Word equations in a uniquely divisible group

Christopher J. Hillar[1†] and Lionel Levine[2‡] and Darren Rhea[3]

[1]*The Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720-5070, USA*
[2]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA USA*
[3]*Department of Mathematics, University of California, Berkeley, CA 94720, USA*

Abstract. We study equations in groups G with unique m-th roots for each positive integer m. A *word equation* in two letters is an expression of the form $w(X, A) = B$, where w is a finite word in the alphabet $\{X, A\}$. We think of $A, B \in G$ as fixed coefficients, and $X \in G$ as the unknown. Certain word equations, such as $XAXAX = B$, have solutions in terms of radicals: $X = A^{-1/2}(A^{1/2}BA^{1/2})^{1/3}A^{-1/2}$, while others such as $X^2AX = B$ do not. We obtain the first known infinite families of word equations not solvable by radicals, and conjecture a complete classification. To a word w we associate a polynomial $P_w \in \mathbb{Z}[x, y]$ in two *commuting* variables, which factors whenever w is a composition of smaller words. We prove that if $P_w(x^2, y^2)$ has an absolutely irreducible factor in $\mathbb{Z}[x, y]$, then the equation $w(X, A) = B$ is not solvable in terms of radicals.

Résumé. Nous étudions des équations dans les groupes G avec les m-th racines uniques pour chaque nombre entier positif m. Une *équation de mot dans deux lettres* est une expression de la forme $w(X, A) = B$, où w est un mot fini dans l'alphabet $\{X, A\}$. Nous pensons $A, B \in G$ en tant que coefficients fixes, et $X \in G$ en tant que inconnu. Certaines équations de mot, telles que $XAXAX = B$, ont des solutions en termes de radicaux: $X = A^{-1/2}(A^{1/2}BA^{1/2})^{1/3}A^{-1/2}$, alors que d'autres tel que $X^2AX = B$ ne font pas. Nous obtenons les familles infinies d'abord connues des équations de mot non solubles par des radicaux, et conjecturons une classification complété. Á un mot w nous associons un polynôme $P_w \in \mathbb{Z}[x, y]$ dans deux variables de permutation, qui factorise toutes les fois que w est une composition de plus petits mots. Nous montrons que si $P_w(x^2, y^2)$ a un facteur absolument irréductible dans $\mathbb{Z}[x, y]$, alors l'équation $w(X, A) = B$ n'est pas soluble en termes de radicaux.

Keywords: absolutely irreducible, polynomials over finite fields, solutions in radicals, uniquely divisible group, word equation

1 Introduction

A group G is called *uniquely divisible* if for every $B \in G$ and each positive integer m, there exists a unique $X \in G$ such that $X^m = B$. We denote the unique such X by $B^{1/m}$, and its inverse by $B^{-1/m}$. In the literature, such groups are also referred to as $\mathbb{Q}$-groups. Note that if it is not the trivial group, then G must be torsion-free, hence infinite. Examples of uniquely divisible groups include the group of positive

†Partially supported by an NSA Young Investigators Grant and an NSF All-Institutes Postdoctoral Fellowship administered by the Mathematical Sciences Research Institute through its core grant DMS-0441170
‡Supported by an NSF Postdoctoral Fellowship

units of a real closed field, unipotent matrix groups, noncommutative power series with unit constant term, and the group of characters of a connected Hopf algebra over a field of characteristic zero [1].

Inspired by trace conjectures in matrix analysis (see section 2), we study here the natural question of which equations in a uniquely divisible group have solutions in terms of radicals. As a motivating example, consider the Riccati equation $XAX = B$ with $A, B \in G$ given and $X \in G$ unknown. This equation has a unique solution X in any uniquely divisible group; moreover, its solution may be written explicitly (albeit in two distinct ways) as

$$X = A^{-1/2}(A^{1/2}BA^{1/2})^{1/2}A^{-1/2} = B^{1/2}(B^{-1/2}A^{-1}B^{-1/2})^{1/2}B^{1/2}. \tag{1}$$

More generally, let $w(X, A)$ be a finite word in the two-letter alphabet $\{X, A\}$. An expression of the form

$$w(X, A) = B \tag{2}$$

is called a *word equation*. We are interested in classifying those word equations that have a solution in every uniquely divisible group. Clearly, the more general situation in which positive rational exponents on X, A, and B are allowed reduces to this one.

The main tool in our analysis is a new combinatorial object $P_w \in \mathbb{Z}[x, y]$, called the *word polynomial*. If $w = A^{a_0}XA^{a_1}X \cdots A^{a_{n-1}}XA^{a_n}$, we define

$$P_w(x, y) := y^{a_0} + xy^{a_0+a_1} + x^2y^{a_0+a_1+a_2} + \cdots + x^{n-1}y^{a_0+\cdots+a_{n-1}}. \tag{3}$$

For a prime p, let $(\mathbb{Z}/p\mathbb{Z})^*$ denote the set of nonzero elements of the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Theorem 1.1 *There exists a uniquely divisible group G with the following property: For all finite words w in the alphabet $\{X, A\}$, if the equation $P_w(x^2, y^2) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p, then there exist elements $A, B, B' \in G$ for which the word equation $w(X, A) = B$ has no solution $X \in G$, and the word equation $w(X, A) = B'$ has at least two solutions $X \in G$.*

We prove this theorem in section 6. The group G is constructed from an infinite collection of pq-groups whose orders are chosen using Dirichlet's theorem on primes in arithmetic progressions.

Despite appearances, Theorem 1.1 yields a computationally efficient sufficient condition for the equation $w(X, A) = B$ to have no solution in G (see Corollary 1.7). To put Theorem 1.1 in context, we next discuss a family of word equations which *do* have solutions in every uniquely divisible group, along with a conjectured complete classification of such words.

In this paper, "word" will always mean a finite word over the alphabet $\{X, A\}$ (unless another alphabet is specified). The word w is called *universal* if (2) has a solution $X \in G$ for every uniquely divisible group G and each two elements $A, B \in G$; if this solution is always unique, then we say that w is *uniquely universal*. A related class of words is those for which (2) has a solution "in terms of radicals." This notion is defined carefully in section 3. Our explorations give evidence for the surprising conjecture that all three of these classes are in fact the same, and can be characterized as follows.

Definition 1.2 *A word w in the alphabet $\{X, A\}$ is totally decomposable if it is the image of the letter X under a composition of maps of the form*

- $\pi_{m,k} : w \mapsto (wA^k)^m w$, *for* $m \geq 1$, $k \geq 0$

- $r : w \mapsto wA$
- $l : w \mapsto Aw$.

For example, the word $w = XAX^2AXAXAX^2AX$ is totally decomposable, as witnessed by the composition $w = \pi_{1,1} \circ \pi_{1,0} \circ \pi_{1,1}(X)$. According to the following lemma, any totally decomposable word is uniquely universal.

Lemma 1.3 *Let G be a uniquely divisible group, and let w be a totally decomposable word. For any $A, B \in G$, the equation $w(X, A) = B$ has a unique solution $X \in G$, and this solution can be expressed in terms of radicals.*

We conjecture the converse: totally decomposable words are the only universal words.

Conjecture 1.4 *Let w be a finite word in the alphabet $\{X, A\}$. The following are equivalent.*

1. *w is totally decomposable.*
2. *w is uniquely universal.*
3. *w is universal.*
4. *$w(X, A) = B$ has a solution in terms of radicals. (see Definition 3.4)*

The implications $(1) \Rightarrow (2) \Rightarrow (3)$ and $(3) \Leftrightarrow (4)$ are straightforward (see section 3). The remaining implication $(4) \Rightarrow (1)$ is the difficult one. Theorem 1.1 arose out of our attempts to prove this implication. It reduces the noncommutative question about word equations to a commutative question about solutions to polynomial equations mod p. More concretely, we use Theorem 1.1 to prove that several infinite families of word equations are not solvable in terms of radicals (Corollary 1.8). To our knowledge, these are the first such infinite families known.

Together with Theorem 1.1, the following would imply Conjecture 1.4.

Conjecture 1.5 *If w is a word that is not totally decomposable, then the equation $P_w(x^2, y^2) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p.*

The questions outlined above (e.g., asking whether a particular word w is universal, or uniquely universal, or solvable in terms of radicals) are examples of decidability questions in first-order theories of groups. Determining whether a set of equations has a solution in a group is known as the *Diophantine problem*. More generally, given a set of axioms for a class of groups, one would like to provide an algorithm which decides the truth or falsehood of any given sentence in the theory. That such an algorithm exists for free groups follows from pioneering work of Kharlampovich and Myasnikov [16] (see also the independent work of Sela [27]), but it is still open whether one exists for free uniquely divisible groups $F^{\mathbb{Q}}$. The Diophantine problem for $F^{\mathbb{Q}}$ does admit such an algorithm [15] although the time complexity of the algorithm described there is likely at least doubly exponential (the proof uses the decidability of Presburger arithmetic). In contrast, Conjecture 1.4 says that the Diophantine problem of a single equation in one variable reduces to an easily verifiable combinatorial condition (total decomposability).

Example 1 *Consider the word $w = X^2AX$, which is not totally decomposable; we use Theorem 1.1 to show that w is not universal. Its word polynomial is $P_w(x, y) = 1 + x + x^2y$. We need to verify that the equation $1 + x^2 + (x^2y)^2 = 0$ has a nonzero solution modulo p, for all sufficiently large primes p. A standard pigeonhole argument shows that for all primes p, there exist $a, b \in \mathbb{Z}$ with $a \neq 0$ such that $1 + a^2 + b^2 \equiv 0 \ (mod\, p)$. If $b = 0$ and $p \geq 5$, then $1 + (-1 + 4^{-1})^2 + (a + a4^{-1})^2 \equiv 0 \ (mod\, p)$ so that we may assume both a, b nonzero. Setting $x = a$ and $y = ba^{-2}$ gives us our solution.*

The word polynomial for the totally decomposable word $v = XAXAX$, on the other hand, is $P_v(x, y) = 1 + xy + x^2y^2$. Let p be a prime greater than 3. If $x, y \in (\mathbb{Z}/p\mathbb{Z})^*$ satisfy $P_v(x^2, y^2) = 0$, then setting $z = x^2y^2$ we have $z^3 = 1$ and $z \neq 1$, which forces $p \equiv 1 \ (mod\, 3)$. □

Recall that a polynomial over a field K is *absolutely irreducible* if it remains irreducible over every algebraic extension of K. The next result shows that to verify Conjecture 1.5 for a particular word w, it suffices to prove that a factor of $P_w(x^2, y^2)$ is absolutely irreducible.

Proposition 1.6 *Suppose $F \in \mathbb{Z}[x, y]$ satisfies $F(0, 0) \neq 0$, and F has a factor $f \in \mathbb{Z}[x, y]$ which is irreducible over $\mathbb{C}[x, y]$. Then the equation $F(x, y) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p.*

Corollary 1.7 *If w is a word in the alphabet $\{X, A\}$ beginning with X, and if $P_w(x^2, y^2)$ has a factor $f \in \mathbb{Z}[x, y]$ such that f is irreducible in $\mathbb{C}[x, y]$, then w is not universal.*

Example 2 *We show the usefulness of Corollary 1.7 by revisiting Example 1. The word $w = X^2AX$ has $P_w(x^2, y^2) = 1 + x^2 + x^4y^2$, which is irreducible over $\mathbb{C}$ (since $1 + x^2$ is not a square in $\mathbb{C}(x)$). It follows that X^2AX is not universal. In contrast, the totally decomposable word $v = XAXAX$ has $P_v(x^2, y^2) = 1 + x^2y^2 + x^4y^4 = (1 + xy + x^2y^2)(1 - xy + x^2y^2)$. Each factor on the right side is irreducible over $\mathbb{Z}$ but factors over $\mathbb{C}$.* □

In Section 7, we use Corollary 1.7 to verify Conjecture 1.5 for the following infinite families of words.

Corollary 1.8 *The following families of words do not have their equations solvable in terms of radicals:*

$$X^nAX^m,\ m, n \geq 1,\ m \neq n;\ \ XA^{m+2n}XA^{m+n}XA^mX,\ m \geq 0, n \geq 1;$$

$$XAX^nAX,\ n \geq 3;\ \ X^2(AX)^nX,\ n \geq 2.$$

Using Corollary 1.7 and the symbolic computation software Maple, we have also verified Conjecture 1.4 for all words of length at most 10. The most difficult part of the computation is to check whether a given bivariate polynomial over $\mathbb{Z}$ is irreducible over $\mathbb{C}$. This can be done in polynomial-time using the algorithm of Gao [8] (and is implemented in Maple).

We do not know if the condition in Corollary 1.7 is sufficient to prove Conjecture 1.4.

Question 1.9 *If the word w is not totally decomposable and begins with X, must $P_w(x^2, y^2)$ have a factor in $\mathbb{Z}[x, y]$ which is irreducible over $\mathbb{C}[x, y]$?*

The remainder of the paper is organized as follows. In section 2, we give additional motivation arising from the BMV trace conjecture in quantum statistical mechanics. In section 3, we review the basic properties of uniquely divisible groups and construct a free uniquely divisible group on two free generators. This construction allows us to define the notion of solvability in terms of radicals, but it is not needed for the proof of Theorem 1.1. Section 4 describes some important examples of uniquely divisible groups.

Most of the standard examples have the property that *every* word equation with nonnegative exponents has a unique solution; the need to construct more exotic groups is part of what makes Conjecture 1.4 so difficult. Section 5 discusses properties of the word polynomial P_w, and section 6 is devoted to the proof of Theorem 1.1. These two sections form the heart of the paper. Finally, Section 7 contains the proof of Corollary 1.8.

2 Background and Motivation

The Lieb-Seiringer formulation [23] of the long-standing Bessis-Moussa-Villani (BMV) trace conjecture [5, 21, 26, 11, 9, 19, 18, 17, 7] in statistical physics says that the trace of $S_{m,k}(A, B)$, the sum of all words of length m in A and B with k Bs, is nonnegative for all $n \times n$ positive semidefinite matrices A and B. In the case of 2×2 matrices, more is true: every word in two positive semidefinite letters has nonnegative trace (in fact, nonnegative eigenvalues). It was unknown whether such a fact held in general until [14] appeared where it was found (with the help of Shaun Fallat) that the word $w = BABAAB$ has negative trace with the positive definite matrices

$$A_1 = \begin{bmatrix} 1 & 20 & 210 \\ 20 & 402 & 4240 \\ 210 & 4240 & 44903 \end{bmatrix} \quad \text{and} \quad B_1 = \begin{bmatrix} 36501 & -3820 & 190 \\ -3820 & 401 & -20 \\ 190 & -20 & 1 \end{bmatrix}.$$

Finding such examples is surprisingly difficult, and randomly generating millions of matrices (from the Wishart distribution) fails to produce them. Nonetheless, it is believed that most words can have negative trace, and it was conjectured [14] that if a word has positive trace for every pair of real positive definite A and B, then it is a palindrome or a product of two palindromes (the converse is well-known). If we replace the words "positive trace" in the previous sentence with "positive eigenvalues," we obtain a weaker conjecture which was also studied in [14]. Further evidence for this conjecture can be found in [12], where it was proved that a generic word has positive definite complex Hermitian matrices A and B giving it a nonpositive eigenvalue.

Positive definite matrices which give a word a negative trace are also potential counterexamples to the BMV conjecture, and it is useful to be able to generate these matrices (see [10, §4.1] and [2, §11] for two such examples). As remarked above, this is difficult since random sampling does not seem to work. This discussion explains some of the subtlety of the BMV conjecture: most words occurring in $S_{m,k}(A, B)$ likely can be made to have negative trace; however, a particular word has a small proportion of matrices which witness this.

Although the set of $n \times n$ positive definite matrices is not a group for $n > 1$, every positive definite matrix has a unique positive definite m-th root for any m. More remarkably, it turns out [13, 2] that every word equation $w(X, A) = B$ with w palindromic (and containing at least one X) has a positive definite solution X for each pair of positive definite A and B (although this solution can be non-unique [2]). Using A_1 and B_1, it follows that any word of the form $wAwAAw$ with $w = w(B, A)$ palindromic (and containing at least one B) can have negative trace. This gives an infinite family verifying the conjecture of [14], and moreover, provides an infinite number of potential counterexamples to the BMV conjecture. The existence proof in [13] uses fixed point methods, although for special cases (e.g. when w contains four or less Xs), one may express solutions X explicitly (and computationally efficiently) in terms of A, B and fractional powers [2, §5]. Computing solutions without using these formal representations "in terms of radicals" is difficult [2, Remark 11.3], and it is believed that most equations do not have solutions expressable in this manner. For instance, there is no known expression for the solution to $XAX^3AX = B$ although there is always a unique positive definite solution [20].

3 Radical words and the free uniquely divisible group

In this section we review some basic properties of uniquely divisible groups, and construct the free uniquely divisible group on two generators. This construction allows us to define precisely the notion of "solvability in terms of radicals" (but we emphasize that the proof of Theorem 1.1 does not rely on this construction). The following lemma shows that rational powers of group elements are well-defined and behave as expected.

Lemma 3.1 *Let G be a uniquely divisible group and $a \in G$. Define $a^{n/m} := (a^n)^{1/m}$ for $n \in \mathbb{Z}$ and $0 \neq m \in \mathbb{N}$, and define $a^0 := 1$. Then if $p, q \in \mathbb{Q}$, we have $(a^p)^q = (a^q)^p = a^{pq}$ and $a^pa^q = a^qa^p = a^{p+q}$.*

A detailed study of uniquely divisible groups can be found in the thesis of Baumslag [3] where they are called *divisible R-groups*. See also [22] for a study of the metabelian case. As remarked in [3], one of the difficulties is that there is no clear normal form for uniquely divisible group elements (for example, see (1) from the introduction). There is, however, the notion of a free uniquely divisible group which comes out of Birkhoff's theory of "varieties of algebras" [6]. Since the construction is simple, we briefly outline the main ideas here. Our perspective is model-theoretic (see [25] for background) although we will use only basic notions from that subject.

Let T be the first-order theory of uniquely divisible groups. The underlying language and axioms of this theory are those of groups, with an additional (countably infinite) set of axioms expressing that every element has a unique m-th root for each positive integer m. Consider the smallest set S of finite, formal expressions containing letters $\{A, B\}$, exponents of the form $^{n/m}$ ($n \in \mathbb{Z}$, $0 \neq m \in \mathbb{N}$), and balanced parentheses that is closed under taking concatenations and powers (and contains the empty expression). For example, S contains the two rightmost expressions in (1).

If G is a uniquely divisible group and $a, b \in G$, then an expression $e = e(A, B) \in S$ defines unambiguously (by Lemma 3.1) an element $e(a, b) \in G$ by replacing letters $\{A, B\}$ with corresponding group elements $\{a, b\}$ and then evaluating the result in G. When two expressions $e, f \in S$ evaluate to the same group element for each pair $a, b \in G$ in every uniquely divisible group G, we write $e \sim f$. For instance, the two rightmost expressions in (1) are equivalent in this way. Although we will not need it here, Gödel's completeness theorem (along with soundness) implies that $e \sim f$ if and only if there is a (finite) formal proof from the axioms of T that they are equal.

Note that $\sim$ is an equivalence relation on S, and we write $[e]$ for the equivalence class containing $e \in S$.

Definition 3.2 *The set $\mathcal{F} := \{[e] : e \in S\}$ with multiplication $[e] \cdot [f] = [ef]$ is called the free uniquely divisible group on letters $L = \{A, B\}$.*

The definition extends in the obvious way to define the free uniquely divisible group on any set L, but (except for a remark at the very end of the paper) we shall only use the case of two generators.

The main facts about $\mathcal{F}$ that we will need are summarized in the following lemma. We remark that any homomorphism $\psi : F \to G$ between uniquely divisible groups is easily seen to satisfy $\psi(a^q) = \psi(a)^q$ for all $a \in F$ and $q \in \mathbb{Q}$.

Lemma 3.3 *$\mathcal{F}$ is a uniquely divisible group. Moreover, $\mathcal{F}$ satisfies a universal property with respect to the map $\theta : L \to \mathcal{F}$ sending $A \mapsto [A]$ and $B \mapsto [B]$: Given any uniquely divisible group G and any map*

$\phi : L \to G$, there exists a unique homomorphism (of uniquely divisible groups) $\psi : \mathcal{F} \to G$ such that $\psi \circ \theta = \phi$.

This discussion allows us to formally define the concept of solution in terms of radicals mentioned in the introduction (specifically, in the statement of Conjecture 1.4).

Definition 3.4 *A word w is called radical (and has equation $w(X, A) = B$ solvable in terms of radicals) if the equation $w(X, [A]) = [B]$ has a solution $X \in \mathcal{F}$.*

We now connect this definition with the idea of word equations having solutions in radicals. Let G be a uniquely divisible group. A subgroup $R \subseteq G$ is *radical* if $x^m \in R$ implies $x \in R$ for all $x \in G$ and all positive integers m. Given a subset H of G, recursively define sets R_n for $n \geq 0$ by setting $R_0 = H$ and

$$R_{n+1} = \{(xy)^q : x, y \in R_n,\ q \in \mathbb{Q}\}.$$

We call the union $\mathcal{R}(H) := \bigcup_{n \in \mathbb{N}} R_n$ the *radical subgroup of G generated by H*. One easily checks that $\mathcal{R}(H)$ is a radical subgroup of G and that it is the intersection of all radical subgroups containing H.

For a uniquely divisible group G and $a, b \in G$, the subgroup $\mathcal{R}(\{a, b\})$ can be thought of as the radical expressions generated by a and b. Given a specific instance of the word equation $w(x, a) = b$, any solution $x \in \mathcal{R}(\{a, b\})$ can be viewed as one "in terms of radicals." Of course, whether a particular word equation in a group has a solution in terms of radicals in this sense depends on the group (and the elements $a, b \in G$). However, as the next lemma shows, a radical word always has such a solution. Note that this verifies the implication $(4) \Rightarrow (3)$ in Conjecture 1.4.

Lemma 3.5 *Let w be a radical word, and let G be a uniquely divisible group. For any $a, b \in G$, the equation $w(x, a) = b$ has a solution x which lies in the radical subgroup $\mathcal{R}(\{a, b\}) \subseteq G$.*

The proof of Lemma 1.3 shows that $(1) \Rightarrow (2)$ in Conjecture 1.4. As the implications $(2) \Rightarrow (3) \Rightarrow (4)$ are trivial, the sole unproved implication is $(4) \Rightarrow (1)$.

4 Examples of Uniquely Divisible Groups

In addition to the free uniquely divisible group encountered in the previous section, there are many interesting examples of uniquely divisible groups. We discuss several of them here, although this list is far from exhaustive.

Recall that a *real closed field* is an ordered field K whose positive elements are squares and such that any polynomial of odd degree with coeffients in K has a zero in K. It follows from the definition that each positive element of a real closed field has a positive m-th root for every positive integer m. Moreover, since the field is ordered, this positive root is unique. The group of positive elements of a real closed field is therefore uniquely divisible.

The rest of our examples are noncommutative. The free group F_2 on the alphabet $\{A, B\}$ may be embedded via the Magnus homomorphism ϕ_M [24] into the algebra $\mathbb{Q}\langle\langle a, b\rangle\rangle$ of noncommutative power series via $A \mapsto 1 + a$ and $B \mapsto 1 + b$. The image of this map is a subgroup of the group D of noncommutative power series with constant term 1. Using the binomial series, it can be shown that D is uniquely divisible (see also Proposition 4.1 below for another proof). In particular, F_2 is a subgroup of a uniquely divisible group.

Our next result shows that every word equation with coefficients in D has a unique solution in that set. However, this solution might not be in the radical subgroup $\mathcal{R}(\phi_M(F_2))$ generated by $\phi_M(F_2)$.

Proposition 4.1 *For any $A_1, \ldots, A_m, B \in D$, the equation $\prod_{i=1}^{m}(A_i X) = B$ has a unique solution $X \in D$.*

Let $\psi : \mathcal{F} \to D$ be the homomorphism of uniquely divisible groups with $\psi([A]) = 1 + a$ and $\psi([B]) = 1 + b$ given by Lemma 3.3. Surprisingly, while the Magnus homomorphism $\phi_M : F_2 \to D$ is an embedding of groups, it is a very old open question whether ψ is also an embedding. As far as we know, Baumslag has the best result on this problem [4], giving injectivity when ψ is restricted to certain one-relator subgroups of $\mathcal{F}$.

Our next example is a matrix group. Let K be a field of characteristic 0 and let UT_n be the group of $n \times n$ *unipotent matrices* over K. These are the upper triangular matrices with coefficients in K with 1's along the diagonal.

Proposition 4.2 *For any $A_1, \ldots, A_m, B \in UT_n$, the equation $\prod_{i=1}^{m}(A_i X) = B$ has a unique solution $X \in UT_n$. (In particular, UT_n is a uniquely divisible group.)*

5 The word polynomial

Given a finite word w over the alphabet $\{X, A\}$, write $w = A^{a_0} X A^{a_1} X \cdots A^{a_{n-1}} X A^{a_n}$ for nonnegative integers $a_0, \ldots, a_n$. The *word polynomial* of w is the polynomial in commuting variables x and y given by (3). For example, the word $w = X^{n-1}AX$ has word polynomial: $P_w(x, y) = 1 + x + x^2 + \cdots + x^{n-2} + x^{n-1}y$. Note that if w ends in X (i.e., $a_n = 0$) then w can be uniquely recovered from P_w.

If u is another word over the same alphabet, the composition $u \circ w$ is the word obtained by replacing each occurrence of the letter X in u by the word w. Although composition of words is not commutative, it can be modeled by multiplication of polynomials according to the following lemma.

Lemma 5.1 *Let u and w be finite words in the alphabet $\{X, A\}$ ending with X, and let m, n be respectively the number of letters in w equal to A, X. Then $P_{u \circ w}(x, y) = P_u(x^n y^m, y) P_w(x, y)$.*

Our next lemma shows another context in which the word polynomial P_w arises: from substituting certain affine transformation matrices for the letters of w.

Lemma 5.2 *Let x, y, z be commuting indeterminates, let $w(X, A)$ be a word, and let m, n be respectively the number of letters in w equal to A, X. Then,* $w\left(\begin{bmatrix} x & z \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} y & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} x^n y^m & P_w(x, y) z \\ 0 & 1 \end{bmatrix}$.

6 Proof of Thoerem 1.1

We begin with the following elementary fact.

Lemma 6.1 *Let G be a group of order n. If m and n are relatively prime, then every element of G has a unique m-th root.*

Let G_i $(i = 1, 2, \ldots)$ be an infinite sequence of finite groups with the following property: For every positive integer m, there exists an N such that

$$m \text{ and } \#G_i \text{ are relatively prime for all } i > N. \tag{4}$$

By Lemma 6.1, these groups have a limiting kind of unique divisibility, which suggests taking the quotient of the direct product of the G_i by their direct sum.

Lemma 6.2 *If $G_1, G_2, \ldots$ is a sequence of finite groups satisfying (4), then $G = \prod_{i=1}^{\infty} G_i \Big/ \bigoplus_{i=1}^{\infty} G_i$ is uniquely divisible.*

This lemma allows us to construct many examples of uniquely divisible groups. Next we describe the sequence of groups G_i that we will use to prove Theorem 1.1.

Let p be an odd prime, and let $q = \frac{p-1}{2}$. Since the group $(\mathbb{Z}/p\mathbb{Z})^*$ of units mod p is cyclic of order $p-1$, we can pick an element $t \in (\mathbb{Z}/p\mathbb{Z})^*$ whose multiplicative order mod p is q (namely, t can be the square of any generator). The powers of t are exactly the nonzero squares, i.e. quadratic residues, in $\mathbb{Z}/p\mathbb{Z}$. We take G_p to be the semidirect product $(\mathbb{Z}/q\mathbb{Z}) \ltimes (\mathbb{Z}/p\mathbb{Z})$, which has the presentation

$$G_p = \langle S, T \ : \ S^t T = TS,\ T^q = 1,\ S^p = 1 \rangle .$$

The group G_p can be realized concretely as the group of affine transformations of $\mathbb{Z}/p\mathbb{Z}$ of the form $z \mapsto az + b$, where $a \in (\mathbb{Z}/p\mathbb{Z})^*$ is a quadratic residue and $b \in \mathbb{Z}/p\mathbb{Z}$ is arbitrary. Thus we can view G_p as the group of all 2×2 matrices of the form $\begin{bmatrix} t^k & b \\ 0 & 1 \end{bmatrix}$ where $k \in \mathbb{Z}/q\mathbb{Z}$ and $b \in \mathbb{Z}/p\mathbb{Z}$. The generators S and T correspond to the affine transformations $z \mapsto z+1$ and $z \mapsto tz$, or the matrices: $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $T = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}$.

Lemma 6.3 *Let $\alpha, \beta \in \{0, \ldots, q-1\}$ and $\gamma \in \{0, \ldots, p-1\}$. For any word $w = w(X, A)$, the following identity holds in the group G_p: $w(S^\gamma T^\beta, T^\alpha) = S^{\gamma P_w(t^\beta, t^\alpha)} T^{\alpha m + \beta n}$, where m and n are respectively the number of A's and X's in w.*

Lemma 6.4 *Let w be a finite word in the alphabet $\{X, A\}$, and let n be the number of letters in w equal to X. Let p be a prime such that $q = \frac{p-1}{2}$ is relatively prime to n. If the equation $P_w(x^2, y^2) = 0$ has a solution $(x, y) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$, then there exist $a, b \in G_p$ for which the word equation $w(X, a) = b$ has no solution $X \in G_p$.*

Proof: Suppose $(x, y) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ solves $P_w(x^2, y^2) = 0$. Since any quadratic residue mod p is a power of t, we can find integers α, δ such that $x^2 = t^\delta$ and $y^2 = t^\alpha$. Let $a = T^\alpha$ and $b = ST^{\alpha m + \delta n}$, where m is the number of letters in w equal to A. By Lemma 6.3, an element $X = S^\gamma T^\beta \in G_p$ solves the word equation $w(X, a) = b$ if and only if

$$S^{\gamma P_w(t^\beta, t^\alpha)} T^{\alpha m + \beta n} = b = ST^{\alpha m + \delta n}. \tag{5}$$

Equating powers of T, we obtain $\beta n \equiv \delta n \pmod{q}$. Since n and q are relatively prime, it follows that $\beta \equiv \delta \pmod{q}$, and hence $t^\beta \equiv t^\delta \pmod{p}$. Now equating powers of S in (5) yields

$$1 \equiv \gamma P_w(t^\beta, t^\alpha) \equiv \gamma P_w(t^\delta, t^\alpha) \equiv \gamma P_w(x^2, y^2) \equiv 0 \pmod{p},$$

so there is no solution X to $w(X, a) = b$ in G_p. $\square$

Proof of Theorem 1.1: Let $\pi_0 = 2, \pi_1 = 3, \ldots$ be the primes in increasing order. By the Chinese remainder theorem, for each $i \geq 1$ there is an integer k_i satisfying

$$\begin{aligned} k_i &\equiv 3 \pmod{4} \\ k_i &\equiv 2 \pmod{\pi_j}, \qquad j = 1, \ldots, i. \end{aligned}$$

By Dirichlet's theorem on primes in arithmetic progression, for each i there exists a prime p_i satisfying $p_i \equiv k_i \pmod{4\pi_1 \dots \pi_i}$. By construction, $\frac{p_i-1}{2}$ is not divisible by any of $2, \pi_1, \dots, \pi_i$. Since $\#G_{p_i} = \frac{p_i(p_i-1)}{2}$, the sequence of groups $G_{p_1}, G_{p_2}, \dots$ satisfies condition (4), so by Lemma 6.2 the quotient group $G = \prod_{i\geq 1} G_{p_i} \Big/ \bigoplus_{i\geq 1} G_{p_i}$ is uniquely divisible.

Now let w be a word in the alphabet $\{X, A\}$, and let n be the number of letters in w equal to X. Let π_{i_0} be the largest prime divisor of n. For $i > i_0$ we have $\frac{p_i-1}{2}$ relatively prime to n. By hypothesis, we can choose $i_1 \geq i_0$ sufficiently large so that the equation $P(x^2, y^2) = 0$ has a solution $(x_i, y_i) \in (\mathbb{Z}/p_i\mathbb{Z})^* \times (\mathbb{Z}/p_i\mathbb{Z})^*$ for all $i > i_1$. By Lemma 6.4, for each $i > i_1$, there exist $a_i, b_i \in G_{p_i}$ for which the word equation $w(X, a_i) = b_i$ has no solution $X \in G_{p_i}$. Let $A, B \in G$ be the images of the sequences $(a_i)_{i>i_1}$ and $(b_i)_{i>i_1}$ under the quotient map $\prod_{i>i_1} G_{p_i} \longrightarrow G$. The equation $w(X, A) = B$ has no solution $X \in G$.

Finally, we prove that there exists $B' \in G$ such that $w(X, A) = B'$ has at least two solutions $X \in G$. Using the result just proved, for all $i > i_1$, the map $G_{p_i} \to G_{p_i}$ sending $g \mapsto w(g, a_i)$ is not surjective, hence not injective. Let $g_i, \tilde{g}_i$ be distinct elements of G_{p_i} such that $w(g, a_i) = w(\tilde{g}_i, a_i)$, and let B' be the image in G of the sequence $(w(g_i, a_i))_{i>i_1}$. Then the images in G of the sequences $(g_i)_{i>i_1}$ and $(\tilde{g}_i)_{i>i_1}$ are distinct solutions to $w(X, A) = B'$. □

7 Word equations not solvable by radicals

In this section, we use Corollary 1.7 to find several infinite families of words that are not universal, and consequently not solvable in terms of radicals.

Lemma 7.1 *Let m and n be distinct positive integers, and let $w = X^m A X^n$. Then $P_w(x^2, y^2)$ is irreducible in $\mathbb{C}[x, y]$.*

Proof: We view the word polynomial $P_w(x^2, y^2) = \frac{x^{2m}-1}{x^2-1} + y^2 x^{2m} \frac{x^{2n}-1}{x^2-1}$ as a polynomial in y with coefficients in $\mathbb{C}(x)$. If $m < n$, then there exists $\zeta \in \mathbb{C}$ such that $\zeta^{2m} \neq 1$ and $\zeta^{2n} = 1$, in which case ζ is a simple pole of $f(x) = x^{-2m}(x^{2m} - 1)/(x^{2n} - 1)$; likewise, if $m > n$, then f has a simple root. Thus f is not a square in $\mathbb{C}(x)$, which implies that $P_w(x^2, y^2)$ is irreducible in $\mathbb{C}(x)[y]$. □

By Corollary 1.7, it follows that for positive integers $m \neq n$, the word equation $X^m A X^n = B$ has no solution in terms of radicals. Our next result shows that for $m \geq 0$ and $n \geq 1$, the word equation $XA^{m+2n}XA^{m+n}XA^mX = B$ has no solution in terms of radicals.

Lemma 7.2 *Let $m \geq 0$ and $n \geq 1$ be integers, and let $w = XA^{m+2n}XA^{m+n}XA^mX$. Then $P_w(x^2, y^2)$ has a factor in $\mathbb{Z}[x, y]$ which is irreducible over $\mathbb{C}[x, y]$.*

Next, we show that the word equation $XAX^nAX = B$ has no solution in terms of radicals if $n \geq 3$.

Lemma 7.3 *Let $n \geq 3$ be an integer, and let $w = XAX^nAX$. Then $P_w(x^2, y^2)$ is irreducible in $\mathbb{C}[x, y]$.*

To further extend these families of words not solvable by radicals, note that under the hypotheses of Theorem 1.1, we can actually derive a slightly stronger conclusion.

Corollary 7.4 *Let u, w be finite words in the alphabet $\{X, A\}$. If $P_w(x^2, y^2) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p, then the word equations $u \circ w(X, A) = B$ and $w \circ u(X, A) = B$ have no solution in terms of radicals.*

A simple substitution also proves the following.

Corollary 7.5 *Let $n \geq 2$ be an integer. The word equation $X^2(AX)^n X = B$ is not solvable by radicals.*

8 Acknowledgements

The authors thank George Bergman, Pavel Etingof, Kiran Kedlaya, Igor Klep, Alexei Miasnikov, Bjorn Poonen, and Kate Stange for helpful conversations.

References

[1] M. Aguiar, N. Bergeron, and F. Sottile. Combinatorial Hopf algebras and generalized Dehn-Sommerville equations. *Compos. Math.*, 142:1–30, 2006.

[2] S. Armstrong and C. Hillar. Solvability of symmetric word equations in positive definite letters. *J. London Math. Soc.*, 76:777–796, 2007.

[3] G. Baumslag. Some aspects of groups with unique roots. *Acta Math.*, 10:277–303, 1960.

[4] G. Baumslag. On the residual nilpotence of certain one-relator groups. *Comm. Pure Appl. Math.*, 21:491–506, 1968.

[5] D. Bessis, P. Moussa, and M. Villani. Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. *J. Math. Phys.*, 16:2318–2325, 1975.

[6] G. Birkhoff. On the structure of abstract algebras. *Proc. Cambridge Philos. Soc.*, 31:433–454, 1935.

[7] B. Collins, K. Dykema, and F. Torres-Ayala. Sum-of-squares results for polynomials related to the Bessis-Moussa-Villani conjecture, 2009. `http://arxiv.org/abs/0905.0420`.

[8] S. Gao. Factoring multivariate polynomials via partial differential equations. *Math. Comput.*, 72(242):801–822, 2003.

[9] D. Hagele. Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 127:1167–1171, 2007.

[10] F. Hansen. Trace functions as Laplace transforms. *J. Math. Phys.*, 47:043504, 2006.

[11] C. Hillar. Advances on the Bessis-Moussa-Villani trace conjecture. *Lin. Alg. Appl.*, 426:130–142, 2007.

[12] C. Hillar and C. R. Johnson. Positive eigenvalues of generalized words in two Hermitian positive definite matrices. In P. Pardalos and H. Wolkowicz, editors, *Novel Approaches to Hard Discrete Optimization*, volume 37 of *Fields Institute Communications*, pages 111–122, 2003.

[13] C. Hillar and C. R. Johnson. Symmetric word equations in two positive definite letters. *Proc. Amer. Math. Soc.*, 132:945–953, 2004.

[14] C. R. Johnson and C. Hillar. Eigenvalues of words in two positive definite letters. *SIAM J. Matrix Anal. Appl.*, 23:916–928, 2002.

[15] O. Kharlampovich and A. Myasnikov. Equations in a free $\mathbf{Q}$-group. *Trans. Amer. Math. Soc.*, 350(3):947–974, 1998.

[16] O. Kharlampovich and A. Myasnikov. Elementary theory of free non-abelian groups. *J. Algebra*, 302(2):451–552, 2006.

[17] I. Klep and M. Schweighofer. Connes' embedding conjecture and sums of Hermitian squares. *Adv. Math.*, 217:1816–1837, 2008.

[18] I. Klep and M. Schweighofer. Sums of Hermitian squares and the BMV conjecture. *J. Stat. Phys.*, 133:739–760, 2008.

[19] P. Landweber and E. Speer. On D. Hagele's approach to the Bessis-Moussa-Villani conjecture. *Lin. Alg. Appl.*, 431:1317–1324, 2009.

[20] J. Lawson and Y. Lim. Solving symmetric matrix word equations via symmetric space machinery. *Lin. Alg. Appl.*, 414:560–569, 2006.

[21] K. J. Le Couteur. Representation of the function Tr$(\exp(A - \lambda B))$ as a Laplace transform with positive weight and some matrix inequalities. *J. Phys. A: Math. Gen.*, 13:3147–3159, 1980.

[22] J. Ledlie. Representations of free metabelian $\mathcal{D}_\pi$-groups. *Trans. Amer. Math. Soc.*, 153:307–346, 1971.

[23] E. H. Lieb and R. Seiringer. Equivalent forms of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 115:185–190, 2004.

[24] W. Magnus. On a theorem of Marshall Hall. *Ann. of Math.*, pages 764–768, 1939.

[25] D. Marker. *Model theory: an introduction*, volume 217 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2010.

[26] P. Moussa. On the representation of tr $\left(e^{A-\lambda B}\right)$ as a Laplace transform. *Rev. Math. Phys.*, 12:621–655, 2000.

[27] Z. Sela. Diophantine geometry over groups. VI. The elementary theory of a free group. *Geom. Funct. Anal.*, 16(3):707–730, 2006.