Computing rank of finite algebraic structures with limited nondeterminism

Jeffrey Finkelstein
Computer Science Department, Boston University
July 15, 2015

Abstract
The rank of a finite algebraic structure with a single binary operation is the minimum number of elements needed to express every other element under the closure of the operation. In the case of groups, the previous best algorithm for computing rank used polylogarithmic space. We reduce the best upper bounds on the complexity of computing rank for quasigroups and for groups, and provide a theoretically efficient algorithm for the latter. This paper proves that with short certificates of correctness, the rank problem for these algebraic structures can be verified by highly restricted models of computation.

Specifically, we prove that the problem of deciding whether the rank of a finite quasigroup, given as a Cayley table, is smaller than a specified number is decidable by a circuit of depth $O(\log \log n)$ augmented with $O(\log^2 n)$ nondeterministic bits (the complexity class of problems decidable by such circuits is denoted $\beta_2\text{FOLL}$). Furthermore, if the quasigroup is a group, then the problem is also decidable by a Turing machine using $O(\log n)$ space and $O(\log^2 n)$ bits of nondeterminism (the complexity class for problems like this is denoted $\beta_2\text{L}$). These are great improvements on the upper bounds for these problems, and for groups, this upper bound implies a deterministic polynomial time algorithm. The lens of limited nondeterminism provides an easy way to improve many simple algorithms, like the ones presented here, and we suspect it will be especially useful for other algebraic algorithms.

1 Introduction
An efficient algorithm computing the rank of a finite algebraic structure (that is, the minimum number of elements required to generate all other elements) benefits...
mathematicians, who use numerical algebra systems for research, cryptographers, who rely on algebraic systems for proofs of security, and theoretical computer scientists, who seek to understand which problems can be solved in a particular model of computation. If the structure is, for example, a finite group, then we can represent this structure in one of two reasonable ways. First, we can represent it as a subset of elements along with a set of equality relations demonstrating how the group operation behaves (known as a group presentation). Second, we can represent it as a table of values for the binary operation under each pair of input elements (known as a Cayley table or multiplication table). These representations offer a tradeoff between representation size and the complexity of deciding properties of the group: the latter representation may be exponentially larger than the former, so an efficient algorithm for the latter may not necessarily be efficient for the former.

Consider the situation when the algebraic structure is the finite cyclic group of order \( n \). A natural presentation of this group is \( \langle a \mid a^n = 1 \rangle \). Since each element in this group can be represented by \( O(\log n) \) bits, the total size of this representation is \( O(\log n) \) bits. In contrast, the Cayley table for this group requires \( O(n^2 \log n) \) bits. Thus, in certain cases, if \( m \) represents the size of the input, an algorithm running in time \( f(m) \) on inputs of the first form runs in time \( O(f(\log m)) \) on inputs of the second form. We can use this to our advantage to construct more efficient algorithms for algebraic problems.

For quasigroups, the previous best algorithm for computing the rank requires polynomial time in addition to a polylogarithmic amount of nondeterministic bits. For groups, the previous best algorithm for computing the rank requires a polylogarithmic amount of space, which induces a superpolynomial time (hence, inefficient) algorithm. Only for certain classes of finite groups was there a polynomial time algorithm. We reduce the best upper bound on the complexity of the rank problem for quasigroups and groups and provide a theoretically efficient algorithm for the latter. This paper proves that with short certificates of correctness, the rank problem for quasigroups and groups can be verified by highly restricted models of computation, and demonstrates how the same strategy can be applied to semigroups and magmas in general.

We prove that the problem of deciding whether the rank of a finite quasigroup, given as a Cayley table, is smaller than a specified number is decidable by a circuit of depth \( O(\log \log n) \) augmented with \( O(\log^2 n) \) nondeterministic bits (the complexity class of problems decidable by such circuits is denoted \( \beta_2 \text{FOLL} \)). Furthermore, if the quasigroup is a group, then the problem is also decidable by a Turing machine using \( O(\log n) \) space and \( O(\log^2 n) \) bits of nondeterminism (the complexity class for problems like this is denoted \( \beta_2 \text{L} \)). The general strategy is to reduce the problem of computing rank to the problem of computing membership; we compute the rank of a group by guessing a small set of candidate generators, then deciding whether each other element in the group can be generated from that set. For the sake of completeness, we show how this strategy applies to semigroups and magmas in general, though the results are less interesting there because these algebraic structures lack two things that the groups and quasigroups have: small generating sets and efficient membership algorithms.
Table 1: We improve algorithms for computing rank of finite algebraic structures.

<table>
<thead>
<tr>
<th></th>
<th>old</th>
<th>new</th>
</tr>
</thead>
<tbody>
<tr>
<td>magma</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>semigroup</td>
<td>NP</td>
<td>NL</td>
</tr>
<tr>
<td>quasigroup</td>
<td>$\beta_2P$ [15]</td>
<td>$\beta_2$FOLL</td>
</tr>
<tr>
<td>group</td>
<td>$L^2$ [13]</td>
<td>$\beta_2$FOLL $\cap \beta_2L$</td>
</tr>
<tr>
<td>ring</td>
<td>NP</td>
<td>NL</td>
</tr>
</tbody>
</table>

Figure 1: $\beta_2$FOLL $\cap \beta_2L$ is a much smaller complexity class than both $L^2$ and $P$.

The problem is in $NL$ for semigroups and in $NP$ for general magmas. Finally, we show that the problem for rings is in $NL$ as well. Table 1 summarizes these improvements, and Figure 1 demonstrates graphically why these improvements are so significant. (We could not find an explicit upper bound for magmas and semigroups, but other papers imply that the problems are in $NP$.)

These are improvements on the previous best upper bounds for these problems. Previously, the best upper bound for computing the rank of a quasigroup given as a Cayley table was $\beta_2P$ [15] Section 5] and for groups, $L^2$ [13] (see also [1], Proposition 6] for a brief description of the algorithm). Our results are an improvement because

$$(\beta_2$FOLL $\cap \beta_2L) \subseteq NL \subseteq (L^2 \cap P).$$

This also improves the result of [1, Theorem 7], which shows that computing the rank of a nilpotent group is in $P$. However, the relationship between
FOLL and \(L\) remains unknown (the best inclusion known is the uninteresting inclusion \(FOLL \subseteq AC^1\)), so the relationship between \(\beta_2FOLL\) and \(\beta_2L\) is unknown as well. Still, for groups, the problem is decidable by an extremely restrictive computational model, and so can be simulated by a deterministic Turing machine in polynomial time. Finally, contrast the complexity of the rank problem for groups with the complexity of computing the rank of a subgroup of a free group. The latter problem is \(P\)-complete, so is not even in \(NC\) unless \(NC = P\) [9, Theorem 4.9] (see also [9, Problem A.8.11]).

Using limited nondeterminism and restrictive models of computation as verifiers may also be useful in examining other problems. The limited nondeterminism lens specifically suggests some opportunities for further research in computational algebra, though it has seen some recent success in other subfields of theoretical computer science (see [7], for example). Here are some avenues for future research.

- Is computing the rank of a quasigroup also in \(\beta_2L\)?
- Is the group rank problem in a smaller complexity class, one contained in both \(\beta_2FOLL\) and \(\beta_2L\)? What is the largest complexity class we can find that is in \(FOLL \cap L\)? This would likely improve all the results in [4].
- Is there a reduction between the problem of computing the rank of a quasigroup and the problem of deciding whether two quasigroups are isomorphic?
- Is the problem of computing the shortest generating sequence for a quasigroup strictly more difficult than the problem of computing the rank of a quasigroup?
- The complexity of group problems, for example, varies based on the succinctness of the representation of the input. In this paper we show that the rank problem is quite easy when the input is given its least succinct representation, the full Cayley table. On the other hand, in the most succinct representation, the group presentation (a set of generators for the group along with relations among the generators), many problems become very difficult, or even undecidable if the group is infinite. For representations of intermediate succinctness, for example a circuit that outputs the entries of the Cayley table, how difficult is the rank problem?

2 Preliminaries

Here, \(\log n\) denotes the base two logarithm of \(n\), for any natural number \(n\). Also, the set \(\{1, \ldots, n\}\) is denoted \([n]\).

2.1 Complexity

\(L\) is the class of languages decidable by a deterministic Turing machine that uses \(O(\log n)\) space on inputs of length \(n\). \(L^2\) is the class of languages decidable
by a deterministic Turing machine that uses \(O(\log^2 n)\) space. \(\text{NL}\) is the class of languages decidable by a nondeterministic Turing machine that uses \(O(\log n)\) space. \(\beta_2 \text{L}\) is the subclass of \(\text{NL}\) in which the nondeterministic Turing machine uses at most \(O(\log^2 n)\) nondeterministic bits. \(\text{FOLL}\) is the class of languages decidable by a \(L\)-uniform family of circuits with polynomial size, unbounded fan-in, and \(O(\log \log n)\) depth. \(\beta_2 \text{FOLL}\) is the class of languages decidable by \(\text{FOLL}\) circuits that have been augmented with \(O(\log^2 n)\) nondeterministic bits (gates with no inputs and one output). \(\text{AC}^0\) and \(\beta_2 \text{AC}^0\) are the restrictions of \(\text{FOLL}\) and \(\beta_2 \text{FOLL}\), respectively, to depth \(O(1)\). In general, the class \(\beta_2 \text{C} (\text{NC}, \text{respectively})\) is the class of languages decidable by \(\text{C}\) machines augmented with \(O(\log^2 n)\) bits (polynomial bits, respectively) of nondeterminism, or equivalently, the class of languages verifiable by \(\text{C}\) machines when given a certificate of length \(O(\log^2 n)\) (polynomial in \(n\), respectively).

If \(L_1\) and \(L_2\) are languages, there is a logarithmic space many-one reduction from \(L_1\) to \(L_2\), denoted \(L_1 \leq^L_{\text{m}1} L_2\), if there is a function \(f\) such that \(f\) is computable in logarithmic space and \(x \in L_1\) if and only if \(f(x) \in L_2\). There is a \(\beta_2 \text{AC}^0\) conjunctive truth-table reduction from \(L_1\) to \(L_2\), denoted \(L_1 \leq^\text{ctt}_i \beta_2 \text{AC}^0 L_2\), if there is a function \(f\) and a polynomial \(p\) such that

- \(f\) is computable in \(\beta_2 \text{AC}^0\),
- \(f(x)\) outputs \((y_1, \ldots, y_{p(n)})\) where \(n\) is the length of \(x\),
- \(x \in L_1\) if and only if \(\bigwedge_{i=1}^{p(n)} y_i \in L_2\).

**Lemma 2.1.** Suppose \(L_1\) and \(L_2\) are languages.

1. If \(L_1 \leq^\text{ctt}_{\text{ctt}} \text{NAC}^0 L_2\) and \(L_2\) is in \(\text{P}\), then \(L_1\) is in \(\text{NP}\).
2. If \(L_1 \leq^\text{ctt}_{\text{ctt}} \text{NAC}^0 L_2\) and \(L_2\) is in \(\text{NL}\), then \(L_1\) is in \(\text{NL}\).
3. If \(L_1 \leq^\text{ctt}_{\text{ctt}} \beta_2 \text{AC}^0 L_2\) and \(L_2\) is in \(\text{FOLL}\), then \(L_1\) is in \(\beta_2 \text{FOLL}\).
4. If \(L_1 \leq^\text{ctt}_{\text{ctt}} \beta_2 \text{AC}^0 L_2\) and \(L_2\) is in \(\text{L}\), then \(L_1\) is in \(\beta_2 \text{L}\).

**Proof.** In each case, let \(f\) denote the reduction and \(M_2\) denote the machine that decides \(L_2\). The machine that decides \(L_1\), call it \(M_1\), simulates \(f\) on its input then runs \(M_2\) on each component of the output of \(f\). The machine \(M_1\) accepts if and only if each of the simulations of \(M_2\) accepts. The correctness of \(M_1\) follows from the correctness of \(f\) and \(M_2\). The only remaining issue is the complexity of \(M_1\).

For the first two cases, we use the fact that \(\text{NAC}^0 \subseteq (\text{NP} \cap \text{NL})\). We define the machine \(M_1\) so that it chooses all its nondeterministic bits at the beginning of its computation. More specifically, if \(f\) requires \(p(n)\) nondeterministic bits, the length of the output of \(f\) is \(q(n)\), and \(M_2\) requires \(r(n)\) nondeterministic bits, then \(M_1\) uses at most \(p(n) + q(n)r(q(n))\) nondeterministic bits, which is a polynomial since \(p\), \(q\), and \(r\) are polynomials.

5
In the first case, the NP machine $M_1$, after choosing a sufficient number of nondeterministic bits, can simulate $f$ in polynomial time and can simulate a polynomial number of instances of $M_2$ (specifically, $q(n)$ instances) in polynomial time. In the second case, the NL machine does the same thing, but requires the fact that logarithmic space computable functions compose.

For the last two cases, we use the fact that $\beta_2\text{AC}^0 \subseteq (\beta_2\text{FOLL} \cap \beta_2\text{L})$. If $L_2$ is in FOLL, we define $M_1$ to be the circuit

$$M_1(x, w) = \bigwedge_{i=1}^{q(n)} M_2(y_i),$$

where $n$ is the length of $x$, the string $w$ is the nondeterministic string of length $O(\log^2 n)$, and $q(n)$ is the polynomial bounding the number of outputs of $f$ on inputs of length $n$. The depth of the $M_1$ circuit is the depth of $f$ plus the depth of $M_2$, which is $O(1) + O(\log \log n)$, or simply $O(\log \log n)$. The number of nondeterministic bits required by $M_1$ is the same as the number of nondeterministic bits required by $f$, which is $O(\log^2 n)$. The circuit is polynomial in size because $f$ is polynomial in size, $M_2$ is polynomial in size, and there are a polynomial number of parallel instances of the circuit $M_2$. Thus $M_1$ is in $\beta_2\text{FOLL}$.

The proof is similar if $L_2$ is in $\text{L}$. The only difference is that instead of a circuit computing the conjunction of $q(n)$ bits, we loop over each $y_i$ and check if each one causes $M_2$ to accept. Since there are a polynomial number of them, indexing them requires only logarithmic space. We also require the fact that logarithmic space computable functions compose.

Finally, if $L$ is a language and $F$ is a function, there is a nonadaptive $\text{AC}^0$ Turing reduction from $L$ to $F$ if there is an $\text{AC}^0$ function $g$ and an $\text{AC}^0$ circuit $C$ such that $x \in L$ if and only if $C(x, F(y_1), \ldots, F(y_m)) = 1$, where $(y_1, \ldots, y_m)$ is the output of $g(x)$ and $m$ is bounded by a polynomial in $|x|$. The function $g$ is called the generator of the reduction and the circuit $C$ is called the evaluator of the reduction.

### 2.2 Algebra

An magma is a set $G$ with a binary operation $\cdot$ that is closed on $G$. Unless otherwise stated, we will only consider finite magmas, in which $G$ is a finite set. The Cayley table of a magma with $n$ elements is the $n \times n$ table whose rows and columns are indexed by the elements of $G$ and where entry $(a, b)$ has value $c$ if $a \cdot b = c$. If the binary operation is associative, the magma is called a semigroup. A semigroup with a unique identity element is called a monoid. If the binary operation has the property that for each $a$ and $b$ in $G$ there are unique elements $x$ and $y$ in $G$ such that $a \cdot x = b$ and $y \cdot a = b$, the magma is called a quasigroup. (In other words, each quasigroup element appears exactly once in each row and each column of the Cayley table of $G$, or the Cayley table is a Latin square.) If a quasigroup is nonempty and associative, then it is a group. Alternately, if a semigroup has an identity and inverses, then it is a group.
Example 2.2. The smallest nonempty quasigroup that is not also a group has three elements, \( \{a, b, c\} \). Its Cayley table is

\[
\begin{array}{c|ccc}
    & a & b & c \\
a & a & b & c \\
b & c & a & b \\
c & b & c & a \\
\end{array}
\]

Examining the table reveals that there is exactly one of each quasigroup element in each row and column. This quasigroup is not associative because \( b \cdot (a \cdot b) = b \cdot b = a \) but \( (b \cdot a) \cdot b = c \cdot b = c \). Also, it has a left identity, \( a \), but no right identity. \( \square \)

Example 2.3. The right zero semigroup is the semigroup in which each element is a right zero. Its Cayley table is

\[
\begin{array}{c|ccc}
    & a & b & c \\
a & a & b & c \\
b & a & b & c \\
c & a & b & c \\
\end{array}
\]

The associativity of this semigroup can be determined by examining all possible triples \( (x, y, z) \) in \( G^3 \) and checking that \( x \cdot (y \cdot z) = (x \cdot y) \cdot z \). Each element of this semigroup is a left identity and a right zero, but there are no right identities, so it is not a group.

Unlike for the Latin square property in the previous example, there is no obvious way to tell whether a binary operation is associative simply by scanning the rows and columns. In other words, given only its Cayley table, determining whether a magma is a quasigroup seems easier than determining whether a magma is a semigroup. However, there is a polynomial time algorithm, attributed to F. W. Light, for deciding whether a magma is associative: it is simply the naïve algorithmic implementation of the associativity condition. \( \square \)

Example 2.4. There is a unique (up to isomorphism) group on three elements \( \{a, b, c\} \). Its Cayley table is

\[
\begin{array}{c|ccc}
    & a & b & c \\
a & a & b & c \\
b & b & c & a \\
c & c & a & b \\
\end{array}
\]

This group is the group of integers under addition modulo three. It is both a semigroup and a quasigroup. \( \square \)

A parenthesization \( P \) of a sequence of magma elements \( (g_1, \ldots, g_k) \) is a binary tree that has the magma elements as its leaves (in the order indicated by the sequence). The parenthesized product of a sequence of magma elements \( (g_1, \ldots, g_k) \) with parenthesization \( P \), denoted \( P(g_1, \ldots, g_k) \), is the element that results from performing the magma operation in the order indicated by the parenthesization.
Example 2.5. Consider \((a, c, a, b)\), a sequence of four elements from the quasigroup defined in [Example 2.2] One parenthesization of this sequence is
\[
\begin{array}{ccc}
  a & c & a \\
  a & b \\
\end{array}
\]
which corresponds to the parenthesized product \(a \cdot ((c \cdot a) \cdot b)\). According to the Cayley table, this product equals \(a\).

Lemma 2.6. For any quasigroup on \(n\) elements given as a Cayley table, any sequence \((g_0, \ldots, g_k)\), and any parenthesization \(P\) of depth \(d\) on that sequence, the parenthesized product \(P(g_0, \ldots, g_k)\) can be computed by an \(L\)-uniform family of unbounded fan-in circuits with size \(O(kn^2 \log n)\) and depth \(O(d)\).

Proof. How does a circuit access and use a Cayley table for a quasigroup? One way for a circuit to compute the product of two quasigroup elements using the Cayley table is via a multiplexer. In the multiplexer, each input has \(O(\log n)\) bits, since each quasigroup element can be represented with \(O(\log n)\) bits and each input is a pair of quasigroup elements. A multiplexer that selects from \(n^2\) inputs, each of length \(O(\log n)\), can be implemented by an unbounded fan-in circuit of depth \(O(1)\) and size \(O(n^2 \log n)\).

If \(S\) is a subset of magma elements, the submagma generated by \(S\), denoted \(\langle S \rangle\), is the closure of \(S\) under the magma operation and under any parenthesization. (For semigroups, and hence for groups, the operation is associative, so the parenthesization is superfluous.) If \(\langle S \rangle = G\), then \(S\) is called a generating set for \(G\). The rank of a magma \(G\), denoted \(\text{rank}(G)\), is the minimum cardinality of a generating set. This terminology extends to semigroups and groups as well.

Example 2.7. Consider the right zero semigroup on \(n\) elements, a generalization of [Example 2.3]. In this semigroup, call it \(G\), we have \(x \cdot y = y\) for each \(x\) and \(y\) in \(G\). The rank of this semigroup must be \(n\). Assume for the sake of producing a contradiction that the rank is strictly less than \(n\). Thus there is an element \(z\) not in the generating set such that \(x_1 \cdots x_n = z\), where each \(x_i\) is an element of the generating set. This is a contradiction with the fact that \(x_1 \cdots x_n = x_n\), since \(x_n\) is a right zero. Therefore the rank of the semigroup must be \(n\).

Example 2.8. Consider the elementary abelian 2-group, \((\mathbb{Z}/2\mathbb{Z})^k\), for some positive integer \(k\). Let \(n\) denote the order of this group, so \(n = 2^k\). The minimum generating set for this group is \(\{e_1, \ldots, e_k\}\), where \(e_i\) is the \(k\)-tuple with a one in the \(i\)th position and a zero in each other position (if we consider the group as a vector space, \(e_i\) is the standard basis vector). Thus the group has a minimum generating set of size \(k\), which is \(\log n\).

---

\(^1\)We avoid representing problems using first-order logic, as in the original definition of FOLL from [4], though the logic definition may provide a more natural representation of this sort of information.
For quasigroups, we consider a slightly more specific notion of “generating.” If \((g_0, \ldots, g_k)\) is a finite sequence of quasigroup elements denoted \(S\) and \(P\) is a parenthesization of that sequence, then the cube of \(S\) with respect to \(P\), denoted \(\text{cube}_P(S)\), is defined

\[
\text{cube}_P(S) = \{ P(g_0, g_1^{\epsilon_1}, \ldots, g_k^{\epsilon_k}) | \epsilon_i \in \{0, 1\} \text{ for each } i \},
\]

where \(g_i^{\epsilon_i}\) denotes \(g_i\) if \(\epsilon_i = 1\) and the empty word if \(\epsilon_i = 0\). The element \(g_0\) has no exponent because the empty word by itself is not an element in a quasigroup. (This is called a “cube” because each vertex of the \(k\)-dimensional Boolean hypercube, when interpreted as a binary string \(\epsilon_1 \cdots \epsilon_k\), yields a quasigroup element.) If \(\text{cube}_P(S) = G\), then \(g\) is called a cube generating sequence of size \(k + 1\) for the quasigroup \(G\). The rank of a quasigroup \(G\), denoted \(\text{rank}(G)\), is the minimum size of a cube generating sequence\(^2\). Contrast the rank of a quasigroup with the rank of a semigroup: the former is the size of a sequence, the latter the size of a set.

Ideally, we would like the notion of rank to be identical for each algebraic structure. If a quasigroup has a cube generating sequence of size at most \(k\), then it has a generating set of size at most \(k\), specifically the set of distinct elements from the sequence. However, we conjecture that for sufficiently large sizes, there is a quasigroup that has a generating set of size strictly less than the size of its minimum cube generating sequence. If this is incorrect, that is if a small generating set implies a small cube generating sequence, we could use the same definition of rank for all our algebraic structures, simplifying our proofs.

Quasigroups have small cube generating sequences, and groups have small generating sets. As of this publication, upper bounds on the size of generating sets for semigroups remain the subject of research\(^8\), although in general, some semigroups of order \(n\) have rank \(n\) (see Example 2.7). Magmas have even less structure than semigroups, and hence lack a meaningful upper bound as well.

An upper bound for the minimum size of a generating set for quasigroups can be proven by the probabilistic method.

**Lemma 2.9** ([6, Theorem 3.3]). Each finite quasigroup with \(n\) elements has a cube generating sequence of size \(O(\log n)\) with a parenthesization of depth \(O(\log \log n)\).

Since a group is a quasigroup, and since a cube generating sequence induces a generating set, the same upper bound can be applied to groups. However, a more specific (and constructive) upper bound can be proven inductively by considering cosets of increasing size.

**Lemma 2.10.** If \(G\) is a finite group of order \(n\) then the minimum size of a generating set is at most \(\log n\), with equality when the group is a finite elementary abelian 2-group.

\(^2\) This is a nonstandard definition of “rank” for quasigroups. Elsewhere, the rank of a quasigroup is the number of blocks in the partition of the quasigroup into conjugacy classes according to the action of the quasigroup on itself.
As stated in the introduction, computing the rank of a magma reduces to the problem of deciding submagma membership. This section reviews the complexity for the membership problem for magmas, semigroups, quasigroups, and groups.

We recall that for magmas the membership problem, given the Cayley table, is in \( \text{NP} \) and for semigroups \( \text{NL} \). We prove that for quasigroups the problem is in \( \text{FOLL} \) and for groups \( \text{L} \). These upper bounds allow us to prove upper bounds...
on the rank problem in the next section. If future work reveals more efficient algorithms for the membership problem for quasigroups or groups, we can provide improved algorithms for computing the rank of these algebraic structures.

The Submagma Membership problem is defined as follows. The inputs are a magma \( G \) given as a Cayley table, a magma element \( h \), and a finite set of magma elements \( S \). The problem is to decide whether \( h \in \langle S \rangle \).

**Lemma 3.1** ([11, Corollary 9]). Submagma Membership is \( \mathbf{P} \)-complete.

The Subsemigroup Membership problem is defined as follows. The inputs are a semigroup \( G \) given as a Cayley table, a semigroup element \( h \), and a finite set \( S \) of semigroup elements. The problem is to decide whether \( h \in \langle S \rangle \).

**Lemma 3.2** ([12]). Subsemigroup Membership is \( \mathbf{NL} \)-complete.

The Cube Membership problem is defined as follows. The inputs are a quasigroup \( G \) given as a Cayley table, a quasigroup element \( h \), a finite sequence of quasigroup elements \( S \), and a parenthesization \( P \) for that sequence. The problem is to decide whether \( h \in \text{cube}_P(S) \).

**Lemma 3.3** (Implicit in [6, Theorem 3.4]). Cube Membership is decidable by an \( \mathbf{L} \)-uniform family of unbounded fan-in circuits with size \( O(2^k kn^2 \log n) \) and depth \( O(d) \), where \( n \) is the order of the quasigroup, \( k \) is the size of the generating sequence, and \( d \) is the depth of the parenthesization.

In particular, if \( k = O(\log n) \) and \( d = O(\log \log n) \), then Cube Membership is in \( \mathbf{FOLL} \).

**Proof.** The input to the circuit is the Cayley table for a quasigroup, a quasigroup element \( h \), a generating sequence \( S \), and a parenthesization \( P \). Suppose \( S = (g_0, \ldots, g_k) \) for some positive integer \( k \). Since the circuit needs to determine if \( h \) is in \( \text{cube}_P(S) \), the circuit accepts if and only if there is some sequence of bits \((\epsilon_1, \ldots, \epsilon_k)\) such that \( h = P(g_0, g_{\epsilon_1}^1, \ldots, g_{\epsilon_k}^k) \). Thus the circuit consists of \( 2^k \) subcircuits joined to a single OR gate, each subcircuit deciding whether one of the \( 2^k \) possible \( k \)-bit sequences \((\epsilon_1, \ldots, \epsilon_k)\) produces \( h \) under the given parenthesization.

The subcircuit corresponding to binary sequence \((\epsilon_1, \ldots, \epsilon_k)\) computes the parenthesized product \( P(g_0, g_{\epsilon_1}^1, \ldots, g_{\epsilon_k}^k) \). Computing the parenthesized product can be implemented in \( O(kn^2 \log n) \) size and \( O(d) \) depth by [Lemma 2.6]. Comparing the element produced this way to the element \( h \) can be done with a constant depth, \( O(\log n) \) size equality comparison circuit.

We conclude that the overall size of the circuit is \( O(2^k kn^2 \log n) \) and the overall depth of the circuit is \( O(d) \).

Although the notion of cube generating sequence will give us a better upper bound for computing quasigroup rank, we prefer to consider the more natural notion of a generating set, as we did for magmas, semigroups, and quasigroups. Since we know from [Lemma 2.9] that we need only consider candidate generating sets of size \( O(\log n) \) and candidate parenthesizations of depth \( O(\log \log n) \), we
use a generalized form of the quasigroup membership problem that allows us to
specify bounds on the generating set size and parenthesization depth.

The **Bounded Subquasigroup Membership** problem is defined as follows. The inputs are a quasigroup \( G \) given as a Cayley table, a quasigroup element \( h \), a finite set \( S \) of semigroup elements, a positive integer \( k \), and a positive integer \( d \). The problem is to decide whether there is a sequence \( s \) in \( S^k \) and a parenthesization of depth \( d \) on \( k \) elements such that \( h = P(s) \). (This is the definition of "\( h \in \langle S \rangle \)”, but with specific size and depth bounds on the binary tree that generates \( h \).) This problem should be at least as difficult as **Cube Membership**: the former requires finding an appropriate sequence and parenthesization, whereas for the latter, they are fixed beforehand.

**Lemma 3.4.** **Bounded Subquasigroup Membership** is decidable by an \( L \)-uniform family of unbounded fan-in circuits with size \( O(n^2 k \log n) \) and depth \( O(d) \), using \( O(k \log n) \) nondeterministic bits.

In particular, if \( k = O(\log n) \) and \( d = O(\log \log n) \), then **Bounded Subquasigroup Membership** is in \( \beta_2\text{FOLL} \).

**Proof.** The algorithm is similar to that of **Lemma 3.3** except now we must nondeterministically choose a sequence and parenthesization. The circuit nondeterministically chooses \( k \) elements of \( S \) and a parenthesization of \( k \) elements of depth \( d \), then accepts if and only if that parenthesized product is \( h \). Choosing \( k \) elements, each of size \( O(\log n) \), requires \( O(k \log n) \) bits and choosing a parenthesization requires \( O(k) \) bits, so the total number of nondeterministic bits required is \( O(k \log n) \). By **Lemma 2.6** computing the parenthesized product requires a circuit of size \( O(kn^2 \log n) \) and depth \( O(d) \). The final equality comparison requires size \( O(\log n) \) and depth \( O(1) \). We conclude that the overall size of the circuit is \( O(kn^2 \log n) \) and the overall depth of the circuit is \( O(d) \). \( \square \)

The **Subgroup Membership** problem is defined as follows. The inputs are a group \( G \) given as a Cayley table, a group element \( h \), and a finite set \( S \) of group elements. The problem is to decide whether \( h \in \langle S \rangle \).

**Lemma 3.5.** **Subgroup Membership** is in \( L \).

**Proof.** The problem is in \( \text{SL} \) by a reduction to **Undirected Path** [5, Section 3], and \( \text{SL} = \text{L} \) [16]. \( \square \)

### 4 Computation of magma rank

Computing submagma membership is where most of the work occurs. Now we need only reduce the rank problem to the membership problem. We do this via a truth-table reduction of relatively low complexity. This section uses these reductions and the results of the previous section to prove the upper bounds on the rank problem as advertised in the introduction.

**Theorem 4.2** proves that for magmas, the rank problem is in \( \text{NP} \), for semigroups \( \text{NL} \), for quasigroups \( \beta_2\text{FOLL} \), and for groups \( \beta_2\text{FOLL} \cap \beta_2\text{L} \). This means
that for groups and semigroups, there is a polynomial time algorithm for computing the rank, and for quasigroups the problem can be verified quickly given a very short witness. We conjecture that for magmas and semigroups, the problems are hard for their respective complexity classes.

The Magma Rank problem is defined as follows. Given the Cayley table of a magma and an integer \( k \) in unary, decide whether the rank of the magma is \( k \) or less. The restrictions of this problem to quasigroups, semigroups, and groups, respectively, are defined similarly. The integer \( k \) is given in unary in order to facilitate the construction of uniform circuit families that decide the problem; since the size of the Cayley table is \( n^2 \log n \) and \( k \) is always at most \( n \), encoding the integer in unary does not cause an exponential increase in the size of the input to the problems.

The reductions in the following lemma are implicit in [6, Theorem 3.4]. That theorem demonstrates a \( \beta_2 \)FOLL algorithm for deciding whether two quasigroups are isomorphic, and the first part of that algorithm determines whether a given sequence of quasigroup elements with a parenthesization is a cube generating sequence.

**Lemma 4.1.**

1. Magma Rank \( \leq_{\text{ctt}} \) Submagma Membership.
2. Semigroup Rank \( \leq_{\text{ctt}} \) Subsemigroup Membership.
3. Quasigroup Rank \( \leq_{\text{ctt}} \) Cube Membership.
4. Quasigroup Rank \( \leq_{\text{ctt}} \) Bounded Subquasigroup Membership.
5. Group Rank \( \leq_{\text{ctt}} \) Subgroup Membership.

**Proof.** First, consider the problem for magmas. Unlike for quasigroups and groups [Lemma 2.9 and Lemma 2.10], we have no general upper bound on the minimum size of a generating set for magmas. Thus, the best we can do is nondeterministically choose a set of \( k \) generators and determine if that set generates the magma, where \( k \) can be as large as the number of elements in the magma.

Let \( g_1, \ldots, g_n \) denote the elements of a magma. The reduction proceeds as follows. On input \((G, k)\), where \( G \) is a magma on \( n \) elements given as its Cayley table and \( k \) is a positive integer given in unary, nondeterministically choose a sequence \( S \) of \( k \) magma elements. Output \(((G, g_1, S), \ldots, (G, g_n, S))\).

Since each magma element can be represented by \( O(\log n) \) bits, the number of nondeterministic bits used is \( O(n \log n) \). By definition of rank,

\[
\text{rank}(G) \leq k \iff \bigwedge_{i=1}^{n} g_i \in \langle S \rangle,
\]

so the reduction is a correct conjunctive truth-table reduction. For semigroups, we apply the exact same reduction.
For quasigroups, in the reduction to the cube membership problem, the only
differences are that we need to nondeterministically choose a parenthesization
as well as a generating sequence, and that we have an upper bound on the
size of the sequence and the parenthesization. By Lemma 2.9, it suffices to
consider inputs to Quasigroup Rank in which $k$ is in $O(\log n)$ and inputs to
Cube Membership in which $P$ is of depth $O(\log \log n)$. The output of the reduction is

$((G, g_1, S, P), \ldots, (G, g_n, S, P))$. Now the number of nondeterministic bits used
is $O(\log^2 n)$, since $S$ is a set of $O(\log n)$ strings, each of length $O(\log n)$. By Lemma 2.9,

$$\text{rank}(G) \leq k \iff \bigwedge_{i=1}^n g_i \in \text{cube}_P(S),$$

so the reduction is a correct conjunctive truth-table reduction.

For the reduction to the bounded quasigroup membership problem, we
still need $O(\log^2 n)$ bits to guess the generating set, but we can let $k = O(\log n)$ and $d = O(\log \log n)$, by Lemma 2.9. Thus the reduction outputs

$((G, g_1, S, k, d), \ldots, (G, g_n, S, k, d))$, and the correctness of the reduction follows
from the fact that

$$\text{rank}(G) \leq k \iff \bigwedge_{i=1}^n g_i \in \langle S \rangle.$$

The proof for groups is again similar. Instead of Lemma 2.9, we invoke
Lemma 2.10 which states that any group of order $n$ has a generating set of size
at most $\log n$. Also, we don’t need to guess a parenthesization (although we still
use $O(\log^2 n)$ nondeterministic bits to guess the generating set). Therefore, the
reduction will output $((G, g_1, S), \ldots, (G, g_n, S))$, and the proof concludes with
the fact that

$$\text{rank}(G) \leq k \iff \bigwedge_{i=1}^n g_i \in \langle S \rangle. \quad \square$$

**Theorem 4.2.**

1. **Magma Rank** is in $\text{NP}$.
2. **Semigroup Rank** is in $\text{NL}$.
3. **Quasigroup Rank** is in $\beta_2^{\text{FOLL}}$.
4. **Group Rank** is in $\beta_2^{\text{FOLL}} \cap \beta_2^{\text{L}}$.

**Proof.** Follows from Lemma 2.1 and Lemma 4.1 along with

1. Lemma 3.1 for magmas,
2. Lemma 3.2 for semigroups,
3. Lemma 3.3 for quasigroups,
Lemma 3.5 for groups.

For groups the problem is also in $\beta_2$FOLL since a group is a quasigroup.

These reductions can be generalized to the problem of computing the size of a minimum generating set for an arbitrary subset of the magma elements. The Generalized Magma Rank problem is defined as follows (and there are analogous problems for the other algebraic structures). Given a magma $G$ as a Cayley table, a finite set $T$ of magma elements, and a natural number $k$ in unary, decide whether there is a set $S$ of size at most $k$ such that $S \subseteq T \subseteq \langle S \rangle$. MAGMA RANK occurs as a special case when choosing $T = G$. Still, this problem reduces to the appropriate membership problem by the same reduction as in Lemma 4.1: nondeterministically choose a subset $S$ of $V$ with $|S| \leq k$, then decide whether each element of $V$ is in $\langle S \rangle$.

We can reprove [6, Theorem 3.4] using this strategy as well. The alternate proof is a reduction from QUASIGROUP ISOMORPHISM to the join of two languages, CUBE MEMBERSHIP and PRODUCT EQUALITY. The latter is the problem of deciding whether two parenthesized products are equal according to the quasigroup operation given as a Cayley table. If the parenthesization is of depth $O(\log \log n)$, this problem is in FOLL by Lemma 2.6.

Theorem 4.3 ([6, Theorem 3.4]). QUASIGROUP ISOMORPHISM is in $\beta_2$FOLL.

Proof. This is a brief overview of the alternate proof. We will show a $\beta_2$AC$^0$ conjunctive normal form truth-table reduction from QUASIGROUP ISOMORPHISM to the join of CUBE MEMBERSHIP and PRODUCT EQUALITY, both of which are in FOLL. Assuming this reduction exists, we conclude using a proof similar to that of Lemma 2.1 that QUASIGROUP ISOMORPHISM is in $\beta_2$FOLL.

The reduction first guesses two cube generating sequences, $g$ for $G$ and $h$ for $H$, both of length $O(\log n)$ and a parenthesization $P$ of depth $O(\log \log n)$, then outputs the conjunction of the following queries.

\begin{align*}
\bigwedge_{g \in G} g & \in \text{cube}_P(g) \quad (1) \\
\bigwedge_{h \in H} h & \in \text{cube}_P(h) \quad (2) \\
\bigwedge_{\epsilon,\eta,\nu \in \{0,1\}^k} (P(g^\epsilon) = P(g^\eta) \cdot P(g^\nu) \iff P(h^\epsilon) = P(h^\eta) \cdot P(h^\nu)) \quad (3)
\end{align*}

The first two formulas ensure that $g$ and $h$ are cube generating sequences for $G$ and $H$, respectively. In the third formula, if $g = (g_0, g_1, \ldots, g_k)$, then $g^\epsilon$ denotes $(g_0^{\epsilon_0}, g_1^{\epsilon_1}, \ldots, g_k^{\epsilon_k})$. This formula checks that the bijection $g_i \mapsto h_i$ is a homomorphism.

Each of the first two formulas comprises $n$ conjunctive queries to CUBE MEMBERSHIP. The last formula comprises a polynomial number of queries in conjunctive normal form to PRODUCT EQUALITY. Thus we have the required reduction.
We conclude this section with a few observations about Theorem 4.2. First, in this proof, we did not use the reduction from Quasigroup Rank to Bounded Subquasigroup Membership, because the closure of \( \beta_2 \text{FOLL} \) under \( \beta_2 \text{AC}^0 \) conjunctive truth-table reductions is \( \text{NFOLL} \), that is, \( \text{FOLL} \) with a polynomial amount of nondeterminism, whereas the closure of \( \text{FOLL} \) under the same reductions is \( \beta_2 \text{FOLL} \), a subset of \( \text{NFOLL} \).

Second, a slight generalization of [15, Theorem 7] already proves that Magma Rank is in (and complete for) the class of problems decidable by a polynomial time Turing machine with \( O(n \log n) \) nondeterministic bits. We have nevertheless included the fact that Magma Rank is in \( \text{NP} \) to highlight the general strategy for proving these upper bounds for each class of algebraic structure.

Third, we can almost show a reduction in the opposite direction of Lemma 4.1. The Submagma Rank problem (a search problem) is defined as follows. Given the Cayley table of a magma and a set of magma elements \( S \), output the rank of \( \langle S \rangle \). (The Submagma Rank problem is more general than the Magma Rank problem: the latter reduces to the former by choosing \( S = G \).)

**Proposition 4.4.** Submagma Membership reduces to the Submagma Rank function by a nonadaptive \( \text{AC}^0 \) Turing reduction making exactly two queries.

**Proof.** We know that for any magma \( G \), any magma element \( h \), and any subset of magma elements \( S \),

\[
h \in \langle S \rangle \iff \text{rank}(\langle S \rangle) = \text{rank}(\langle S \cup \{h\} \rangle).
\]

(This is analogous to the corresponding situation in linear algebra: a vector \( h \) is in the span of a set of vectors \( S \) exactly when the rank of \( S \) does not increase when \( h \) is added to \( S \).) Thus the generator of the reduction is the function \((G, h, S) \mapsto ((G, S), (G, S \cup \{h\}))\) and the evaluator of the reduction compares \( \text{rank}(S) \) to \( \text{rank}(\langle S \cup \{h\} \rangle) \) for equality. \( \square \)

However, this reduction is not satisfying, because the Submagma Rank is essentially the Magma Rank problem when the input is provided as a set of generators instead of as a Cayley table. As stated in the reduction, this representation may be exponentially smaller than the Cayley table representation.

Finally, although the precise relationship between \( \text{FOLL} \) and \( \text{L} \) is unknown, \( \text{FOLL} \) does not contain any class containing the Parity problem. Since Parity is in \( \text{L} \), we know \( \text{FOLL} \) does not contain \( \text{L} \). Stated in a slightly more general way, \( \text{FOLL} \) cannot be hard under \( \text{AC}^0 \) many-one reductions for any complexity class that contains Parity [4, Proposition 2.1]. This is true even when the circuit is augmented with a polylogarithmic number of nondeterministic bits [6, Section 4]. This gives an immediate improvement to the upper bound of the Quasigroup Rank problem.

**Theorem 4.5.** Quasigroup Rank is not hard under \( \text{AC}^0 \) many-one reductions for any complexity class containing Parity.
Specifically, Quasigroup Rank is not hard for any of the classes in the inclusion chain

$$ACC^0 \subseteq TC^0 \subseteq NC^1 \subseteq L \subseteq NL \subseteq (LOGCFL \cup DET).$$

5 Computation of ring rank

The previous section provides an upper bound on the computational complexity of computing the rank of a group. Is it interesting to ask about the “rank” of rings or fields, and what does rank mean for these algebraic structures? A ring comprises an additive group and a multiplicative monoid with an additional distributivity property relating the two. We represent it as a pair of Cayley tables, one for the group and one for the monoid.

Since our purpose for asking questions about the rank of an algebraic structure is to determine the smallest number of elements required to generate all other elements, we define the rank of a ring $R$ to be the minimum cardinality of a set $S$ such that $\langle S \rangle = R$, where $\langle S \rangle$ is the closure of $S$ under both addition and multiplication. The rank of a ring is bounded above by the minimum of the rank of its additive group and the rank of its multiplicative monoid. This section applies the same strategy used for Magma Rank to determine the complexity of Ring Rank: reduce the rank problem to the membership problem, then show an upper bound on the membership problem.

Theorem 5.3 below reveals that our best upper bound for computing the rank of a ring given as a pair of Cayley tables is $NL$, the same as our best upper bound for computing the rank of a semigroup. This makes sense because the hardness in this problem lies in computing the rank of the underlying monoid (which is just a semigroup with an identity element, so the worst-case complexity of the rank problem for monoids and for semigroups is exactly the same). Improving the algorithm for computing the rank of a monoid (or semigroup) will immediately improve the algorithm for computing the rank of a ring.

Not all rings have an interesting rank problem. In the special case that the ring is a finite domain (that is, it has no nontrivial zero divisors), then by Wedderburn’s little theorem [14, Theorem 3 § 11.1], the ring is a finite field. The multiplicative group of any finite field is cyclic [14, Theorem 7 § 6.4], so the rank of a finite field, and hence any finite domain, is one and the computational problem is uninteresting. But in general, for arbitrary (commutative or non-commutative) rings, the problem has nontrivial complexity.

Lemma 5.1. Ring Rank $\leq_{\text{ctt}}^{2\text{AC}^0}$ Subring Membership.

Proof. The reduction is identical to the ones in Lemma 4.1 and uses only $O(\log^2 n)$ bits of nondeterminism because any generating set must generate the additive group, and the additive group has a generating set of size at most $\log n$ by Lemma 2.10.

Lemma 5.2. Subring Membership $\leq_m$ Directed Path.
Proof. Given a ring \( R \) (expressed as two Cayley tables, one for addition and one for multiplication), a ring element \( r \), and a set of ring elements \( S \), construct a directed, labeled graph as follows. The set of vertices is the set of all ring elements. There is an edge from \( x \) to \( y \) labeled \((a, b)\) if \( xa - b = y \). Let \( G \) be the subgraph induced by edges labeled by pairs \((a, b)\) where both \( a \) and \( b \) are elements of \( S \). Output the subgraph \( G \), the source vertex 1, and the target vertex \( r \).

The choice of edges \((x, y)\) where \( xa - b = y \) for some \( a \) and \( b \) in \( S \) deserves some justification. First, choosing the subgraph induced by edges of the form \((a, 0)\) yields the Cayley graph of the ring’s underlying multiplicative monoid. Similarly, choosing the subgraph induced by edges of the form \((1, b)\) yields the Cayley graph of the ring’s underlying additive group. Second, the subring test states that a subset of a ring is a subring if it is closed under multiplication and subtraction and contains the identity element. Thus the transitive closure of the vertex representing the multiplicative identity under edges labeled with elements from the subset \( S \) is guaranteed to be a subring.

Looping over each pair of ring elements \((x, y)\) and each pair \((a, b)\) requires \(O(\log n)\) space for a ring with \( n \) elements. Deciding whether to add a labeled edge from \( x \) to \( y \) requires a constant number of Cayley table lookups, again requiring \(O(\log n)\) space. Thus the reduction is computable in logarithmic space.

To prove correctness, suppose \( r \) is in the subring generated by \( S \). One way to see that there is a path from 1 to \( r \) is to consider the sequence of multiplications and additions that produce \( r \) from 1. Let \((c_1, \ldots, c_n)\) be the sequence of ring elements and \((*_1, \ldots, *_n)\) be the finite sequence of ring operations, each one either an addition or a multiplication, such that \( 1 *_1 c_1 \cdots *_n c_n = r \). The sequences must be finite because the cardinality of the ring is finite. Then the path from vertex 1 to vertex \( r \) is the sequence of edges \(((a_1, b_1), \ldots, (a_n, b_n))\), where \((a_i, b_i)\) is \((c_i, 0)\) if \(*_i\) is multiplication or \((1, -c_i)\) if \(*_i\) is addition, for each \( i \in [n] \).

For the converse, suppose there is a path from 1 to \( r \) of length \( k \) in the graph \( G \), where \( k \leq n \). Let \(((a_1, b_1), \ldots, (a_k, b_k))\) be the labels along the edges of that path. To facilitate the closed-form representation of \( r \), let \( b_0 = -1 \) and \( a_{k+1} = 1 \). By construction of the graph \( G \), we know \( r = \cdots ((1a_1 - b_1)a_2 - b_2) \cdots a_k - b_k \), or more concisely,

\[
r = - \sum_{i=0}^{k} \left[ b_i \prod_{j=i+1}^{k+1} a_j \right].
\]

As stated previously, the transitive closure of 1 in \( G \) is the subring generated by \( S \). By the formula above, \( r \) is in the transitive closure of \( S \), so it is in the subring generated by \( S \).

**Theorem 5.3.** Ring Rank is in NL.

**Proof.** Directed Path is in (and complete for) NL, and NL is closed under logarithmic space many-one reductions, so Subring Membership is in NL, by
Lemma 5.2 Using the fact that $\beta_2 AC^0$ is a subset of $NAC^0$, Lemma 2.1 and Lemma 5.1 imply that $RING$ RANK is in NL.

If $\text{SUBRING$ MEMBERSHIP}$ reduces to $\text{SUBGROUP$ MEMBERSHIP}$ by a sufficiently tight reduction, then we could improve the upper bound for $RING$ RANK to $\beta_2 L$. However, such a reduction seems unlikely, since access to both addition and multiplication should allow more ring elements to be generated from a given set than access to addition alone. The converse reduction seems unlikely as well, since an arbitrary group is not necessarily abelian, and a non-abelian group does not admit a ring structure. Even for abelian groups, the previous concern about access to both addition and multiplication applies. We conjecture that the two problems are incomparable with respect to many-one reductions of sufficiently low complexity.

As stated in the introduction to this section, the rank of a ring is bounded above by the rank of its subgroup and of its monoid. If this inequality is tight, then $RING$ RANK reduces to the join of $GROUP$ RANK and $MONOID$ RANK, by computing the minimum of the solution to each of these two problems. However, we conjecture that there is a finite ring whose rank is strictly less than the rank of both its group and its monoid.

References


