. □

For most classes of finite simple groups, Guralnick et al. [7] obtained a presentation for each member $G$ that is very short compared to $|G|$. Subsection 4.2 contains some detail on Ree groups.
Theorem 4.3. [7, Thm. A] There exists a constant $C_0$ such that any non-abelian finite simple group, with the possible exception of a Ree group of type $2G_2$, has a presentation with at most $C_0$ generators and relations and total 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 4.4. The class of finite simple groups, excluding the Ree groups of type $2G_2$, is log-compressible. The same class is log$^4$-compressible using $\Sigma_3$-sentences.

Proof. For simple cyclic 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 $\log n + \log q = O(\log |G|)$. 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 [10] show that $|G|$ is at least $q^n$.

Now the result follows from the foregoing theorem using Lemma 4.2 (i) and (ii), respectively. □

4.2. Short first-order descriptions via interpretations. It remains to treat the class of Ree groups of type $2G_2$, which are called twisted. While the (untwisted) Chevalley groups of type $G_2$ exist over any field $\mathbb{F}$ as the automorphism group of the octonian algebra over $\mathbb{F}$, the (twisted) groups $2G_2$ exist only over fields of characteristic 3 which have an automorphism $\sigma$ 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 seen 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 $\sigma$ on the entries of the matrix (see [6, Section 13.4]).

Proposition 4.5. The class of Ree groups of type $2G_2$ is log-compressible via $\Sigma_d$-sentences for some constant $d$.

No short presentations are known for these Ree groups. Instead, we use first-order interpretations between groups and finite fields in order to derive the proposition from Lemma 3.1.

Suppose that $L, K$ are languages in a finite signature. 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 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 quotient fields, which can be interpreted in the given integral 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 $n^2$, $D$ 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 $\Delta$. It decodes $G$ from $F$ using first-order formulas, so that $G = \Delta(F)$ is an $L$-structure.

**Definition 4.6.** Suppose that $L, K$ are languages in a finite signature, and that classes $C \subseteq M(L), D \subseteq M(K)$ are given. We say that a function $\Delta$ as above is a uniform interpretation of $C$ in $D$ if for each $G \in C$, there is $F \in D$ such that $G = \Delta(F)$.

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

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

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

**Proposition 4.7.** Suppose that $L, K$ are languages in a finite signature, and that classes $C \subseteq M(L), D \subseteq M(K)$ are given. Suppose furthermore that

1. there is a uniform interpretation $\Delta$ without parameters of $C$ in $D$,
2. there is a uniform interpretation $\Gamma$ with parameters of $D$ in $C$, and
3. there is an $L$-formula $\eta$ involving parameters such that for each $G \in C$ there is a list of parameters $\overline{p}$ in $G$ so that $\eta$ defines an isomorphism of $G$ with $\Delta(\Gamma(G, \overline{p}))$.

If $D$ is log-compressible, then so is $C$.

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

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

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

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 4.6, we have $|\psi| = O(\log |G|)$. \qed

Note that if $\varphi$ is a $\Sigma_k$ sentence, then $\psi$ is a $\Sigma_{k+c}$ sentence for a constant $c$ depending only on the interpretations and the formula $\eta$. Thus, if $D$ is
log-compressible using $\Sigma_k$ sentences, then $\mathcal{C}$ is log-compressible using $\Sigma_{k+c}$ sentences.

The previous proposition allows us to deal with the class of twisted Ree groups using a result of Ryten. Note that the class of difference fields $(\mathbb{F}_{3^{2k+1}}, \text{Frob}_{3^{k+1}}^3)$, $k \in \mathbb{N}$, is denoted $\mathcal{C}_{(1,2,3)}$ there. The following is a special case of the more general result of Ryten.

**Theorem 4.8.** (by [9], 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 $(\mathbb{F}_{3^{2k+1}}, \text{Frob}_{3^{k+1}}^3)$. The hypotheses of Prop. 4.7 can be satisfied via uniform interpretations $\Delta, \Gamma$ and a formula $\eta$ in the language of groups.

The details of the proof are contained in Ch.5 of [9]. Since they require quite a bit of background on groups of Lie type, we merely indicate how to obtain the required formulas. The group $2G_2(\mathbb{F})$ has Lie rank 1, and so behaves similarly to the group $SL_2(\mathbb{F})$, which is also of Lie rank 1. The formulas required for Prop. 4.7 are essentially the same in both cases. Since most readers will be more familiar with $SL_2$, we use this group rather than $2G_2$ to make the required subgroups more explicit.

The uniform interpretation $\Delta$ 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 $G_2$ - and hence its subgroup $2G_2$ - 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 $2G_2$ of $G_2$ are then uniformly defined in the language of difference fields by expressing that its elements induce linear transformations (of the affine group $G_2$) that commute with the field automorphism $\sigma$.

The uniform interpretation with parameters $\Gamma$ 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 $SL_2(\mathbb{F})$) and 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 $U_+ \times T$ 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 $\eta: H \cong \Delta(\Gamma(H, \eta))$ 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 $\eta$.

**Proof of Proposition 4.5.** By Theorem 4.8 the class $\mathcal{C}$ of Ree groups of type $^2G_2(q)$ is uniformly parameter interpretable in the class $\mathcal{D}$ of finite difference fields $(F_{3^{2k+1}}, \text{Frob}_3^k)$. By Lemma 3.1, the class $\mathcal{D}$ is log-compressible using $\Sigma_3$ sentences. By Proposition 4.7 and the remark after its proof, this implies that the class $\mathcal{C}$ is log-compressible via $\Sigma_d$ sentences for some constant $d$. (We estimate that $d \leq 10$.)

**Remark 4.9.** In fact, Ryten proves that for fixed Lie type $L$ and rank $n$, the class of finite simple groups $L_n$ 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 Prop. 4.7 there is a formula $\delta$ in the first-order language for $K$ that defines for each $F \in \mathcal{D}$ an isomorphism between $F$ and $\Gamma(\Delta(F), \pi)$. Via Proposition 4.7 this yields a proof that each class of finite simple groups is log-compressible. However, since there are infinitely many such classes it would need some further effort in order to show that there is a uniform constant which works for all classes. We circumvented the problem by using the results of [7] instead.

**Remark 4.10.** By the remark after Lemma 2.4 and the proofs above, for each finite simple group $G$, we can in fact obtain a description of binary length $O(\log(|G|))$. Note that for a given $\psi$ of binary length $O(|G|)$, one can decide whether $G \models \psi$ in deterministic work space logarithmic in $|G|$. Thus, given a simple group $G$ we can find a shortest description within logarithmic space.

Suppose a finite group $G$ is given as an input string of length $O(|G|^2 \log |G|)$ using the multiplication table. It is easy to see that simplicity of $G$ can be decided in polynomial time. We can strengthen this to deterministic logarithmic space. The algorithm tries out all the sentences that are possible descriptions of a finite group and have length $O(\log |G|)$. If one of them is satisfied the algorithm outputs "yes", otherwise "no".

## 5. Describing general finite groups

In this section we prove our second main result, Theorem 1.3, which states that the class of finite groups is $\log^3$-compressible.

Recall straight line programs and the Reachability Lemma from the proof of Lemma 2.6. We will apply a stronger version of the lemma, which is the version Babai and Szemeredi [2] actually proved. We briefly review this following [1].

For a finite group $G$ and $A, T \subseteq G$, a straight line program (SLP) computes $A$ from $T$ if it computes each member of $A$ from $T$. Let $\text{cost}(A \mid T)$ be the shortest reduced length of a straight line program computing $A$ from $T$. Informally, the full Reachability Lemma states that given a generating set
T, we only once need a straight line program of length $\log^2(|G|)$. This long program is for pre-processing, which yields a “better” set of generators $A$: every $g$ in $G$ can be obtained within cost $2\log(|G|)$ from $A$. The fact that $A$ has to be computed only once will save on the length of formulas below.

**Lemma 5.1** ([2, 1]). Let $G$ be a finite group and $T$ a set of generators of $G$. Then there is $A \subseteq G$ with cost($A \mid T$) $< \log^2(|G|)$, and cost($g \mid A$) $< 2\log |G|$ for any $g \in G$.

**Remark 5.2.**
(1) It is well known that any finite simple group has a generating set of size at most 2; see e.g. [5].
(2) For a finite group $G$, any properly ascending chain of subgroups

$$G_0 = 1 < G_1 < \ldots < G_{r-1} < G_r = G$$

has length bounded by $\log |G|$ since we have $|G_i : G_{i-1}| \geq 2$. In particular, any finite group $G$ has a generating set of size at most $\log |G|$. The group $(\mathbb{Z}/2\mathbb{Z})^r$ shows that this bound is optimal.

We write $M \models \varphi$ if $M$ is the only model of $\varphi$ up to isomorphism.

**Lemma 5.3.** Let $G$ be a finite group, so that $G$ has a composition series

$$1 = G_0 \triangleleft G_1 \triangleleft \ldots \triangleleft G_r = G$$

with simple factors $H_i := G_i/G_{i-1}$, $i = 1, \ldots, r$. Suppose that $H_i \models \varphi_i$ for each $i \leq r$. Then $G \models \varphi$ for some sentence $\varphi$ of length

$$O((\log |G|) \sum_i |\varphi_i| + \log^2 |G|).$$

**Proof.** We first introduce some notation. Consider an (ordered) generating set $T = \{t_1, \ldots, t_n\}$ for a finite group $H$. Recall that by Lemma 2.4, the formula $\alpha_n(g; t_1, \ldots, t_n)$ of length $O(n + \log |H|)$ expresses that $g \in \langle T \rangle$. We write

$$\alpha_n(H; T) = \{g \in H : H \models \alpha_n(g; t_1, \ldots, t_n)\},$$

which is a subgroup of $H$.

By Remark 5.2 there is an ascending sequence of generating sets $T_i$ for $G_i$, $0 < i \leq r \leq \log |G|$, with $T_i \setminus T_{i-1}$ of size at most 2. Then $G$ is up to isomorphism the only group such that there exists such a sequence, with $T_r$ generating $G$, and such that the conditions obtained in (a) and (b) below hold.

$$(a) \bigwedge_{0 < i < r} \alpha_{n_i}(G; T_i)/\alpha_{n_{i-1}}(G; T_{i-1}) \models \varphi_i.$$  

The statement inside the $\bigwedge$ can be expressed by a formula $\psi_i$ involving free variables corresponding to the elements of $T_r$. To obtain $\psi_i$, firstly, we restrict the quantifiers in $\varphi_i$ to $\alpha_{n_i}(G; T_i)$. Then we replace each occurrence of “$t = s$” in $\varphi_i$ by “$ts^{-1} \in \alpha_{n_{i-1}}(G; T_{i-1})$”. Since we replace each symbol in $\varphi_i$ by a string of length $O(\log |G|)$, we have $|\psi_i| = O(|\varphi_i| \log |G|)$.

(b) For each $i$ we describe the group structure of $G_{i+1}$ in terms of $T_{i+1}$. $G_0$ is trivial. If we have already described $G_i$ in terms of $T_i$, it suffices to determine the result of multiplication of elements of the form $tg, sh$ for $t, s \in$...
$T_{i+1} \setminus T_i, g, h \in G_i$. This amounts to describing the action of $s \in T_{i+1} \setminus T_i$ on $G_i$ by conjugation, because $(tg)(sh) = ts(s^{-1}gs)h$.

Since $T_i$ generates $G_i$, it suffices in fact for each $i \leq r$ and each $u \in T_i$, $s \in T_{i+1} \setminus T_i$, to determine $s^{-1}us$ as is a certain element $h_{s,u}$ obtained as a term in the generators in $T_i$. However, if for each $s, u$ we picked a SLP of length $< \log^2(|G|)$ computing $h_{s,u}$ from $T_i$, we would only obtain a length of $O(\log^4 |G|)$ in the final tally. Instead, we pick a (pre-processing) set $A_{i-1}$ for $G_i$ and $T_i$ according to Lemma 5.1: a SLP $\mathcal{H}_i$ of reduced length $< \log^2(|G|)$ computes $A_i$. Next for each $s, u$ as above we pick a straight line program $\mathcal{L}_{s,u}$ of reduced length $< 2 \log |G|$ computing $h_{s,u}$ from $A_{i-1}$.

The formula for condition (b) is the following. We work inside an array of existential quantifiers of length $O(\log^2 |G|)$ referring to the elements of the sets $T_i$, which is inherited from the formula in condition (a).

The formula is a conjunction over indices $i$ with $0 \leq i < r$. A conjunction indexed $i$ starts with a string of at most $\log^2 |G|$ existential quantifiers referring to the elements of $A_i$.

- We first express that $\mathcal{H}_i$ computes $A_i$ from $T_i$. This takes a length of $O(\log^2(|G|))$ since the number of new existential quantifiers needed is the reduced length of $\mathcal{H}_i$.
- Next for each $u \in T_i$, $s \in T_{i+1} \setminus T_i$ we express that $\mathcal{L}_{s,u}$ computes $h_{s,u}$ from $A_i$. This takes a length of $O(\log(|G|))$ for a similar reason.

In total we have $O(\log^2(|G|))$ as well for this part.

By Remark 5.2 we have $r \leq \log |G|$. Thus the conjunction obtained in (b) has length $O(\log^3(|G|))$ as required. $\square$

We are now in a position to prove Theorem 1.3:

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

Proof. Let $G$ be a finite group with composition series

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

and simple factors

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

By Theorem 1.2, for each $H_i$ there is a sentence $\varphi_i$ with $|\varphi_i| = O(\log |H_i|)$ describing $H_i$. The result now follows from Lemma 5.3. $\square$

The exponent 3 is optimal even for $p$-groups by the Higman-Sims formula, which states that for some $\epsilon > 0$, for large enough $n$ there are at least $p^{\epsilon n^3}$ non-isomorphic groups of order $p^n$. The argument is similar to [1, Prop. 8.6].

Finally we consider the length of descriptions with a bounded number of quantifier alternations.

Theorem 5.4. For some $p$, the class of finite groups is $\log^9$-compressible via $\Sigma_p$ sentences.

Proof. We only note the necessary modifications to the previous arguments. In Lemma 5.3, for some small $d$ we can choose $\Sigma_d$-descriptions $\varphi_i$ of length $O(\log^4 |G|)$ via Propositions 4.4 and 4.5. Instead of the $\alpha_k$ we use the existential generation formulas $\beta_k$ from Lemma 2.6. They have length $O(\log^4 |G|)$,
so we obtain that $|\psi_i| = O(|\varphi_i| \log^4|G|) = O(\log^8|G|)$. We conclude the argument as in the proof of Theorem 1.3. It is clear that the number of quantifier alternations remains bounded. □

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(1):79–112, 1997.

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

[3] 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(6):3277–3285, 2011.

[4] H. Buhrman, M. Li, J. Tromp, and P. Vitányi. Kolmogorov random graphs and the incompressibility method. SIAM Journal on Computing, 29(2):590–599, 1999.

[5] T. Burness, M. Liebeck, and A. Shalev. Generation and random generation: from simple groups to maximal subgroups. Advances in Mathematics, 248:59–95, 2013.

[6] R. Carter. Simple groups of Lie type, volume 22. John Wiley & Sons, 1989.

[7] R. M. Guralnick, W. M. Kantor, M. Kassabov, and A. Lubotzky. Presentations of finite simple groups: a quantitative approach. J. Amer. Math. Soc., 34:711–774, 2008.

[8] W. Hodges. Model Theory. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1993.

[9] M. J. Ryten. Model theory of finite difference fields and simple groups. PhD thesis, University of Leeds, 2007. Available at https://www1.maths.leeds.ac.uk/pure/staff/macpherson/ryten1.pdf.

[10] R. A. Wilson. The finite simple groups, volume 251 of Graduate Texts in Mathematics. Springer-Verlag London, Ltd., London, 2009.