LENGTH OF POLYNOMIALS OVER FINITE GROUPS

GÁBOR HORVÁTH AND CHRYSTOPHER L. NEHANIV

Abstract. We study the length of polynomials over finite simple non-Abelian groups needed to realize Boolean functions. We apply the results for bounding the length of 5-permutation branching programs recognizing a Boolean set. Moreover, for Boolean and general functions on these groups, we present upper bounds on the length of shortest polynomials computing an arbitrary n-ary Boolean or general function, or a function given by another polynomial.

1. Introduction

Computational models are based on functionally complete algebras, that is, algebras over which every function can be built up from variables, constants and the basic operations of the algebra. The most well-known functionally complete algebra is the two-element Boolean algebra, which is used as a basis for contemporary computers. Nevertheless, other functionally complete algebras exist. Maurer and Rhodes [14] proved that a finite group is functionally complete if and only if it is simple and non-Abelian. Then Krohn, Maurer and Rhodes [11] proved that any Boolean function can be realized by a finite state sequential machine based on a finite simple non-Abelian group. At the end of their paper they suggest to write some forthcoming paper on the algorithmic aspects of such realizations which, unfortunately, never came to exist. The present paper was motivated by trying to fill some of the gaps left by them by estimating the length of a polynomial realizing a given function over a given finite simple non-Abelian group.

The length of polynomials needed to realize a given (Boolean or more general) function has been investigated for several different algebras. Not surprisingly, most of these results concern the two-element Boolean

2010 *Mathematics Subject Classification.* 20F10, 20D06, 20D15, 68Q05, 68Q70.

Key words and phrases. length of polynomial functions, simple non-Abelian groups, nilpotent groups, branching program, permutation branching program.

The first author was partially supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, by the European Union and the European Social Fund through project Supercomputer, the national virtual lab (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0010), and by the Hungarian Scientific Research Fund (OTKA) grant no. K109185. The research by both authors leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318202.

algebra (see e.g. [21]). There are some sporadic results for rings, e.g. short representing polynomials were given for the squareroot function in [1]. There exist some results on the length of unary polynomials over finite groups as well [16], but no estimates can be found in the literature for the n-ary case. Just recently, some particular polynomials for certain special functions were computed in [19] for certain functionally complete algebras. The authors of that paper used a computational search method (genetic programming) to search for discriminator polynomials, Mal'cev polynomials and majority polynomials for particular three- and four-element functionally complete algebras. It turns out that even for such small algebras it is quite difficult to find these polynomials. For example, the exhaustive search to compute a short discriminator polynomial over a particular four-element functionally complete algebra would take about 10^{38} years by their estimation. After a week of running time, their genetic programming method was not able to provide a discriminator polynomial for the algebra either (see [19] for further details).

In our paper we consider two types of functions over a finite simple non-Abelian group $\mathbf{G}$ in Section 3. A Boolean function can easily be represented over $\mathbf{G}$ by a function $f\colon \{1,g\}^n \to \{1,g\}$ for some nontrivial $g \in \mathbf{G}$, where 1 corresponds to *false* and g to *true*. In Theorem 7 we provide an upper bound for a shortest polynomial realizing an arbitrary such f. The proof is based on a recent result of Wilson [22], which uses some parts of the classification of finite simple groups. Then in Theorem 9 we prove an upper bound on the length of an arbitrary function $f\colon \mathbf{G}^n \to \mathbf{G}$. Finally, Theorem 10 gives a lower bound on the length of a 'longest' n-ary function based on an elementary counting argument. This puts the upper bounds obtained in Theorems 7 and 9 into perspective.

Sections 4 and 5 are devoted to two applications of the results of Section 3. In Section 4 we consider branching programs. Branching programs were first defined by Lee [12] as an alternative to Boolean circuits. Since then branching programs have been thoroughly investigated (see e.g. [2, 5, 6, 8, 18] from the past few years). Krohn, Maurer and Rhodes proved in [11] that a finite state sequential machine can compute an arbitrary Boolean function if it is based on a finite simple non-Abelian group. A direct consequence of this result, but which was proven independently by Barrington [4], is that a language can be recognized by an $O(\log n)$ depth, polynomial size Boolean circuit if and only if it can be recognized by a polynomial length branching program over a finite simple non-Abelian group. (Here, by polynomial we mean polynomial in n, which is the arity of the Boolean function.) In fact, Barrington gives an upper bound on the length of the branching program required, depending on the depth of a Boolean circuit which recognizes the particular language. Using the results of Theorem 7, in

Corollary 11 we give a different upper bound on the length of a branching program required, and compare it to Barrington's bound. We find that our bound is better for almost all functions than the one provided by Barrington's construction.

In Section 5 we consider function realization for other finite, but not necessarily non-Abelian or simple, groups. Of course, if a group is not functionally complete, then not every function can be represented as a polynomial. Nevertheless, it would be interesting to know the length of a shortest representing polynomial for a given function that can be represented. In Theorem 12 we show that for groups with nilpotency class d the length of a minimal realizing polynomial for a representable n-ary function $f\colon \mathbf{G}^n \to \mathbf{G}$ is at most $c \cdot n^d$ for some c depending on the group, and this bound is almost the best possible. Corollary 13 is a direct consequence of Theorem 9 which provides a bound for finite simple non-Abelian groups. We suspect that similar upper bounds could be given for arbitrary groups, provided that the length of polynomials over a group $\mathbf{G}$ can be estimated by the length of polynomials over $\mathbf{N}$ and over $\mathbf{G}/\mathbf{N}$ for some normal subgroup $\mathbf{N}$. Finally, we close the paper with some open problems in Section 6.

2. Preliminaries

Let $\mathbf{G}$ be a finite group. A *polynomial* (or *word*) over $\mathbf{G}$ is a product of variables, inverses of variables, and constants from $\mathbf{G}$. For example, $xgy^{-1}x$ is a polynomial over $\mathbf{G}$ for some $g \in \mathbf{G}$ and variables x, y . Let p be a polynomial over $\mathbf{G}$. The *length* of p (denoted by $\|p\|$) is defined recursively:

(1) the length of a variable, of an inverse of a variable or of a constant is 1: $\|x_i\| = \left\|x_i^{-1}\right\| = \|g\| = 1$ $(1 \leq i \leq n,\ g \in \mathbf{G})$;
(2) the length of a product is the sum of the lengths of the factors: $\|p_1p_2\| = \|p_1\| + \|p_2\|$.

For example, the length of the polynomial $xgy^{-1}x$ is

$$\left\|xgy^{-1}x\right\| = \|xg\| + \left\|y^{-1}x\right\| = \|x\| + \|g\| + \left\|y^{-1}\right\| + \|x\| = 4.$$

The *number of variable occurrences* of p (denoted by $v\left(p\right)$) is the number of occurring variables in p, counting multiplicities. The precise definition is the same as for the length, except $v\left(g\right) = 0$ for any $g \in \mathbf{G}$. For example, the number of variable occurrences of $xgy^{-1}x$ is

$$v\left(xgy^{-1}x\right) = v\left(xg\right) + v\left(y^{-1}x\right) = v\left(x\right) + v\left(g\right) + v\left(y^{-1}\right) + v\left(x\right) = 3.$$

A polynomial p *realizes* a function $f\colon \mathbf{G}^n \to \mathbf{G}$ if for all $(a_1, \ldots, a_n) \in \mathbf{G}^n$ we have $f\left(a_1, \ldots, a_n\right) = p\left(a_1, \ldots, a_n\right)$. We say that $f\colon \mathbf{G}^n \to \mathbf{G}$ is a *polynomial function* if f can be realized by a polynomial. The group $\mathbf{G}$ is *functionally complete* if every function $f\colon \mathbf{G}^n \to \mathbf{G}$ can be realized by a polynomial. A finite $\mathbf{G}$ is functionally complete if and only if $\mathbf{G}$ is simple and non-Abelian [14]. The length of a polynomial function f over $\mathbf{G}$ is the length of a shortest polynomial realizing f:

$$\|f\| = \min \{ \|p\| : p \text{ realizes } f \} .$$

Similarly, for the minimal number of variable occurrences:

$$v(f) = \min \{ v(p) : p \text{ realizes } f \} .$$

If f is a non-realizable function over $\mathbf{G}$, then let $\|f\| = v(f) = \infty$.

The first lemma lists some basic observations. It connects the number of necessary variable occurrences with composition of functions and the length of a function with the number of variable occurrences.

Lemma 1. *For polynomial functions $f, g_1, \dots, g_n$ over $\mathbf{G}$ we have*

$$v(f) \leq \|f\| \leq 2v(f) + 1, \tag{1}$$

$$v(f(g_1, \dots, g_n)) \leq v(f) \cdot \max_{1 \leq i \leq n} v(g_i) . \tag{2}$$

Proof. Let p be a polynomial realizing f for which $v(p) = v(f)$. Let p' be the polynomial which we obtain from p by collecting the neighbouring constants into one constant. Then between two variables at most one constant can occur, thus $\|p'\| \leq 2v(p') + 1 = 2v(f) + 1$. As p' realizes f, (1) follows.

For proving (2), let $v_1, \dots, v_n$ be the number of occurrences of the variables $x_1, \dots, x_n$ in p realizing f, where $v(p) = v(f)$. Then

$$\begin{aligned} v(f(g_1, \dots, g_n)) &\leq \sum_{i=1}^{n} v_i v(g_i) \\ &\leq \sum_{i=1}^{n} v_i \cdot \max_{1 \leq i \leq n} v(g_i) = v(f) \cdot \max_{1 \leq i \leq n} v(g_i) . \end{aligned}$$

□

Example 1. Take $\mathbf{G} = A_5$, the alternating group on 5 points, and let $f(x, y) = (123)x(123)(123)y(123)$. Now, $v(f) = 2$, as f depends on both its variables. Therefore it has an at most length 5 realization by multiplying the constants between x and y: $(123)x(132)y(123)$. Furthermore, if $g_1(x, y) = xyx^{-1}y^{-1}$ and $g_2(x, y) = x(234)y^{-1}x$, then $f(g_1(x, y), g_2(x, y))$ can be realized by

$$(123)\ xyx^{-1}y^{-1}\ (132)\ x(234)y^{-1}x\ (123),$$

which has $4 + 3 \leq 2 \cdot 4$ variable occurrences.

In the following lemma we create a 'short' polynomial for an n-ary version of a binary polynomial using logarithmic depth iteration. The idea is similar as how one constructs the n-ary AND function from the binary one. From now on, by log we always mean $\log_2$.

Lemma 2. *Let p be a binary polynomial over $\mathbf{G}$. Define the following polynomials: $p^{(1)}(x_1) = x_1$, $p^{(2)}(x_1, x_2) = p(x_1, x_2)$ and for every integer $n > 2$ let*

$$(3) \quad p^{(n)}(x_1, \dots, x_n) \\ = p\left(p^{(\lfloor n/2 \rfloor)}\left(x_1, \dots, x_{\lfloor n/2 \rfloor}\right), p^{(\lceil n/2 \rceil)}\left(x_{\lfloor n/2 \rfloor + 1}, \dots, x_n\right)\right).$$

Let $V = v(p)$. If $V \geq 2$, then $v\left(p^{(n)}\right) < V \cdot n^{\log V}$. If both x_1 and x_2 each occur exactly twice in p, then $v\left(p^{(n)}\right) \leq \frac{3}{2}n^2 - \frac{3}{2}n + 1$.

Proof. By induction on n (considering the cases where n is odd and where n is even), it is straightforward to prove that $v\left(p^{(n)}\right)$ is increasing in n. If n is a power of 2, then $v\left(p^{(n)}\right) = V^{\log n}$. Thus for arbitrary n we have $v\left(p^{(n)}\right) \leq V^{\lceil \log n \rceil} < V^{1 + \log n} = V \cdot n^{\log V}$. The other inequality can be proved by induction, as well. □

Example 2. Take $p(x_1, x_2)$ to be the commutator of x_1 and x_2, that is $p(x_1, x_2) = [x_1, x_2] = x_1 x_2 x_1^{-1} x_2^{-1}$. Then

$$\begin{aligned} p^{(3)}(x, y, z) &= [x, [y, z]] = xyzy^{-1}z^{-1}x^{-1}zyz^{-1}y^{-1}, \\ p^{(4)}(x, y, z, w) &= [[x, y], [z, w]] = [x, y]\,[z, w]\,[x, y]^{-1}\,[z, w]^{-1} \\ &= xyx^{-1}y^{-1}zwz^{-1}w^{-1}yxy^{-1}x^{-1}wzw^{-1}z^{-1}. \end{aligned}$$

Thus,

$$\begin{aligned} v\left(p^{(2)}\right) &= 4 = 4^{\log 2} = \frac{3}{2} \cdot 2^2 - \frac{3}{2} \cdot 2 + 1, \\ v\left(p^{(3)}\right) &= 10 = \frac{3}{2} \cdot 3^2 - \frac{3}{2} \cdot 3 + 1, \\ v\left(p^{(4)}\right) &= 16 = 4^{\log 4} \leq 19 = \frac{3}{2} \cdot 4^2 - \frac{3}{2} \cdot 4 + 1. \end{aligned}$$

We need some results from group theory. Throughout the paper, the commutator of a and b is $[a, b] = aba^{-1}b^{-1}$, and the conjugate of a by b is $a^b = bab^{-1}$, and multiplication of permutations is carried out from right to left. For general background for group theory we refer to [17].

In the proofs of Theorems 7 and 9 in Section 3 the following recent result of Wilson is crucial.

Theorem 3 ([22, Theorem 1])**.** *Let $\mathbf{G}$ be a finite group. Then the following are equivalent:*

1. *$\mathbf{G}$ is solvable;*
2. *no non-trivial element g is the product of 56 commutators of the form $[g^h, g^k]$ (with $h, k \in \mathbf{G}$).*

That is, for finite simple non-Abelian groups there exist elements g $(\neq 1), h_1, k_1, \dots, h_{56}, k_{56}$ such that $g = \prod_{i=1}^{56} \left[g^{h_i}, g^{k_i}\right]$. This fact combined with Lemma 2 will provide us a short n-ary version of the Boolean AND function.

Finally, in the special case of alternating groups we need the following.

Lemma 4. *Let $u \in \mathbf{A}_m$ (for some $m \geq 5$) be nontrivial and let C_u denote the conjugacy class of u in $\mathbf{A}_m$. Let $D_u = C_u \cup C_{u^{-1}}$ and let D_u^2 denote the set $\{ u_1 \cdot u_2 \mid u_1, u_2 \in D_u \}$. If u is a product of disjoint 2-cycles and moves every point, then D_u^2 contains a product of two disjoint 3-cycles. Otherwise, D_u^2 contains a 3-cycle.*

Proof. It is well-known (see e.g. [17, p. 299, 11.1.5]) that if u is *not* the product of disjoint odd cycles of pairwise different lengths (considering 1-cycles as well) then any $v \in \mathbf{A}_m$ having the same cycle structure as u lies in the same conjugacy class of u. Moreover, if u is the product of odd cycles of pairwise different lengths (considering 1-cycles as well) then the set of elements of $\mathbf{A}_m$ having the same cycle structure as u is the disjoint union of two conjugacy classes.

We choose a cycle of maximal length in u. Let k be its length. Without loss of generality we can assume that this cycle is the $c_k = (1, \ldots, k)$ cycle in u. Note that by [17, p. 299, 11.1.5] if $k \leq 4$, then the conjugacy class C_u contains every element of $\mathbf{A}_m$ with the same cycle-structure as u. We distinguish five cases.

(1) $k \geq 5$. Let $v = c_k^{-1} u$, $v' = v^{-1} = (1\,3)(2\,4) \cdot v^{-1} \cdot (1\,3)(2\,4)$, $c_k' = (2, 1, 4, 3, k, k-1, \ldots, 5) = (1\,3)(2\,4) \cdot c_k^{-1} \cdot (1\,3)(2\,4)$, and let $u' = c_k' \cdot v'$. Then $u' \in C_{u^{-1}} \subseteq D_u$ and (multiplying from right to left)
$$u' \cdot u = c_k' v' \cdot c_k v = c_k' c_k \cdot v' v = c_k' \cdot c_k = (2\,k\,4)\,.$$

(2) $k = 4$. Let $v = c_k^{-1} u$, $c_k' = (1\,2\,4\,3)$ and let $u' = c_k' v^{-1}$. Then $u' \in C_u \subseteq D_u$ (since u and u' have the same cycle-structure and $k \leq 4$) and (multiplying from right to left) $u' \cdot u = (1\,4\,2)$.

(3) $k = 3$. Let $v = c_k^{-1} u$ and let $u' = c_k \cdot v^{-1}$. Now $u' \in C_u \subseteq D_u$ (since u and u' have the same cycle-structure and $k \leq 4$) and (multiplying from right to left) $u' \cdot u = (1\,3\,2)$.

(4) $k = 2$ and u stabilizes an element from $\{ 1, \ldots, m \}$. Without loss of generality we can assume that $u = (1\,2)\,v$ stabilizes 3, then let $u' = (1\,3)\,v^{-1}$. Now $u' \in C_u \subseteq D_u$ (since u and u' have the same cycle-structure and $k \leq 4$) and (multiplying from right to left) $u' \cdot u = (1\,2\,3)$.

(5) $k = 2$ and u moves all the elements from $\{ 1, \ldots, m \}$. Then u is the product of 2-cycles. Without loss of generality we can assume that $u = (1\,2)\,(3\,4)\,(5\,6) \cdot v$. Let $u' = (1\,6)\,(2\,3)\,(4\,5) \cdot v^{-1}$. Then $u' \in C_u \subseteq D_u$ (since u and u' have the same cycle-structure and $k \leq 4$) and (multiplying from right to left) $u' \cdot u = (1\,3\,5) \cdot (2\,6\,4)$.

□

Corollary 5. *Let $u \in \mathbf{A}_m$ (for some $m \geq 5$) be nontrivial and let $t \in \mathbf{A}_m$ be a 3-cycle. Then*

1. *t can be generated as a product of at most 4 conjugates of u and u^{-1},*
2. *u is the product of at most $\lfloor m/2 \rfloor$ conjugates of t and t^{-1}.*

Proof. (1) follows easily from Lemma 4 if u is not the product of disjoint 2-cycles moving every point. Otherwise one can obtain some w, a product of two disjoint 3-cycles, as the product of two conjugates of u and u^{-1}. Then applying Lemma 4 to w provides the result. For (2) see e.g. [3, Chapter 3]. □

Finally, we will need the following:

Theorem 6 ([13, Theorem 1.1])**.** *There exists a positive c_0 such that the following holds: for all finite simple non-Abelian groups $\mathbf{G}$, for every subset $S \subseteq \mathbf{G}$, $S \not\subseteq \{1\}$ closed under conjugation, and for every $m \geq c_0 \log |\mathbf{G}| / \log |S|$ we have $S^m = \mathbf{G}$.*

3. Length of functions over finite simple groups

First, we provide an upper bound on the length of polynomials realizing Boolean-type functions. Let $\exp \mathbf{G}$ denote the exponent of $\mathbf{G}$, i.e. the least $n > 0$ for which g^n is the identity for all $g \in \mathbf{G}$.

Theorem 7. *Let $\mathbf{G}$ be a functionally complete group. Then there exists $g\ (\neq 1) \in \mathbf{G}$ such that for every n-ary function $f\colon \{1, g\}^n \to \{1, g\}$ over $\mathbf{G}$, we have*

$$\|f\| \leq 448 \cdot n^8 \cdot e,$$

where $e = |f^{-1}(g)|$, ($e \leq 2^n$).
If $\mathbf{G} = \mathbf{A}_m$ for some $m \geq 5$, then

$$\|f\| \leq \left(3n^2 - 3n + 2\right) \cdot e + 1.$$

Remark 8. Note, that a Boolean function in disjunctive normal form has essentially length $n \cdot e$.

Proof of Theorem 7. We apply Wilson's result: by Theorem 3 there exist elements $g\ (\neq 1), h_1, k_1, \ldots, h_{56}, k_{56} \in \mathbf{G}$ such that $g = \prod_{i=1}^{56} \left[g^{h_i}, g^{k_i}\right]$. Let $p(x_1, x_2) = \prod_{i=1}^{56} \left[x_1^{h_i}, x_2^{k_i}\right]$ and for every $n \geq 3$ let $p^{(n)}$ be the polynomial defined by (3) of Lemma 2. Note that $p^{(n)}(g, \ldots, g) = g$, and $p^{(n)}$ attains 1 if any of the variables is 1. Now, we have $v\left(p^{(n)}\right) < v(p) \cdot n^{\log v(p)} < 224 \cdot n^8$ by Lemma 2.

Let $f\colon \{1, g\}^n \to \{1, g\}$ be arbitrary taking non-identity values e times. Let $A = \{(a_1, \ldots, a_n) \in \{1, g\}^n : f(a_1, \ldots, a_n) = g\}$, then $|A| = e$. Let $q_1(x) = x^{-1}g$ and $q_g(x) = x$. For every $(a_1, \ldots, a_n) \in A$ let $p_{a_1,\ldots,a_n}(x_1, \ldots, x_n) = p^{(n)}(q_{a_1}(x_1), \ldots, q_{a_n}(x_n))$, then $v(p_{a_1,\ldots,a_n}) \leq$

$v\left(p^{(n)}\right)$ by (2) of Lemma 1. Now, $\prod_{(a_1,\dots,a_n)\in A} p_{a_1,\dots,a_n}$ realizes f, hence applying (1) of Lemma 1 we obtain

$$\|f\| \leq 1 + 2 \cdot \sum_{(a_1,\dots,a_n)\in A} \left(224n^8 - 1\right) \leq 448 \cdot n^8 \cdot e.$$

If $\mathbf{G} = \mathbf{A}_m$ ($m \geq 5$), then by choosing $g = (12345)$, $h = (24)(35)$, $k = (235)$, we have

$$g = (12345) = [(13542), (14523)] = \left[g^k, g^h\right].$$

Now, we choose $p(x_1, x_2) = \left[x_1^k, x_2^h\right] = kx_1k^{-1}hx_2h^{-1}kx_1^{-1}k^{-1}hx_2^{-1}h^{-1}$. Then both x_1 and x_2 occur twice in p. As before, applying Lemmas 2 and 1 finishes the proof. □

One wonders if a similar bound could be obtained by not using Wilson's result but only elementary methods. It is not too hard to prove a bound of $O(n^c \cdot e)$, where c is a constant *depending on the group.* Furthermore, bounding c by a universal constant is equivalent to finding some constant in (2) of Theorem 3 for finite simple non-Abelian groups where the constant 56 appears. Considering that the proof of Theorem 3 in [22] uses Thompson's classification of minimal simple groups [20], an elementary proof to bound c in a universal manner is unlikely.

Theorem 9. *Let $\mathbf{G}$ be a functionally complete group, $N = |\mathbf{G}|$. Let f be an n-ary function over $\mathbf{G}$ taking non-identity values e times ($e \leq N^n$). Then the following inequality holds:*

$$\|f\| \leq 100352 \cdot K^2 \cdot N^8 \cdot n^8 \cdot e,$$

where $K \leq \min(c_0 \log N, \text{number of conjugacy classes of } \mathbf{G})$ with c_0 the universal constant from Theorem 6. If $\mathbf{G} = \mathbf{A}_m$ for some $m \geq 5$, then

$$\|f\| \leq 9 \cdot m \cdot N^2 \cdot n^2 \cdot e.$$

Proof. We begin the same way as in the proof of Theorem 7. By Theorem 3 there exists elements $g\ (\neq 1), h_1, k_1, \dots, h_{56}, k_{56} \in G$ such that $g = \prod_{i=1}^{56} \left[g^{h_i}, g^{k_i}\right]$. Let $p(x_1, x_2) = \prod_{i=1}^{56} \left[x_1^{h_i}, x_2^{k_i}\right]$ and for every $n \geq 3$ let $p^{(n)}$ be the polynomial defined by (3) of Lemma 2. By Lemma 2 we have $v\left(p^{(n)}\right) < v(p) \cdot n^{\log v(p)} < 224 \cdot n^8$, moreover $p^{(n)}(g, \dots, g) = g$, and $p^{(n)}$ attains 1 if any of the variables is 1.

We claim that for every $1 \neq u \in \mathbf{G}$ there exists a unary polynomial $r_u(x)$ such that $r_u(1) = 1$, $r_u(u) = g^{-1}$ and $v(r_u) \leq c_0 \log N$. Indeed, by Theorem 6 there exists a universal constant c_0 (i.e. not depending on $\mathbf{G}$ or on f) such that the conjugacy class of u generates $\mathbf{G}$ in at most $c_0 \log N$ steps. That is, there exist elements $s_1, \dots, s_{K_u}$ (for some $K_u \leq c_0 \log N$) such that $g^{-1} = u^{s_1} \dots u^{s_{K_u}}$. Then the polynomial $r_u(x) = x^{s_1} \dots x^{s_{K_u}}$ has the required properties. Note, that K_u can be chosen to be less than the number of conjugacy classes of $\mathbf{G}$, as well.

(The set $\{ u^{y_1} \dots u^{y_k} : y_1, \dots, y_t \in \mathbf{G} \}$ is closed under conjugation, thus increases by at least one conjugacy class if t increases by 1.)

Similarly, for every $u \in \mathbf{G} \setminus \{ 1 \}$ there exists a unary polynomial $r'_u(x)$ such that $r'_u(1) = 1$, $r'_u(g) = u$ and $v(r'_u) \leq K'_u$, where $K'_u \leq c_0 \log N$ and K'_u can be chosen to be less than the number of conjugacy classes of $\mathbf{G}$, as well. Let $K = \max_{u \in \mathbf{G} \setminus \{1\}} \{ K_u, K'_u \}$. Then K is less than the number of conjugacy classes of $\mathbf{G}$, and $K \leq c_0 \log N$.

Let $u_1, \dots, u_{N-1}$ be the non-identity elements of $\mathbf{G}$. Let

$$\chi(x) = p^{(N-1)} \left(gr_{u_1}(x), \dots, gr_{u_{N-1}}(x) \right),$$

$$\chi_{a_1,\dots,a_n}(x_1, \dots, x_n) = p^{(n)} \left(\chi\left(x_1 a_1^{-1}\right), \dots, \chi\left(x_n a_n^{-1}\right) \right), \ (a_i \in \mathbf{G})$$

$$q(x_1, \dots, x_n) = \prod_{\substack{(a_1,\dots,a_n) \in \mathbf{G}^n \\ 1 \neq u = f(a_1 \dots, a_n)}} r'_u\left(\chi_{a_1,\dots,a_n}(x_1, \dots, x_n)\right).$$

Then χ is the characteristic function of 1, that is $\chi(1) = g$, and χ attains 1 at any other substitution. Similarly, $\chi_{a_1,\dots,a_n}$ is the characteristic function of the tuple $(a_1, \dots, a_n)$, i.e. $\chi_{a_1,\dots,a_n}(a_1, \dots, a_n) = g$ and $\chi_{a_1,\dots,a_n}$ attains 1 on every other n-tuple. Thus $q(a_1, \dots, a_n) = f(a_1, \dots, a_n)$ for every $a_1, \dots, a_n \in \mathbf{G}$. By (2) of Lemma 1 we have

$$\begin{aligned} v(q) &\leq \max_{u \in G \setminus \{1\}} v(r_u) \cdot v\left(p^{(N-1)}\right) \cdot v\left(p^{(n)}\right) \cdot \max_{u \in G \setminus \{1\}} v(r'_u) \cdot e \\ &< 50176 \cdot K^2 \cdot N^8 \cdot n^8 \cdot e. \end{aligned}$$

Applying (1) of Lemma 1 we obtain the desired bound.

If $\mathbf{G} = \mathbf{A}_m$, then we can give better estimates. By choosing $g = (123)$, $h = (243)$, $k = (154)$, we have

$$g = (123) = [(235), (142)] = \left[g^k, g^h\right].$$

Now, we choose $p(x_1, x_2) = \left[x_1^k, x_2^h\right] = k x_1 k^{-1} h x_2 h^{-1} k x_1^{-1} k^{-1} h x_2^{-1} h^{-1}$. Then both x_1 and x_2 occur twice in p. Lemma 2 yields $v\left(p^{(n)}\right) < \frac{3}{2} n^2$. Note that $v(r_u) \leq 4$ by (1) of Corollary 5, and $v(r'_u) \leq \lfloor m/2 \rfloor$ by (2) of Corollary 5. Then by (2) of Lemma 1 we have

$$\begin{aligned} v(q) &\leq \max_{u \in G \setminus \{1\}} v(r_u) \cdot v\left(p^{(N-1)}\right) \cdot v\left(p^{(n)}\right) \cdot \max_{u \in G \setminus \{1\}} v(r'_u) \cdot e \\ &< 9 \lfloor m/2 \rfloor \cdot N^2 \cdot n^2 \cdot e. \end{aligned}$$

Applying (1) of Lemma 1 we obtain the desired bound. □

Finally, to put these upper bounds into context, we give a lower bound on the length of a 'longest' n-ary function.

Theorem 10. *Let* $\mathbf{G}$ *be a functionally complete group and let* $N = |\mathbf{G}|$. *For every* $\varepsilon > 0$ *and for sufficiently large* n *(depending on* ε*) there exists an* n*-ary function* f *over* $\mathbf{G}$*, such that*

$$\|f\| \geq \frac{\log N}{1 + \varepsilon} \cdot \frac{N^n}{\log n}.$$

Proof. We use a simple counting argument. The number of polynomials of length at most l is at most $(2n+N+1)^l$, since at every position of a polynomial there is either a constant, a variable, an inverse of a variable, or nothing at all. Let f be a longest n-ary function, let $L = \|f\|$. As the number of n-ary functions is N^{N^n}, we obtain $N^{N^n} \leq (2n+N+1)^L$, and thus

$$N^n \cdot \log N \leq L \cdot \log(2n+N+1).$$

Let us fix $\varepsilon > 0$. For $n \geq \max\left(3^{1/\varepsilon}, N+1\right)$ we have

$$N^n \cdot \log N \leq L \cdot \log 3n \leq L \cdot (1+\varepsilon) \cdot \log n.$$

□

4. Bounded-width branching programs

An n-input *branching program* of length s over a monoid $\mathbf{M}$ is a sequence $B = \langle i_1, f_1, g_1 \rangle \dots \langle i_s, f_s, g_s \rangle$ with $1 \leq i_j \leq n$ and $f_i, g_i \in \mathbf{M}$. On the input $(a_1, \dots, a_n) \in \{0,1\}^n$ the instruction $\langle i, f, g \rangle$ is evaluated to f if $a_i = 1$ and to g if $a_i = 0$. The program is evaluated as the product of the evaluated instructions. This assigns to a program B a function $B^* \colon \{0,1\}^n \to \mathbf{M}$:

$$B^*(a_1, \dots, a_n) = h_{i_1} \dots h_{i_s}, \text{ where } h_{i_j} = \begin{cases} f_j, & \text{if } a_{i_j} = 1, \\ g_j, & \text{if } a_{i_j} = 0. \end{cases}$$

Let us fix a subset $F \subseteq \mathbf{M}$. We say that a set $A \subseteq \{0,1\}^n$ is recognized by the branching program B if

$$B^*(a_1, \dots, a_n) \in F \iff (a_1, \dots, a_n) \in A.$$

If $\mathbf{M}$ is a permutation group over w elements, then we use the term *permutation branching program of width* w, or shortly w-PBP. We say that a 5-PBP B *five-cycle recognizes* $A \subseteq \{0,1\}^n$ if there exists a five-cycle $g \in \mathbf{S}_5$ such that $B^*(a_1, \dots, a_n) = g$ if $(a_1, \dots, a_n) \in A$ and $B^*(a_1, \dots, a_n) = 1$ if $(a_1, \dots, a_n) \notin A$.

Barrington [4] proved that if a subset $A \subseteq \{0,1\}^n$ can be recognized by a Boolean circuit of depth d, then it can be 5-cycle recognized by a 5-PBP of length at most 4^d. Note that putting together the proof from [4] and Theorem 3, one can have Barrington's result for arbitrary nonsolvable groups with branching program length at most $(4 \cdot 56)^d$. However, we can prove another upper bound (not depending on d but only on n) using Theorem 7:

Corollary 11. *Let $A \subseteq \{0,1\}^n$. Then A is five-cycle recognized by a 5-PBP of length at most $\frac{3}{2}n^2 \cdot \min\{|A|, 2^n - |A|\}$.*

Proof. The proof of Theorem 7 provides a 5-cycle element $g \in \mathbf{A}_5$ and a polynomial q for which $v(q) \leq \frac{3}{2}n^2|A|$ and $q(g^{a_1}, \dots, g^{a_n}) = g$ if $(a_1, \dots, a_n) \in A$ and $q(g^{a_1}, \dots, g^{a_n}) = 1$, otherwise. Let $k = v(q)$ and $q(x_1, \dots, x_n) = c_1 y_1 c_2 y_2 \dots c_k y_k c_{k+1}$, where $c_j \in \mathbf{A}_5$ for $1 \leq j \leq k+1$

and each y_j is either x_i or x_i^{-1} for some $1 \leq i \leq n$. Let B be the following 5-PBP: the jth instruction of B is (for $1 \leq j \leq k-1$)

- $\langle i, c_j g, c_j \rangle$, if $y_j = x_i$;
- $\langle i, c_j g^{-1}, c_j \rangle$, if $y_j = x_i^{-1}$;

and the kth instruction is

- $\langle i, c_k g c_{k+1}, c_k c_{k+1} \rangle$, if $y_k = x_i$;
- $\langle i, c_k g^{-1} c_{k+1}, c_k c_{k+1} \rangle$, if $y_k = x_i^{-1}$.

Then B recognizes the set A.

Let A^c denote the complement of A. If $|A| > 2^n - |A| = |A^c|$, then instead of q we consider the polynomial $g \cdot (q')^{-1}$, where q' is a polynomial for which $v(q') \leq \frac{3}{2} n^2 |A^c|$, $q'(g^{a_1}, \ldots, g^{a_n}) = g$ if $(a_1, \ldots, a_n) \in A^c$ and $q'(ga_1, \ldots, g^{a_n}) = 1$, otherwise. The construction of B is similar as in the other case. □

Almost every n-ary function is recognized by a circuit of depth at least $n - \log\log n$ [21, Theorem 4.1, p. 97]. (A property P holds for *almost all* functions if the ratio of the number of n-ary functions for which P holds to the total number of n-ary functions tend to 1 as $n \to \infty$.) In particular, Barrington's construction [4] provides an upper bound of at least $4^n / \log^2 n$ on the length needed to five-cycle recognize almost every n-ary function. By Corollary 11 any n-ary function can be five-cycle recognized by a 5-PBP of length at most $\frac{3}{4} \cdot n^2 \cdot 2^n$.

5. Length of polynomial functions over finite groups

Finally, we consider the length of polynomial functions over finite groups. In particular, if a function can be represented by a polynomial, then it can be represented by a short polynomial, as well. For example, if $\mathbf{G}$ is a commutative (i.e., Abelian) group and p is an n-ary polynomial over $\mathbf{G}$, then there exists an n-ary polynomial p' such that $p'(a_1, \ldots, a_n) = p(a_1, \ldots, a_n)$ for every $(a_1, \ldots, a_n) \in \mathbf{G}^n$ and $\|p'\| \leq 1 + n \cdot (\exp \mathbf{G} - 1)$. Moreover, one can find p' from p using $O(\|p\|)$ time and $O(n)$ space. A similar result for nilpotent groups can be obtained using commutator calculus [15, Chapter 3]:

Theorem 12. *Let $\mathbf{G}$ be a finite nilpotent group with nilpotency class d. Let p be an n-ary polynomial over $\mathbf{G}$. Then there exists an n-ary polynomial p' such that $p(a_1, \ldots, a_n) = p'(a_1, \ldots, a_n)$ for every $(a_1, \ldots, a_n) \in \mathbf{G}^n$ and*

$$\|p'\| \leq c \cdot n^d,$$

where c depends only on $\mathbf{G}$. Moreover, for every $\varepsilon > 0$ and for sufficiently large n (depending on ε) there exists an n-ary polynomial function f over $\mathbf{G}$ such that

$$\|f\| \geq \frac{1}{d^d (1 + \varepsilon)} \cdot \frac{n^d}{\log n}.$$

Proof. Let $N = |\mathbf{G}|$. We use the definition of the *weight of a basic commutator* from [15, 31.51]: Let $T = \mathbf{G} \cup \{ x_1, \dots, x_n \}$ and assume any (fixed) linear order $\preceq$ on T. Then the elements of T are basic commutators of weight 1. If basic commutators of weight $< k$ are defined and ordered extending $\preceq$, then define basic commutators of weight k as $[c_i, c_j]$, where the sum of weights of c_i and c_j is k. Then we extend the order $\preceq$ by $c_i \prec c_j$ if the weight of c_i is strictly smaller than the weight of c_j, and use any ordering among basic commutators of the same weight.

By [15, 31.52] every n-ary polynomial over $\mathbf{G}$ with nilpotency class d is equivalent to a product of basic commutators of the form

$$\cdots \prod_{s,t,u \in T} [[s,t],u]^{k_{s,t,u}} \prod_{s,t \in T} [s,t]^{k_{s,t}} x_n^{k_n} \dots x_1^{k_1} g,$$

where every basic commutator has weight at most d, and the occurring basic commutators appear in the order of $\succeq$.

We count the number of basic commutators of weight l. First, one chooses the l elements of the basic commutator in at most $(n+N)^l$-many ways. One can put in brackets into each such basic commutator in $\frac{1}{l+1}\binom{2l}{l}$-many ways (this is exactly the Catalan-number, see e.g. [7, Chapter 4]). Then, each basic commutator of weight l can be expanded to a group polynomial of length at most 4^{l-1}. This can be proved by induction on l: for basic commutator expressions p and q of weight t and $l-t$, $[p,q]$ can be expanded to a polynomial of length at most

$$2 \cdot \left(4^{t-1} + 4^{l-t-1}\right) \leq 2 \cdot \left(4^{l-2} + 4^{l-2}\right) = 4^{l-1}.$$

Thus every polynomial has length at most

$$1 + \exp \mathbf{G} \cdot \sum_{l=1}^{d} 4^{l-1} \cdot (n+N)^l \cdot \frac{1}{l+1}\binom{2l}{l} \leq$$

$$\exp \mathbf{G} \cdot \sum_{l=1}^{d} (16n + 16N)^l \leq d \cdot \exp \mathbf{G} \cdot (16n + 16N)^d \leq c \cdot n^d,$$

for $c \leq d \cdot \exp \mathbf{G} \cdot 16^d (N+1)^d$ and for all $n \geq 1$. (For the first inequality we used $\frac{1}{l+1}\binom{2l}{l} \leq 2^{2l} = 4^l$, and applied another factor of 4 to get rid of the additional 1 at the beginning.)

For proving the lower bound, we use a simple counting argument similar to the proof of Theorem 10. The number of polynomials of length at most l is at most $(2n+N+1)^l$, since at every position of a polynomial there is either a constant, a variable, an inverse of a variable, or nothing at all. Let f be a longest n-ary polynomial function, let $L = \|f\|$. As the number of n-ary functions realized by polynomials is more than $2^{\left(\frac{n}{d}\right)^d}$ [9, Section 1.3], we obtain $2^{\left(\frac{n}{d}\right)^d} \leq (2n+N+1)^L$, and thus

$$\frac{1}{d^d} \cdot n^d \leq L \cdot \log (2n + N + 1).$$

Let us fix $\varepsilon > 0$. For $n \geq \max \left(3^{1/\varepsilon}, N + 1\right)$ we have

$$\frac{1}{d^d} \cdot n^d \leq L \cdot \log 3n \leq L \cdot (1 + \varepsilon) \cdot \log n.$$

□

Theorem 9 immediately gives an estimate on the length of polynomials for finite simple non-Abelian groups.

Corollary 13. *Let* $\mathbf{G}$ *be a finite simple non-Abelian group. Let* p *be an* n*-ary polynomial over* $\mathbf{G}$*, and let* e *denote the number of* n*-tuples where* p *attains a non-identity element. Then there exists an* n*-ary polynomial* p' *such that* $p(a_1, \ldots, a_n) = p'(a_1, \ldots, a_n)$ *and*

$$\|p'\| \leq c \cdot n^8 \cdot e,$$

where c *depends only on* $\mathbf{G}$*.*

6. Open problems

A natural problem arises immediately after one defines the length and variable occurrence for a function as a minimum length and variable occurrence of its realizing polynomials. Namely, whether these two minima can attain their value on the same polynomial. We conjecture that it is not always the case, we have no counterexample, though.

Problem 1. Is it true that for every function f there exists p such that $\|f\| = \|p\|$ and $v(f) = v(p)$?

Comparing the results of Theorems 9 and 10, one wonders what the best possible estimate on the length of functions for finite simple non-Abelian groups could be.

Problem 2. Let $\mathbf{G}$ be a finite simple non-Abelian group, and let $f \colon \mathbf{G}^n \to \mathbf{G}$ be an arbitrary function. Determine the length of a shortest polynomial realizing f.

In Section 5 we presented some upper bounds on the length of a polynomial realizing an arbitrary polynomial function. It would be interesting to know whether similar bounds can be applied for arbitrary finite groups.

Problem 3. Let $\mathbf{G}$ be a finite group, and let $f \colon \mathbf{G}^n \to \mathbf{G}$ be a polynomial function. Determine the length of a shortest polynomial realizing f.

In particular, we believe that $f \colon \mathbf{G}^n \to \mathbf{G}$ can be represented by a polynomial built up from polynomials over $\mathbf{N}$ and $\mathbf{G}/\mathbf{N}$ for some normal subgroup $\mathbf{N}$ of $\mathbf{G}$. This has been proven (in a more general setting) for $\mathbf{G} \simeq \mathbf{N} \times \mathbf{K}$, where $(|\mathbf{N}|, |\mathbf{K}|) = 1$ [10, Corollary 2], or when $\mathbf{N}$ is a non-Abelian minimal normal subgroup of $\mathbf{G}$ [10, Corollary 14].

Problem 4. Let $\mathbf{G}$ be a finite group, $\mathbf{N}$ be one of its normal subgroups. Assume that an arbitrary n-ary function over $\mathbf{N}$ has length at most $s(n)$, and an arbitrary function over $\mathbf{G}/\mathbf{N}$ has at most length $t(n)$. Determine the length of a shortest polynomial realizing an arbitrary n-ary function over $\mathbf{G}$.

The algorithmic aspect of finding a short polynomial realizing a polynomial function is interesting, as well.

Problem 5. Let $\mathbf{G}$ be a finite group, and let $p\colon \mathbf{G}^n \to \mathbf{G}$ be a polynomial. Is there a polynomial algorithm in $\|p\|$ to find a shortest polynomial realizing p?

References

[1] S. J. Agou, M. Deléglise, and J.-L. Nicolas. Short polynomial representations for square roots modulo p. *Des. Codes Cryptogr.*, 28(1):33–44, 2003.

[2] M. Ajtai. A non-linear time lower bound for Boolean branching programs. *Theory Comput.*, 1:149–176, 2005.

[3] Z. Arad, J. Stavi, and M. Herzog. *Products of conjugacy classes in groups*, volume 1112 of *Lecture Notes in Mathematics*. Springer, 1985.

[4] D. A. Barrington. Bounded-width polynomial-size branching programs recognize exactly those languages in NC^1. *J. Comput. System Sci.*, 38(1):150–164, 1989. 18th Annual ACM Symposium on Theory of Computing (Berkeley, CA, 1986).

[5] B. Bollig. Property testing and the branching program size of Boolean functions (extended abstract). In *Fundamentals of Computation Theory*, volume 3623 of *Lecture Notes in Comput. Sci.*, pages 258–269. Springer, Berlin, 2005.

[6] M. Braverman, S. Cook, P. McKenzie, R. Santhanam, and D. Wehr. Branching programs for tree evaluation. In *Mathematical foundations of computer science 2009*, volume 5734 of *Lecture Notes in Comput. Sci.*, pages 175–186. Springer, Berlin, 2009.

[7] J. H. Conway and R. K. Guy. *The Book of Numbers.* Copernicus, New York, 1996.

[8] K. A. Hansen. Constant width planar branching programs characterize ACC^0 in quasipolynomial size. In *Twenty-Third Annual IEEE Conference on Computational Complexity*, pages 92–99. IEEE Computer Soc., Los Alamitos, CA, 2008.

[9] G. Higman. The orders of relatively free groups. In *Proc. Internat. Conf. Theory of Groups (Canberra, 1965)*, pages 153–165. Gordon and Breach, New York, 1967.

[10] K. Kaarli and P. Mayr. Polynomial functions on subdirect products. *Monatsh. Math.*, 159(4):341–359, 2010.

[11] K. Krohn, W. D. Maurer, and J. Rhodes. Realizing complex boolean functions with simple groups. *Information and Control*, 9(2):190–195, 1966.

[12] C. Y. Lee. Representation of switching circuits by binary-decision programs. *Bell System Tech. J.*, 38:985–999, 1959.

[13] M. W. Liebeck and A. Shalev. Diameters of finite simple groups: sharp bounds and applications. *Ann. of Math. (2)*, 154(2):383–406, 2001.

[14] W. D. Maurer and J. L. Rhodes. A property of finite simple non-abelian groups. *Proc. Amer. Math. Soci.*, 16:552–554, 1965.

[15] H. Neumann. *Varieties of Groups.* Springer-Verlag, Berlin, 1967.

[16] S. D. Scott. The arithmetic of polynomial maps over a group and the structure of certain permutational polynomial groups. I. *Monatsh. Math.*, 73:250–267, 1969.

[17] W. R. Scott. *Group Theory.* Dover Publications Inc., New York, second edition, 1987.

[18] J. Šíma and S. Žák. A polynomial time constructible hitting set for restricted 1-branching programs of width 3. In *SOFSEM 2007: Theory and practice of computer science*, volume 4362 of *Lecture Notes in Comput. Sci.*, pages 522–531. Springer, Berlin, 2007.

[19] L. Spector, D. M. Clark, I. Lindsay, B. Barr, and J. Klein. Genetic programming for finite algebras. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2008).* ACM Press, 2011. in press.

[20] J. G. Thompson. Nonsolvable finite groups all of whose local subgroups are solvable. *Bull. Amer. Math. Soc.*, 74:383–437, 1968.

[21] I. Wegener. *The Complexity of Boolean Functions.* John Wiley & Sons Ltd, B. G. Teubner, Stuttgart, 1987.

[22] J. S. Wilson. Finite axiomatization of finite soluble groups. *J. London Math. Soc.*, 74(3):566–582, 2006.

Institute of Mathematics, University of Debrecen, Pf. 12, Debrecen, 4010, Hungary
E-mail address: `ghorvath@science.unideb.hu`

Centre for Computer Science & Informatics Research, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB, United Kingdom
E-mail address: `C.L.Nehaniv@herts.ac.uk`