Laboratoire de l'Informatique du Parallélisme
École Normale Supérieure de Lyon
Unité Mixte de Recherche CNRS-INRIA-ENS LYON
n° 5668

Quantum Automata and Algebraic Groups

Harm Derksen ,
Emmanuel Jeandel ,
Pascal Koiran

23rd July 2003

Research Report N° 2003-39

École Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : +33(0)4.72.72.80.37 Fax : +33(0)4.72.72.80.80 Adresse électronique : `lip@ens-lyon.fr`

CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Quantum Automata and Algebraic Groups

Harm Derksen ,
Emmanuel Jeandel and Pascal Koiran

23rd July 2003

Abstract

We show that several problems which are known to be undecidable for probabilistic automata become decidable for quantum finite automata. Our main tool is an algebraic result of independent interest: we give an algorithm which, given a finite number of invertible matrices, computes the Zariski closure of the group generated by these matrices.

Keywords: quantum automata, probabilistic automata, undecidability, algebraic groups, algebraic geometry

Résumé

Nous montrons ici que plusieurs problèmes indécidables pour des automates probabilistes sont décidables pour des automates quantiques. Ce rsultat s'appuie sur un algorithme intéressant en soi, qui, étant donné des matrices inversibles, calcule la cloture de Zariski du groupe engendré par ses matrices.

Mots-clés: Automates quantiques, automates probabilistes, indécidabilité, groupes algébriques, géométrie algébrique

QUANTUM AUTOMATA AND ALGEBRAIC GROUPS

HARM DERKSEN, EMMANUEL JEANDEL, AND PASCAL KOIRAN

ABSTRACT. We show that several problems which are known to be undecidable for probabilistic automata become decidable for quantum finite automata. Our main tool is an algebraic result of independent interest: we give an algorithm which, given a finite number of invertible matrices, computes the Zariski closure of the group generated by these matrices.

1. INTRODUCTION

The development of the theory of computation has led to the study of various models of computation, e.g., finite automata, boolean circuits, Turing machines, cellular automata... Due to the recent interest in quantum computation, quantum counterparts of the main classical models of computation (including the four models listed above) have been defined. It is especially fruitful to compare these models to their probabilistic counterparts. The best known result in this direction is Shor's quantum factoring algorithm, which runs in polynomial time despite the fact – or rather the belief – that no classical algorithm, deterministic or probabilistic, can factor integers in polynomial time.

In this paper we show that several problems which are known to be undecidable for probabilistic automata become decidable for quantum finite automata. We work with the "measure once" model of quantum automata of Moore and Crutchfield [20]. The main other model is the "measure many" model of Kondacs and Watrous [17]. Further comparisons between probabilistic and quantum automata can be found in [5]. The main focus of these three papers is the study of the languages recognized by quantum finite automata. Our main tool is an algebraic result of independent interest: we give an algorithm which, given a finite number of invertible matrices, computes the Zariski closure of the group generated by these matrices. The problem of finding the Zariski closure of matrix groups also appears naturally in other areas. For example, it is well-known the the Zariski closure of the monodromy group of a Fuchsian system of differential equations is the differential Galois group (see [7] for an introduction to differential Galois theory).

Harm Derksen is partially supported by NSF, grant DMS 0102193.

2. Probabilistic and Quantum Automata

In this section we recall the definitions of probabilistic and quantum automata, and obtain our decidability results. The remainder of the paper is devoted to our group-theoretic algorithm.

2.1. **Probabilistic Automata.**

Formally, a probabilistic automaton is a tuple $\mathcal{A} = (Q, q_0, Q_f, \Sigma, (X_a)_{a\in\Sigma})$ where $Q = \{1, \ldots, q\}$ is a finite set of states, $q_0 \in Q$ is the initial state, $Q_f \subseteq Q$ is the set of final states, and Σ is a finite alphabet. Each matrix X_a is a $q \times q$ stochastic matrix: $(X_a)_{ij}$ is the probability of going from state i to state j when a is the input letter. For instance, if the rows of all X_a contain exactly one 1 (and $q-1$ zeros) we recover the familiar model of deterministic finite automata. Another degenerate case is obtained when $|\Sigma| = 1$. In this case, our probabilistic automaton is essentially a finite state Markov chain.

In order to define the language accepted by a probabilistic automaton, we need to fix a threshold $\lambda \in [0, 1]$. A word $w = w_1 \ldots w_n \in \Sigma^*$ is accepted if the probability of ending up in Q_f upon reading w is at least λ. This condition can be conveniently expressed in a matrix formalism. Let π be the column vector of size q such that $\pi_i = 1$ if $i = q_0$ and $\pi_i = 0$ otherwise. Let η be the column vector of size q such that $\eta_i = 1$ if $i \in Q_f$ and $\eta_i = 0$ otherwise. Finally, let $ACC_w = \pi^T X_w \eta$ where $X_w = X_{w_1} \cdots X_{w_n}$. The word w is accepted if $ACC_w > \lambda$. Note that the row vector $\pi^T X_w$ can be interpreted as a probability distribution over Q.

It turns out that one cannot decide whether the set of accepted words is empty, even if λ and the entries of the X_a are rational numbers. In fact, the following problems are all undecidable [8, 22].

(1) Is there $w \in \Sigma^*$ such that $ACC_w \geq \lambda$?
(2) Is there $w \in \Sigma^*$ such that $ACC_w \leq \lambda$?
(3) Is there $w \in \Sigma^*$ such that $ACC_w = \lambda$?
(4) Is there $w \in \Sigma^*$ such that $ACC_w > \lambda$?
(5) Is there $w \in \Sigma^*$ such that $ACC_w < \lambda$?

A threshold λ is said to be isolated if there exists $\epsilon > 0$ such that

$$|ACC_w - \lambda| \geq \epsilon$$

for every $w \in \Sigma^*$. This definition is motivated in particular by the fact that probabilistic automata with isolated thresholds accept exactly the same languages as deterministic finite automata [23]. Unfortunately, the following two basic problems are undecidable [4, 6, 8].

(6) Given $\mathcal{A}$ and λ, decide whether λ is isolated.
(7) Given $\mathcal{A}$, decide whether there exists a threshold λ which is isolated.

2.2. **Quantum Automata.**

After reading a word w, a probabilistic automaton is in a probability distribution of the form $\sum_{i\in Q} \alpha_i e_i$ where $(e_1, \ldots, e_n)$ is the canonical basis of $\mathbb{R}^q$. In quantum automata this probability distribution is replaced by a *superposition* $\sum_{i\in Q} \alpha_i e_i$ of unit ℓ^2 norm.

Instead of stochastic matrices we must therefore work with matrices X_a which conserve the norm, i.e., with orthogonal matrices. More generally, one could allow matrices with complex coefficients (i.e., unitary matrices) but we shall stick to orthogonal matrices throughout the paper. The definition of ACC_w is changed accordingly: in a quantum automaton the probability of accepting a word w is $ACC_w = ||\pi^T X_w P||^2$ where P is the matrix of orthogonal projection on the subspace spanned by the final states (hence $P_{ii} = 1$ if $i \in Q_f$, and the other entries of P are null). The other definitions are unchanged.

Problems 1 through 7 clearly make sense for quantum automata. The first three problems remain undecidable [13]. As far as quantum automata are concerned, the main result of this paper is that the last four problems become decidable.

Theorem 1. *Problems (4) through (7) are decidable for quantum automata.*

For this theorem to make sense, one must explain how the entries of the matrices X_a are finitely represented. One may for instance assume that they are algebraic numbers, which can be represented by their minimal polynomial and an isolating interval. More general solutions are possible: see Remark 1 at the end of this section.

Note also that there is nothing quantum about our decision algorithms: they are classical algorithms about a quantum model of computation. The decidability of problems (4) and (5) has also been obtained in [13] by a slightly different method. It is known that problems (1) through (5) are undecidable for the measure-many model [13], but the status of problems (6) and (7) is unknown.

Let $\langle X_a \rangle_{a \in \Sigma}$ be the group generated by the matrices X_a, and let $G(\mathcal{A})$ be the closure (for the Euclidean topology on $\mathbb{R}^{q^2}$) of this group. Thus $G(\mathcal{A})$ is a compact group of orthogonal matrices. This group plays a crucial role in our proofs. We now illustrate this point on problem 6. First, we need an easy lemma.

Lemma 1. *The group $G(\mathcal{A})$ is equal to the closure of the monoid generated by the matrices X_a.*

Proof. The inclusion from right to left is obvious. For the converse, note that there exists (by compactness) for each matrix X_a a sequence $(n_k)_{k \geq 1}$ such that $X_a^{n_k}$ converges to the identity matrix as $k \to +\infty$.

Hence $X_a^{-1} = \lim_{k \to +\infty} X_a^{n_k - 1}$ and the result follows. □

Proposition 1. *The two following properties are equivalent.*

(i) *the threshold λ is isolated.*
(ii) *There exists $\epsilon > 0$ such that $|\ ||\pi^T g P||^2 - \lambda\ | \geq \epsilon$ for every $g \in G(\mathcal{A})$.*

Proof. By Lemma 1, the set $\{||\pi^T g P||^2;\ g \in G(\mathcal{A})\} \subseteq [0, 1]$ is the closure of $\{ACC(w);\ w \in \Sigma^*\}$. □

Instead of checking property (i) directly, we may therefore check property (ii). It is not immediately clear how this can be done algorithmically. Here, two miracles happen. The first miracle is that the group $G(\mathcal{A})$ is algebraic (in other words, the Euclidean closure of $\langle X_a \rangle_{a\in\Sigma}$ is equal to its Zariski closure). This follows from the general fact that a compact group of real matrices is algebraic [21]. The second miracle is that there is an algorithm – presented in the next section – which from the matrices X_a computes a system of polynomial equations defining the Zariski closure of $\langle X_a \rangle_{a\in\Sigma}$. Checking (ii) then amounts to deciding whether a first-order sentence of the language of ordered fields is true in the field of real numbers. It has been known since Tarski that this can be done algorithmically (more efficient algorithms and further references can be found in [2] or [24]).

The algorithms for problems 4, 5 and 7 are almost identical. We leave it to the reader to write down the corresponding first-order sentences.

Remark 1. *As mentioned above, Theorem 1 applies to matrices X_a with entries in a field $K \subset \mathbb{R}$ bigger than the field of real algebraic numbers. For instance we may give a (finite) transcendence basis $\mathcal{B}$ of K, and represent the entries of X_a as algebraic numbers over $\mathcal{B}$. This purely algebraic information is sufficient to compute the group $G(\mathcal{A})$. Once the group is computed we have to decide a first-order sentence of the field of real numbers, and we therefore have to compute the sign of polynomial functions of the elements of $\mathcal{B}$. In order to do this we only need to assume that we have access to an algorithm which for any element x of $\mathcal{B}$ and any $\epsilon > 0$ computes a rational number q such that $|x - q| < \epsilon$. We use the algebraic information to determine whether a polynomial takes the value zero, and if not we use approximations to determine its sign.*

3. Algebraic Groups

Let K be a field and let $\overline{K}$ be its algebraic closure. Suppose that $\{X_1, X_2, \ldots, X_k\} \subset \mathrm{M}_n(K)$ is a finite set of invertible $n \times n$ matrices. Let $G = \langle X_1, X_2, \ldots, X_k \rangle$ be the subgroup of $\mathrm{GL}_n(\overline{K})$ generated by $X_1, X_2, \ldots, X_k$. In this section we will present an algorithm to compute $\overline{G}$, the Zariski closure of G in $\mathrm{GL}_n(\overline{K})$. For the applications to quantum automata we may assume that $X_1, \ldots, X_k$ are unitary. It is therefore possible for a reader interested primarily in quantum automata to skip cases 1 of section 3.2 and the case of unipotent matrices in section 3.3. If the entries of the matrices $X_1, \ldots, X_k$ are algebraic numbers, one may also skip case 3 of section 3.2.

The ring of polynomial functions on $\mathrm{GL}_n(\overline{K})$ is generated by the coordinate functions $x_{i,j}$, $1 \leq i, j \leq n$ and the function $x_0 = 1/\det(x_{i,j})$. The coordinate ring of $\mathrm{GL}_n(\overline{K})$ can therefore be identified with

$$R_{\overline{K}} = \overline{K}[x_{1,1}, x_{1,2}, \ldots, x_{n,n}, x_0]/(\det(x_{i,j})x_0 - 1).$$

To compute the Zariski closure of G means that we have to find generators of the ideal

$$I_{\overline{K}} = \{f \in R_{\overline{K}} \mid f(X) = 0 \text{ for all } X \in G\}.$$

Since G is a subgroup of $\mathrm{GL}_n(K) \subset \mathrm{GL}_n(\overline{K})$, $I_{\overline{K}}$ is generated by polynomials in

$$R_K = K[x_{1,1}, x_{1,2}, \dots, x_{n,n}, x_0]/(\det(x_{i,j})x_0 - 1).$$

If we define

$$I_K = \{f \in R_K \mid f(X) = 0 \text{ for all } X \in G\}$$

then $I_{\overline{K}}$ will be just the ideal in $R_{\overline{K}}$ generated by I_K. We will discuss an algorithm that produces a finite number of generators $f_1, f_2, \dots, f_r$ of the ideal I_K. If $X \in \mathrm{GL}_n(\overline{K})$ then it is easy to check whether $X \in \overline{G}$, namely

$$X \in \overline{G} \Leftrightarrow f_1(X) = f_2(X) = \cdots = f_r(X) = 0.$$

Without loss of generality we may assume that the field K is finitely generated as a field over $\mathbb{Q}$ or over a finite field. In fact, we may take K as the smallest field that contains all entries of all matrices in $\{X_a\}_{a\in\Sigma}$. After some preparation, we will first discuss the case where G is generated by only one matrix. This then will be used to describe an algorithm for the Zariski closure of a matrix group with an arbitrary (finite) number of generators.

3.1. **Gröbner bases techniques.** We briefly summarize the main results we will need from Gröbner bases theory. We assume that the field K is finitely generated (as a field) over $\mathbb{F}_p$ for some prime number or over $\mathbb{Q}$.

Suppose that A and B are affine varieties over a field K, and $\psi : A \to B$ is a morphism of affine varieties. If $H \subset A$ is a Zariski closed subset, then one can compute $\overline{\psi(H)}$, the Zariski closure of the image, using Gröbner basis elimination techniques. The morphism $\psi : A \to B$ corresponds to a ring homomorphism $\psi^\star : K[B] \to K[A]$ of the coordinate rings. Given the generators of the vanishing ideal $\mathfrak{h} \subset K[A]$ of H, one can compute generators of the ideal $(\psi^\star)^{-1}(\mathfrak{h})$ which is the vanishing ideal of $\overline{\psi(H)}$.

One situation where we will apply this is the following. Suppose that A and B are Zariski closed subsets of $\mathrm{GL}_n(\overline{K})$. Let $\overline{A \cdot B}$ be the Zariski closure of

$$A \cdot B = \{XY \mid X \in A, Y \in B\}.$$

Since the multiplication map $m : \mathrm{GL}_n(\overline{K}) \times \mathrm{GL}_n(\overline{K}) \to \mathrm{GL}_n(\overline{K})$ is a morphism of affine varieties, we will be able to compute

$$\overline{A \cdot B} = \overline{m(A \times B)}.$$

If S is a ring of finite type over K, and $\mathfrak{a}$ is an ideal in S given by its generators, then generators of the radical ideal $\sqrt{\mathfrak{a}}$ can be computed. The first radical ideal algorithms assumed characteristic 0. However, in recent publications algorithms have been suggested that compute radical ideals over base fields which are finitely generated over a finite field or over $\mathbb{Q}$ (see [11, 16, 19]).

This can be applied to compute the integral closure of S if S is a domain using De Jong's algorithm (see [14, 9]). Following Becker and Weispfenning, one can also compute the primary decomposition of an ideal $\mathfrak{a}$ of S (see [3]). We emphasize that these algorithms (using the radical ideal algorithms mentioned before) will work over any field K as general as our assumptions.

If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in a ring S of finite type over K, then the colon ideal

$$(\mathfrak{a} : \mathfrak{b}) = \{f \in S \mid f\mathfrak{b} \subseteq \mathfrak{a}\}$$

can also be computed with Gröbner basis techniques (see for example [9, 1.2.4]).

3.2. **Finding multiplicative relations.** Let K be a field that is finitely generated over the rational numbers or over a finite field. Suppose that $\lambda_1, \lambda_2, \dots, \lambda_n \in K^\star$. Consider the group homomorphism $\varphi : \mathbb{Z}^n \to K^\star$ defined by

$$\varphi(a_1, a_2, \dots, a_n) = \lambda_1^{a_1} \lambda_2^{a_2} \cdots \lambda_n^{a_n}.$$

We will discuss an algorithm that finds generators of the kernel of φ. We distinguish three cases.

case 1: *K is a finite field.* The field $K^\star$ is then a finite cyclic group. It is elementary to compute the kernel of a homomorphism between two finitely generated abelian groups.

case 2: *K is a number field*, a finite algebraic extension of $\mathbb{Q}$ of degree d. For $a = (a_1, a_2, \dots, a_n) \in \mathbb{Z}^n$, we define

$$|a| = \max\{|a_1|, |a_2|, \dots, |a_n|\}.$$

Recall that an absolute value $|.|_\nu$ is said to be normalized if:

- $|x|_\nu = x$ if $x \in \mathbb{Q}$, $x > 0$ and $|.|_\nu$ is archimedean.
- $|p|_\nu = 1/p$ if $|.|_\nu$ is p-adic.

The other absolute values are obtained by multiplication by a constant. In the following we only consider normalized absolute values (for all these matters we refer the reader to [29]). The height $h(\lambda)$ of λ is defined by

$$h(\lambda) = \frac{1}{d} \sum_v \max\{\log |\lambda|_v, 0\},$$

where the sum extends over all normalized absolute values on K. For $\lambda \in K$ we have $h(\lambda) = 0$ if and only if λ is a root of unity.

One approach to find the kernel of φ is to observe that

$$\lambda_1^{a_1} \lambda_2^{a_2} \cdots \lambda_n^{a_n}$$

is a root of unity if and only if

$$a_1 \log |\lambda_1|_v + a_2 \log |\lambda_2|_v + \cdots + a_n \log |\lambda_n|_v = 0$$

for all absolute values. We will not work out the details here. Instead we will give an explicit bound by Masser. From this bound it is clear that the generators of the kernel of φ can be found constructively.

We define η to be the infimum of $h(\lambda)$ over all $\lambda \in K$ that are not roots of unity. We also define ω to be the largest integer d such that K contains a d-th root of unity. We also define

$$h = \max\{h(\lambda_1), h(\lambda_2), \ldots, h(\lambda_n), \eta\}.$$

Theorem 2 (Masser, [18]). *The kernel of φ is generated by elements $a \in \mathbb{Z}^n$ with*

$$|a| \leq n^{n-1}\omega(h/\eta)^{n-1}.$$

We still have to show that all the constants in the inequality can be effectively computed. If K contains an ω-th root of unity then the degree of the extension $K : \mathbb{Q}$ must be at least $\phi(\omega)$ where ϕ is Euler's function. From this follows that one can easily bound ω in terms of the degree of the extension d.

Estimating η is more difficult. If α is not an algebraic integer, then $h(\alpha) \geq (\log 2)/d$ because $|\alpha|_v \geq 2$ for some valuation v. Lower bounds on the height of algebraic integers are not so easily obtained, and several bounds have been proposed in the literature ([29], section 3.6). For our purpose any effective bound will do, for instance the recent bound

$$h(\alpha) \geq \frac{1}{4d}\left(\frac{\log\log d}{\log d}\right)^3$$

due to Voutier [26].

case 3: *K has transcendental elements.* The field K contains a field F where $F = \mathbb{Q}$ or F is the finite field $\mathbb{F}_p$ for some prime number p. Note that F is a perfect field. Let t be an indeterminate and consider the ring

$$S = F[\lambda_1 t, \lambda_2 t, \ldots, \lambda_n t, t] \subseteq K[t].$$

The reason that we consider this ring S is that the quotient field of S contains the elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ and that the only invertible elements of S are constant functions (as we will prove below). This allows us then to think geometrically. We would like to view $\lambda_1, \ldots, \lambda_n$ as divisors on the affine variety corresponding to S. In order to do so, we need S to be integrally closed. With De Jong's algorithm we can compute the integral closure $\widetilde{S}$ of S (see [14]). This algorithm works for any domain of finite type over a perfect field (see [9, §1.6, §1.5]). Since $K[t]$ is integrally closed, $\widetilde{S}$ is contained in $K[t]$. Let L be the integral closure of F within S. It follows from [28, Theorem 6.7.3] that the intersection of $\widetilde{S}$ and K is equal to L. Now L is again a field, and L is a finite algebraic extension of F. This means that L is a number field, or L is a finite field. We have that $\widetilde{S}^\star$, the set of invertible elements in $\widetilde{S}$ is equal to $L^\star$.

Divisors on $\mathrm{Spec}(\widetilde{S})$ correspond to height 1 prime ideals in $\widetilde{S}$. For every $\mathfrak{p}$ we denote its zero set (which is a divisor) by $D_\mathfrak{p}$. Whenever $\mathfrak{p}$ is a height one prime ideal, the localization $\widetilde{S}_\mathfrak{p}$ is a discrete valuation ring (see [10, Theorem 11.5]). We have a valuation $v_\mathfrak{p}$ on the quotient field of $\widetilde{S}$ such $v_\mathfrak{p}(f) \geq 0$

if and only if $f \in S_{\mathfrak{p}}$. The valuation $v_{\mathfrak{p}}$ is normalized such that $v_{\mathfrak{p}}$ reaches exactly all values in $\mathbb{Z}$. For any f in the quotient field of $\widetilde{S}$, we define its Weil divisor as the formal sum

$$\operatorname{div}(f) = \sum_{\mathfrak{p}} v_{\mathfrak{p}}(f)[D_{\mathfrak{p}}],$$

where $\mathfrak{p}$ runs over all height one prime ideals. Let $\operatorname{Div}(\widetilde{S})$ be the group of Weil divisors on $\widetilde{S}$. For any rational function f, $\operatorname{div}(f) = 0$ if and only if $f \in \widetilde{S}^\star = L^\star$ (because $\widetilde{S}$ is the intersection of all localizations of height 1 prime ideals, see [10, Corollary 11.4]).

We have a natural homomorphism of abelian groups

$$\widetilde{\varphi} : \mathbb{Z}^n \to \operatorname{Div}(\widetilde{S})$$

defined by

$$(b_1, \dots, b_n) \mapsto b_1 \operatorname{div}(\lambda_1 t) + b_2 \operatorname{div}(\lambda_2 t) + \cdots + b_n \operatorname{div}(\lambda_n t) - (b_1 + b_2 + \cdots + b_n) \operatorname{div}(t).$$

We have that

$$\widetilde{\varphi}(b_1, \dots, b_n) = 0 \Leftrightarrow \lambda_1^{b_1} \cdots \lambda_n^{b_n} \in \widetilde{S}^\star = L^\star$$

Generators of the kernel of $\widetilde{\varphi}$ can be computed as follows. Let $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ be all the height 1 prime ideals corresponding to the divisors appearing in $\operatorname{div}(\lambda_1 t), \dots, \operatorname{div}(\lambda_n t), \operatorname{div}(t)$. These prime ideals can be found by computing the primary decompositions of the ideals $(\lambda_1 t), \dots, (\lambda_n t), (t)$. We will write v_i instead of $v_{\mathfrak{p}_i}$. If $f \in \widetilde{S}$, then $v_i(f)$ can be computed because

$$\begin{aligned}
v_i(f) \geq r \quad &\Leftrightarrow \quad \mathfrak{p}_i^r \widetilde{S}_{\mathfrak{p}_i} \subseteq f \widetilde{S}_{\mathfrak{p}_i} \\
&\Leftrightarrow \quad g \mathfrak{p}_i^r \subseteq (f) \text{ for some } g \in \widetilde{S} \setminus \mathfrak{p}_i \\
&\Leftrightarrow \quad ((f) : \mathfrak{p}_i^r) \not\subseteq \mathfrak{p}_i.
\end{aligned}$$

In particular, we can compute all $v_i(\lambda_j t)$ and all $v_i(t)$ for all i and j.

Note that $v_i(\lambda_j) = v_i(\lambda_j t) - v_i(t)$. Now $\widetilde{\varphi}(a_1, a_2, \dots, a_n) = 0$ if and only if

$$a_1 v_i(\lambda_1) + a_2 v_i(\lambda_2) + \cdots + a_n v_i(\lambda_n) = 0$$

for $i = 1, 2, \dots, r$. We can solve these equations and we find generators of $\ker(\widetilde{\varphi})$. Let $a^{(1)}, a^{(2)}, \dots, a^{(s)}$ be generators of $\ker(\widetilde{\varphi})$. The kernel of φ is contained in the kernel of $\widetilde{\varphi}$. To find generators of $\ker(\varphi)$ we proceed as follows. Let $\mu_i = \varphi(a^{(i)}) \in L^\star$, $i = 1, 2, \dots, s$. Then

$$\varphi(b_1 a^{(1)} + \cdots + b_s a^{(s)}) = \mu_1^{b_1} \cdots \mu_s^{b_s}. \tag{1}$$

We already have seen how to compute generators of the module of all $(b_1, b_2, \dots, b_s) \in \mathbb{Z}^s$ such that the righthandside of (1) is equal to 1. This then gives us explicit generators of the kernel of φ.

3.3. **Zariski closure of cyclic groups.** We will now discuss how one can compute the Zariski closure of a group generated by a single invertible matrix $X \in \mathrm{M}_n(K)$. Using linear algebra, one can find a matrix $Y \in \mathrm{GL}_n(K)$ such that YXY^{-1} is in Jordan normal form. (We may have to replace K by an finite algebraic extension of itself.) Without loss of generality we may assume that X is in Jordan normal form. We can effectively write down the multiplicative Jordan decomposition

$$X = X_s X_u$$

where X_s is semisimple and X_u is unipotent. In fact X_s is just the diagonal part of X, and X_n is equal to $X_s^{-1}X$. Since X_s and X_u commute, we have that

$$\overline{\langle X\rangle} = \overline{\overline{\langle X_s\rangle} \cdot \overline{\langle X_n\rangle}}.$$

Because of the previous section, this reduces the problem to computing the Zariski closure of $\langle X\rangle$ where X is either semisimple or unipotent.

Suppose now that X is a unipotent matrix. If the characteristic of the field K is positive, then X will have finite order. In that case $\langle X\rangle$ is equal to its Zariski closure and it easily can be computed. Let us assume for a moment that the characteristic of K is equal to 0. Define Z by

$$Z = \log(X) = \sum_{i=1}^{\infty} (-1)^{i-1} \frac{(X-I)^i}{i}.$$

Note that the infinite sum actually only runs up to $i = n-1$ since X is unipotent. The matrix Z is nilpotent. Define $\varphi : \overline{K} \to \mathrm{GL}_n(\overline{K})$ by

$$t \mapsto \exp(tZ) = \sum_{i=0}^{\infty} \frac{t^i Z^i}{i!} = \sum_{i=0}^{n-1} \frac{t^i Z^i}{i!}.$$

For any integer k we have $\varphi(k) = X^k$. Since the integers are Zariski dense in K, we see that the Zariski closure of $\langle X\rangle$ is the Zariski closure of the image of φ. Again the Zariski closure of the image of φ can be computed using a Gröbner basis elimination.

Assume that X is diagonal, say

$$X = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

(and K can again be of arbitrary characteristic). The group of diagonal matrices is isomorphic to $T = (\overline{K}^\star)^n$. The coordinate ring of T (over K) is isomorphic to the ring of Laurent polynomials

$$U = K[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}].$$

The ideal I of the Zariski closure of $\langle X\rangle$ is generated by all $f \in L$ such that

$$f(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k) = 0$$

for all $k \in \mathbb{Z}$. Define (as in the previous subsection) a group homomorphism $\varphi : \mathbb{Z}^n \to K^\star$ by

$$\varphi(a_1, a_2, \ldots, a_n) = \lambda_1^{a_1} \lambda_2^{a_2} \cdots \lambda_n^{a_n}.$$

Let J be the ideal of all

$$x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} - 1$$

with $(a_1, a_2, \ldots, a_n) \in \ker(\varphi)$. If $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{Z}^n$, then we have

$$x_1^{a_1+b_1} x_2^{a_2+b_2} \cdots x_n^{a_n+b_n} = x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} (x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} - 1) + (x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} - 1) \in$$

$$(x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} - 1, x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} - 1).$$

From this it is easy to see that if S is a set of generators of $\ker(\varphi)$, then then the ideal J is generated by all

$$x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} - 1$$

with $(a_1, a_2, \ldots, a_n) \in S$. In the previous subsection we have seen how to find a set of generators of the kernel of φ. This gives us a way to find generators of the ideal J. With the lemma below, we obtain in this way a set of generators of the ideal I, the vanishing ideal of the Zariski closure of $\langle X \rangle$.

Lemma 2. *We have $J = I$.*

Proof. Clearly $J \subseteq I$. If $J \neq I$ then one can choose $f \in I \setminus J$ such that

$$f = \sum_{i=1}^{r} b_i m_i$$

with $b_1, b_2, \ldots, b_r \in K$ and $m_1, \ldots, m_r$ Laurent monomials. Choose f such that r is minimal. Let $\mu_i = m_i(\lambda_1, \ldots, \lambda_n)$. Note that $\mu_i \neq \mu_j$ for $i \neq j$, because otherwise $m_i m_j^{-1} - 1 \in J$ and $f - b_i m_j (m_i m_j^{-1} - 1) \in I \setminus J$ would have fewer terms than f. Now

$$0 = f(\lambda_1^k, \ldots, \lambda_n^k) = \sum_{i=1}^{r} b_i \mu_i^k$$

for all k. Since the vectors

$$\begin{pmatrix} 1 \\ \mu_1 \\ \mu_1^2 \\ \vdots \\ \mu_1^{r-1} \end{pmatrix}, \begin{pmatrix} 1 \\ \mu_2 \\ \mu_2^2 \\ \vdots \\ \mu_2^{r-1} \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \mu_r \\ \mu_r^2 \\ \vdots \\ \mu_r^{r-1} \end{pmatrix}$$

are linearly independent, it follows that $b_1 = b_2 = \cdots = b_r = 0$ which leads to a contradiction. $\square$

3.4. **An algorithm for the Zariski closure of matrix groups.** We are now able to present the algorithm which computes the Zariski closure of the group generated by given $n \times n$ invertible matrices $X_1, X_2, \dots, X_k$.

Algorithm 1.

1. `input:` matrices $X_1, X_2, \dots, X_k \in \mathrm{GL}_n(K)$.
2. $H := \{I\}$
3. $S := \{I, X_1, X_2, \dots, X_k\}$
4. `repeat`
5. $\quad H' := H$
6. $\quad S' := S$
7. $\quad$ `for` Y `in` S `do`
8. $\quad\quad H := \overline{H \cdot \overline{\langle Y \rangle}_0}$
9. $\quad\quad H := \overline{H \cdot YHY^{-1}}$
10. $\quad\quad G = S \cdot H$
11. $\quad\quad$ `for` Z `in` S `do`
12. $\quad\quad\quad$ `if` $YZ \notin G$ `then` $S := S \cup \{YZ\}$
13. `until` $H' = H$ `and` $S' = S$
14. `output:` G

Throughout the algorithm G and H are Zariski closed subsets of $\mathrm{GL}_n(K)$, and S is a finite subset of $\mathrm{GL}_n(K)$. The reader should be aware that G and H are represented by an ideal in the coordinate ring of $\mathrm{GL}_n(K)$ throughout the algorithm. We clarify some of the steps.

line 8: Here we compute the Zariski closure $\overline{\langle B \rangle}$ of the group $\langle Y \rangle$ generated by the matrix Y as discussed in section 3.3. Using an algorithm for primary decomposition, we can find the connected component of the identity in $\overline{\langle Y \rangle}$. This component is denoted by $\overline{\langle Y \rangle}_0$. We compute the Zariski closure of the product of H and $\overline{\langle Y \rangle}_0$ and assign it to H.

line 9: We conjugate H with Y. Conjugation with Y induces an automorphism of GL_n and also an automorphism of the coordinate ring of GL_n. If we apply conjugation with Y^{-1} to the vanishing ideal of H, then we get the vanishing ideal of YHY^{-1}. We compute the Zariski closure of the product of H and YHY^{-1} and assign it to H.

line 10: G is a finite union of cosets of H. For each coset of H we can compute the vanishing ideal since left multiplication in GL_n induces an automorphism of the coordinate ring of GL_n. Then the vanishing ideal of G is the intersection of the vanishing ideals of all cosets. This can be computed using Gröbner basis techniques.

Let $\widetilde{G}$ be the Zariski closure of the group generated by $X_1, X_2, \dots, X_k$. Our goal is to prove that the algorithm terminates and that the output is $\widetilde{G}$. In order to do this, we first give various invariants.

Lemma 3. *Throughout the algorithm we have*

(a) *H is an irreducible variety containing the identity I.*
(b) *$S \cdot H$ contains $I, X_1, X_2, \dots, X_k$.*
(c) *$S \cdot H$ is contained in the Zariski closure $\widetilde{G}$ of $\langle X_1, X_2, \dots, X_k \rangle$.*

Proof. (a) If A and B are irreducible, then so is $\overline{A \cdot B}$ (since it is the Zariski closure of the image of an irreducible variety under a morphism). Note that if $B \in \mathrm{GL}_n$ then $\overline{\langle B \rangle}$ is an algebraic group and $\overline{\langle B \rangle}_0$ is a connected algebraic group. Any connected algebraic group is always irreducible. At the beginning in line 2, H is irreducible. Throughout the algorithm H remains irreducible, since it remains irreducible in lines 8 and 9.

(b) After execution of line 3 we have that $S \cdot H$ contains $I, X_1, X_2, \dots, X_k$. Throughout the algorithm S and H never get smaller.

(c) This is certainly true after execution of line 3. It is easy to check that after execution of line 8, 9 or 12, $S \cdot H$ remains to be contained in the Zariski closure of $\langle X_1, X_2, \dots, X_k \rangle$.

□

Lemma 4. *In each iteration of the* `repeat-until` *loop just before the execution of line 13, the following statements are true:*

(a) *For every $Y, Z \in H'$ we have $YZ \in H$.*
(b) *For every $Y, Z \in S'$ we have $YZ \in S \cdot H$.*
(c) *For every $Y \in S'$ we have $YH'Y^{-1} \subseteq H$.*
(d) *For every $Y \in S'$, some positive power of Y lies in H.*

Proof. (a) From the for statement with $Y = I$ we see that H contains $\overline{H' \cdot H'}$ because of line 5 and 9.

(b) This is clearly true after the execution of lines 11 and 12.

(c) This is clearly true after execution of line 5 and 9.

(d) Some positive power of Y lies in the connected component of $\overline{\langle Y \rangle}$, because $\overline{\langle Y \rangle}$ is an algebraic group. Now (d) follows from line 5 and 8.

□

Theorem 3. *The algorithm terminates and the output is $\widetilde{G}$, the Zariski closure of $\langle X_1, X_2, \dots, X_k \rangle$.*

Proof. Let H_i and S_i be the values of H and S respectively at the end of the `repeat-until` loop, just before line 13. We have that H_i is an irreducible Zariski closed subset of GL_n by Lemma 3(a) and that

$$H_1 \subseteq H_2 \subseteq H_3 \subseteq \cdots$$

Hence we must have

$$H_l = H_{l+1} = H_{l+2} = \cdots$$

for some l because GL_n has finite dimension. We write $\widetilde{H} = H_l$. We claim that $\widetilde{H}$ is a normal subgroup of $\widetilde{G}$. It suffices to show that $\widetilde{H}$ is closed under conjugation by $X_1, X_2, \ldots, X_k$. For every i we have $X_i \widetilde{H} X_i^{-1} = X_i H_l X_i^{-1} \subseteq H_{l+1} = \widetilde{H}$ by Lemma 4(c). Since $\widetilde{H}$ is a normal subgroup of $\widetilde{G}$ we can form the quotient group $\widetilde{G}/\widetilde{H}$ which is again a linear algebraic group. Consider the sequence

$$S_l/\widetilde{H} \subset S_{l+1}/\widetilde{H} \subset S_{l+2}/\widetilde{H} \subset \cdots$$

Note that every inclusion is a strict inclusion. For any i, $S_i/\widetilde{H}$ consists of elements of finite order in $\widetilde{G}/\widetilde{H}$ by Lemma 4(d). Let $\widetilde{S}$ be the union of all $S_l, S_{l+1}, \ldots$. The quotient $\widetilde{S}/\widetilde{H}$ is a group since it is stable by multiplication (this follows from Lemma 4(b)) and stable by inverse (this follows from Lemma 4(d)). This group must be finite by Theorem 4, and the loop therefore terminates.

After termination it is clear that $G = S \cdot H$ is closed under multiplication by Lemma 4(a),(b),(c). Now G is a Zariski closed subgroup of GL_n by Lemma 5 below. Also G is contained in $\widetilde{G}$. The group G contains $I, X_1, X_2, \ldots, X_k$, so this implies that G contains $\widetilde{G}$. We conclude that $G = \widetilde{G}$. □

Lemma 5. *Let H be a nonempty Zariski closed subset of GL_n such that $H \cdot H$ is contained in H. Then H is an algebraic subgroup of GL_n.*

Proof. We have to show that H contains the identity I and that H is closed under inverse. Let $g \in H$. For every i we have that $g^{i+1}H$ is a Zariski closed subset of $g^i H$. We get

$$H \supseteq gH \supseteq g^2 H \supseteq g^3 H \supseteq \cdots$$

By the Noetherian property, $g^i H = g^{i+1} H$ for some i. But then we get also $g^{-1}H = H$. Since $g \in H$ we have $I = g^{-1} g \in H$. Because $I \in H$ we have $g^{-1} I = g^{-1} \in H$. □

Theorem 4. *Suppose that K is a field and $G \subset \mathrm{GL}_n(K)$ is a subgroup. If every element of G has finite order, then G must be finite.*

A periodic group is a group for which every element has finite order. The general Burnside problem asks whether every finitely generated periodic group is necessarily finite. Although there are counterexamples now, Schur proved that the general Burnside problem is true for subgroups of $\mathrm{GL}_n(\mathbb{C})$ (see [27]). Kaplansky generalized Schur's result to subgroups of $\mathrm{GL}_n(K)$ where K can be an arbitrary field (see [15]).

Remark 2. *We did not attempt to optimalize the running time of the algorithm for the Zariski closure of matrix groups. Instead, we described an algorithm that will work in the most general setting. In characteristic 0, one might replace H by its tangent space at the identity. The algorithm should* *then be modified accordingly. This way one may avoid Gröbner basis computations in the algorithm and one may end up with an algorithm that is actually practical.*

Remark 3. *A related easier problem is to decide whether a given finitely generated matrix group is finite. Some efficient algorithms for this are known, see* [1], [25] *and* [12].

Acknowledgment

P.K. would like to thank Etienne Ghys for pointing out reference [21]. The main algorithm was inspired by discussions with Joris van der Hoeven.

References

[1] L. Babai, R. Beals and D. Rockmore. Deciding finiteness for matrix groups in deterministic polynomial time. In *Proc. ISSAC '93*, pages 117–126. ACM Press, New York, 1993.

[2] S. Basu, R. Pollack, and M.-F. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *Journal of the ACM*, 43(6):1002–1045, 1996.

[3] Thomas Becker and Volker Weispfenning. *Gröbner bases, A computational approach to commutative algebra.* volume 141 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, 1993.

[4] A. Bertoni. The solution of problems relative to probabilistic automata in the frame of the formal languages theory. In *Vierte Jahrestagung der Gesellschaft für Informatik*, volume 26 of *Lecture Notes in Computer Science*, pages 107–112. Springer, Berlin, 1975.

[5] A. Bertoni and M. Carpentieri. Analogies and differences between quantum and stochastic automata. *Theoretical Computer Science*, 262:69-81, 2001.

[6] A. Bertoni, G. Mauri, and M. Torelli. Some recursively unsolvable problems relating to isolated cutpoints in probabilistic automata. In *Proc. 4th International Colloquium on Automata, Languages and Programming*, volume 52 of *Lecture Notes in Computer Science*, pages 87–94. Springer, Berlin, 1977.

[7] F. Beukers. Differential Galois theory. In *From number theory to physics (Les Houches, 1989)*, pages 413–439. Springer, Berlin, 1992.

[8] V. Blondel and V. Canterini. Undecidable problems for probabilistic automata of fixed dimension. `http://www.inma.ucl.ac.be/~blondel`.

[9] Harm Derksen and Gregor Kemper. Computational invariant theory. *Invariant Theory and Algebraic Transformation Groups* **I**, volume 130 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2002.

[10] D. Eisenbud. *Commutative algebra with a view toward algebraic geometry.* volume 150 of *Graduate Texts in Mathematics*. Springer, New York, 1995.

[11] Elisabetta Fortuna, Patrizia Gianni and Barry M. Trager. Derivations and radicals of polynomial ideals over fields of arbitrary characteristic. *Journal of Symbolic Computation*, 33:609–625, 2002.

[12] G. Ivanyos. Deciding finiteness for matrix semigroups over function fields over finite fields. *Israel Journal of Mathematics*, 124:185–188, 2001.

[13] E. Jeandel. Indécidabilité sur les automates quantiques. Master Thesis, Ecole Normale Supérieure de Lyon, 2002.

[14] Theo de Jong. An algorithm for computing the integral closure. *Journal of Symbolic Computation*,26(3):273–277, 1998.

[15] I. Kaplansky, *Fields and rings*, University of Chicago.

[16] Gregor Kemper. The calculation of radical ideals in positive characteristic. To appear in *Journal of Symbolic Computation*
[17] Attila Kondacs and John Watrous. On the power of quantum finite state automata. In *Proc. 38th IEEE Symposium on Foundations of Computer Science*, pages 66–75, 1997.
[18] David W. Masser. Linear relations on algebraic groups. In *New advances in transcendence theory (Durham, 1986)*, pages 248–262. Cambridge Univ. Press, Cambridge, 1988.
[19] Ryutaroh Matsumoto. Computing the radical of an ideal in positive characteristic. *Journal of Symbolic Computation* 32(3):263–271, 2001.
[20] C. Moore and J. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237(2):257–306, 2000.
[21] A. Onishchik and E. Vinberg. *Lie groups and algebraic groups*. Springer Verlag, 1990.
[22] Azaria Paz. *Introduction to Probabilistic Automata*. Academic Press, New York, N.Y., 1971.
[23] Michael O. Rabin. Probabilistic automata. *Information and Control*, 6:230–245, 1963.
[24] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. parts I, II, III. *Journal of Symbolic Computation*, 13(3):255–352, March 1992.
[25] D. N. Rockmore, K.-S. Tan and R. Beals. Deciding finiteness for matrix groups over function fields. *Israel Journal of Mathematics*, 109:93-116, 1999.
[26] P. Voutier. An effective lower bound for the height of algebraic numbers. *Acta Arithmetica*, 74(1):81-95, 1996.
[27] I. Schur, *Über Gruppen periodischer Substitutionen*, Sitzungsber. Preuss. Akad. Wiss., 619–627.
[28] Wolmer V. Vasconcelos, *Computational methods in commutative algebra and algebraic geometry*, volume 2 of *Algorithms and Computation in Mathematics*, Springer-Verlag, Berlin, 1998.
[29] M. Waldschmidt. *Diophantine approximation on linear algebraic groups*. Springer, 2000.

Harm Dersken, Department of Mathematics, University of Michigan, Ann Arbor.
E-mail address: `hderksen@umich.edu.`

Emmanuel Jeandel, Laboratoire de l'Informatique du Parallélisme, Ecole Normale Supérieure de Lyon.
E-mail address: `Emmanuel.Jeandel@ens-lyon.fr`

Pascal Koiran, Laboratoire de l'Informatique du Parallélisme, Ecole Normale Supérieure de Lyon.
E-mail address: `Pascal.Koiran@ens-lyon.fr`