6-27-1995

Deciding Finiteness for Matrix Groups Over Function Fields

Robert Beals
Institute for Advanced Study

Daniel N. Rockmore
Dartmouth College

Ki-Seng Tan
Columbia University

DECIDING FINITENESS FOR MATRIX GROUPS OVER FUNCTION FIELDS

Robert Beals
Daniel N. Rockmore
Ki-Seng Tan

Technical Report PCS-TR94-227

Revised 6/95

Deciding Finiteness for Matrix Groups over Function Fields

Robert Beals
Institute for Advanced Study
School of Mathematics
Princeton, NJ 08540

Daniel N. Rockmore*
Department of Mathematics Dartmouth College
Hanover, NH 03755

Ki-Seng Tan
Dept. of Mathematics
Columbia University
NY, NY 10027

June 27, 1995

Abstract

Let S be any *finite* subset $GL_n(\mathbf{F}(t))$ where $\mathbf{F}$ is a field. In this paper we give algorithms to decide if the group generated by S is finite. In the case of characteristic zero, slight modifications of earlier work of Babai, Beals and Rockmore [1] give polynomial time deterministic algorithms to solve this problem. The case of positive characteristic turns out to be more subtle and our algorithms depend on a structure theorem proved here, generalizing a theorem of Weil. We also present a fairly detailed analysis of the size of finite subgroups in this case and give bounds which depend upon the number of generators. To this end we also introduce the notion of the **diameter** of a finitely generated algebra and derive some upper bounds related to this quantity. In positive characteristic the deterministic algorithms we present are exponential. A randomized algorithm based on ideas of the Meat-Axe is also given. While not provably efficient, the success of the Meat-Axe suggests the randomized algorithm will be useful.

1 Introduction

The *finiteness problem* for finitely generated groups can be stated as follows: let Γ be an infinite group, then give an efficient algorithm for deciding for any finite subset $S \subset \Gamma$, if $G = \langle S \rangle \leq \Gamma$ is finite. Such an algorithm is said to *decide finiteness for subgroups of* Γ.

In the case of $\Gamma = GL(n, K)$, for K a number field, [1] gives polynomial time (in $|S|, n$ and $[K : \mathbf{Q}]$) deterministic and randomized (Monte Carlo) algorithms to decide finiteness. The purpose of this note is to extend those results to the case of function fields, i.e. fields $K(t)$, where K is a field (possibly of

*D. Rockmore supported in part by an NSF Math Sciences Posdoctoral Fellowship as well as NSF DMS Award 9404275

positive characteristic) and t is an indeterminate, so that $K(t)$ represents the field of rational functions over K. Thus, any element $f \in K(t)$ can be written as

$$f = \frac{p(t)}{q(t)}$$

for some polynomials $p(t)$, $q(t) \in K[t]$.

The ideas of [1] extend fairly readily to $K(t)$ when K has characteristic 0. Consequently, in this case polynomial time algorithms for deciding finiteness are obtained (Theorem 2.1). The techniques used in characteristic 0 depend quite heavily on two conditions which do not extend to positive characteristic:

(1) The enveloping algebra of any finite matrix group over a field of characteristic 0 will be semisimple.

(2) Finite subgroups of $GL(n, K(t))$ are necessarily conjugate to finite subgroups of $GL(n, K)$.

Thus it would seem that new ideas may be needed to give efficient algorithms in positive characteristic. Our results in positive characteristic are a first step in this direction. Towards the goal of efficient algorithms to determine finiteness in this case we prove the following result:

Theorem 3.3 *Let $G \leq GL(n, \mathbf{F}_q(t))$ be finite. Then G is conjugate to a subgroup of $GL(n, \mathbf{F}_q(t))$ of the form*

$$\begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ 0 & 0 & \vdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$$

where the $A_i \in GL(n, \mathbf{F}_q)$ and the elements in the upper triangle are all in $\mathbf{F}_q[t]$ and of bounded degree.

Theorem 3.3 is in fact a special case of a general decomposition theorem for matrix groups over local fields (Theorem 3.1), extending an earlier result of Weil ([12], Theorem 1).

Thus, Theorem 3.3 essentially reduces the problem of deciding finiteness in positive characteristic to finding invariant subspaces of $V \cong (\mathbf{F}_q(t))^n$. Once an invariant subspace is found it is easy to check if the associated restricted representation is defined only over $\mathbf{F}_q$. At present we can only give provably exponential deterministic algorithms for deciding finiteness in this case.

However in this case randomized techniques should be very useful. Parker's "Meat-Axe" [9] is a randomized algorithm used to decompose modular representations of finite groups. While to date, there are no theoretical bounds for its expected running time, a wealth of experience shows it to be a very useful and efficient algorithm. We outline how an adaptation of this method could be used in our situation.

An interesting difference in the case of positive characteristic is that here, arbitrarily large finite subgroups may occur. We devote a fair amount of Section 3 (as well as an Appendix) to deriving bounds on the size of finitely generated subgroups in terms of the number of generators. Partially to this end we introduce and briefly discuss the notion of the diameter of finitely generated algebra.

2 Characteristic Zero

By the results of [1], to give a polynomial time algorithm to decide finiteness for $K(t)$ for $char(K) = 0$, it is enough to give a polynomial time reduction to the case of $GL(n, \mathbf{Q})$.

Theorem 2.1 *Let $\langle S \rangle = G \leq GL(n, \mathbf{Q}(t))$. Then in polynomial time either G can be transformed into an equivalent subgroup of $GL(n, \mathbf{Q})$, or shown to be infinite.*

Lemma 2.2 *Let $G \leq GL(n, \mathbf{Q}(t))$ be finite, then G is conjugate over $\mathbf{Q}(t)$ to a subgroup of $GL(n, \mathbf{Z})$.*

Proof: By ([1], Proposition 2.3) it is enough to show that if G is finite then G is conjugate over $GL(n, \mathbf{Q}(t))$ to a subgroup of $GL(n, \mathbf{Q})$. For this, notice that if G is finite, then there exists a finite extension K of $\mathbf{Q}$ such that the defining representation of G in $GL(n, \mathbf{Q}(t))$ is equivalent to a representation of G in $GL(n, K)$. In this case, there exists a matrix $X \in GL(n, K(t))$ such that

$$XAX^{-1} \in GL(n, K) \tag{1}$$

for all $A \in G \leq GL(n, \mathbf{Q}(t))$. Since G is finite, there exists a rational number α such that α is not a root of any denominator in any $A \in G$, X or X^{-1}. For any matrix $B \in GL(n, K(t))$ let $B|_\alpha$ denote the evaluation of B at α (assuming it is defined). Thus, by (1),

$$X|_\alpha \; A|_\alpha \; X^{-1}|_\alpha = A'|_\alpha = A' \tag{2}$$

where the last equality follows from the fact that $A' \in GL(n, K)$ by (1).

Since $G \leq GL(n, \mathbf{Q}(t))$ and $\alpha \in \mathbf{Q}$, $A|_\alpha \in GL(n, \mathbf{Q})$. Thus, combining (1), (2) and the fact that $X^{-1}|_\alpha = (X|_\alpha)^{-1}$ we have that for all $A \in G$,

$$X^{-1}|_\alpha \; X \; A \; X^{-1} \; X|_\alpha = A|_\alpha \in GL(n, \mathbf{Q}).$$

■

Corollary 2.3 *If $G \leq GL(n, \mathbf{Q}(t))$ is finite, then $trace(A) \in \mathbf{Z}$ for all $A \in G$.*

Lemma 2.2 shows that if G is finite, then $dim_{\mathbf{Q}}(env_{\mathbf{Q}}(G)) \leq n^2$. Consequently, this gives a simple test for infiniteness. Independent elements for $env_{\mathbf{Q}}(G)$ can be generated until either more than n^2 of these elements are obtained, or a basis over $\mathbf{Q}$ of dimension less than n^2 is found. In the former case the group is infinite. In the latter case, the basis over $\mathbf{Q}$ can be used to potentially find a representation of G in $GL(d, \mathbf{Q})$ for $d \leq n^2$ or again prove infiniteness (cf. Lemma 2.7).

Proposition 2.4 *If $G \leq GL(n, \mathbf{Q}(t))$ is finite then $env(G)$ is semisimple over $\mathbf{Q}(t)$.*

Proof: This is a simple application of Maschke's theorem ([3], Theorem 10.8).

■

By successive matrix multiplications and Gaussian elimination we have the following result.

Lemma 2.5 *Let $\langle S \rangle = G \leq GL(n, \mathbf{Q}(t))$. Then a basis $A_1, \ldots, A_d$ for $env_{\mathbf{Q}(t)(G)}$ (so $d \leq n^2$) can be constructed in polynomial time. Furthermore, the A_i can be taken to be in S^{i-1}.*

Lemma 2.6 *Let all notation be as in Lemma 2. Let $T = ((trace(A_i A_j))$. Then $env_{\mathbf{Q}(t)}(G)$ is semisimple iff $T \in GL(n, \mathbf{Q}(t))$.*

Proof: This is essentially a classical result of Dickson [4], restated in [2].

■

Lemma 2.7 *Let $G \leq GL(n, \mathbf{Q}(t))$ be such that $trace(A) \in \mathbf{Z}$ for all $A \in G$. If $env_{\mathbf{Q}(t)}(G)$ is semisimple, then in polynomial time an isomorphism can be constructed between G and a subgroup of $GL(d, Q)$ where $d = dim(env_{\mathbf{Q}(t)}(G))$.*

Proof: As in Lemma 2.5, let $A_1, \ldots, A_d$ be a basis for $env_{\mathbf{Q}(t)}(G)$ with $A_i \in S^{i-1} \subset G$. Let $A \in G$. Then we may write

$$A = \sum_{i=1}^{d} \alpha_i A_i \tag{1}$$

where each $\alpha_i \in \mathbf{Q}(t)$. We will show that in fact, each $\alpha_i \in \mathbf{Q}$.

To show this, let $\underline{\alpha} = (\alpha_1, \ldots, \alpha_d)^T \in \mathbf{Q}(t)^d$. Let $T = ((trace(A_i A_j))$. By the semisimplicity of $env_{\mathbf{Q}(t)}(G)$, we know that $T \in GL(n, \mathbf{Q}(t))$. However, by the assumed condition on the traces of the elements of G, and the fact that each $A_i A_j \in G$, we see that in fact, $T \in GL(d, \mathbf{Q})$. Thus, let $\underline{\tau} = T\underline{\alpha}$. Note that

$$\begin{aligned} \tau_i &= \textstyle\sum_{j=1}^{d} trace(A_i A_j)\alpha_j \\ &= trace(A_i \textstyle\sum_{j=1}^{d} \alpha_j A_j) \\ &= trace(A_i A). \end{aligned}$$

Thus, again, assuming that each element in G has trace in $\mathbf{Z}$, we see that $\underline{\tau} \in \mathbf{Z}^d$. Hence $\underline{\alpha} = T^{-1}\underline{\tau} \in \mathbf{Q}^d$.

Consequently, by considering the action of G on the $\mathbf{Q}$-span of $A_1, \ldots, A_d$ we get a representation of G in $GL(d, \mathbf{Q})$. The representation is faithful since $AA_i = A_i$ for all i implies that $A = 1$. Furthermore, it suffices to determine the action of S on $A_1, \ldots, A_d$. In polynomial time we can either find this faithful representation of G in $GL(d, \mathbf{Q})$ or we will discover that for some $A \in S$, the product AA_i cannot be represented as a $\mathbf{Q}$ linear sum of the $A_1, \ldots, A_d$ and then G is not finite.

■

In fact, as H. Bass has pointed out to us, it is easy to see that this argument works for any finite number of indeterminates and thus,

Theorem 2.8 *Let $\langle S \rangle = G \leq GL(n, \mathbf{Q}(t_1, \ldots, t_m))$ for independent indeterminates $t_1, \ldots, t_m$. Then in polynomial time either G can be transformed into an equivalent subgroup of $GL(n, \mathbf{Q})$, or shown to be infinite.*

3 Positive Characteristic

It appears that the arguments used in the characteristic zero case are not applicable to positive characteristic. For example, in general, if the group is finite, the associated enveloping algebra will not be semisimple. Also, subgroups of $GL(n, \mathbf{F}_q(t))$ can be arbitrarily large. (When $n = 2$, consider the upper triangular subgroups generated by elements with monomials of differing degrees in the upper corner.) Consequently, new ideas seem to be necessary. As a first step we prove a structure theorem for finite subgroups of $GL(n, \mathbf{F}_q(t))$. This will follow from a much more general decomposition theorem, generalizing slightly a classical result of Weil [12] as well as other work related to computing L-series for modular functions over function fields [11].

Notice that a subgroup of $GL(n, \mathbf{F}_q(t))$ is finite if and only if its enveloping algebra over $\mathbf{F}_q$ is finite, or equivalently, finite dimensional. The structure theorem (Theorem 3.3) allows us to give some bounds on the size of the enveloping algebra over $\mathbf{F}_q$, and consequently some coarse bounds on the size of finite subgroups of $GL(n, \mathbf{F}_q(t))$ generated by a fixed number of elements (Corollary 3.5).

With these results in hand we may then give some naive algorithms for determining finiteness. We anticipate that there is much room for improvement.

3.1 Structure Theorems for $GL(n, \mathbf{F}_q(t))$

We first introduce some notation which will be in use throughout this section.

Let $K = \mathbf{F}_q((\frac{1}{t}))$, and $\mathcal{O} = \mathbf{F}_q[[\frac{1}{t}]]$. Let $\Gamma = GL(n, \mathbf{F}_q[t]) \subset GL(n, K)$. Define

$$\mathcal{R} = \{\underline{r} = (r_1, \ldots, r_n) | r_i \in \mathbf{Z}, 0 \leq r_1 \leq r_2 \leq \cdots \leq r_n\}$$

and for each $\underline{r} \in \mathcal{R}$, define

$$\rho_{\underline{r}} = \begin{pmatrix} t^{r_1} & 0 & \cdots & 0 \\ 0 & t^{r_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t^{r_n} \end{pmatrix}.$$

Theorem 3.1 *With the notation above, for each $g \in GL(n, K)$ there exist $\gamma \in \Gamma$, $\underline{r} \in \mathcal{R}$, $\xi \in GL(n, \mathcal{O})$ and $\zeta \in \mathcal{Z}$ such that*

$$g = \gamma \cdot \rho_{\underline{r}} \cdot \xi \cdot \zeta.$$

The proof of Theorem 2 requires the following lemma (a decomposition lemma) which generalizes the dimension two result given by Lemma 3 in [11].

Lemma 3.2 *(Decomposition lemma) Let*

$$\mathcal{S} = \{(i; a_{i+1}, \ldots, i_n) \mid i = 1, \ldots, n;\ j = 1, \ldots, n-i;\ a_{i+j} \in \mathbf{F}_q\}.$$

For each $s \in \mathcal{S}$, let

$$\rho'_s = \begin{pmatrix} I_i & 0 & 0 \\ 0 & \pi & a_{i+1} \ \ldots a_n \\ 0 & 0 & I_{n-i-1} \end{pmatrix}$$

where I_m denotes the $m \times m$ identity matrix.

Then

$$GL(n, \mathcal{O}) \cdot \begin{pmatrix} \pi & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = \bigcup_{s \in \mathcal{S}} \rho'_s \cdot GL(n, \mathcal{O}).$$

Assuming the decomposition lemma, we prove Theorem 3.3.

Proof: (of Theorem 3.3) After multiplication by a suitable element $\zeta \in \mathcal{Z}$ we can assume that all the entrices of g are in $\mathcal{O}$.

By the *elementary divisor theorem* ([8], Theorem 3.8), there exists $\sigma \in GL(n, \mathcal{O})$ and $s_n \geq s_{n-1} \geq \cdots \geq s_1 \geq 0, \qquad s_i \in \mathbf{Z}$ such that

$$g \in \sigma \cdot \begin{pmatrix} \pi^{s_1} & 0 & \cdots & 0 \\ 0 & \pi^{s_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pi^{s_n} \end{pmatrix} \cdot GL(n, \mathcal{O}).$$

Now we have the factorizations

$$\begin{pmatrix} \pi^{s_1} & 0 & \cdots & 0 \\ 0 & \pi^{s_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pi^{s_n} \end{pmatrix} = \begin{pmatrix} \pi & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \end{pmatrix}^{s_1} \begin{pmatrix} 1 & 0 \cdots & 0 & \\ 0 & \pi & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}^{s_2} \cdots \begin{pmatrix} 1 & 0 \cdots & 0 & \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pi \end{pmatrix}^{s_n}$$

and

$$\begin{pmatrix} I_i & 0 & 0 \cdots 0 \\ 0 & \pi & 0 \cdots 0 \\ 0 & 0 & I_{n-i-1} \end{pmatrix} = E_{1,i} \cdot \begin{pmatrix} \pi & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \cdot E_{1,i}^{-1}$$

where for any $m > 0$, I_m denotes the $m \times m$ identity matrix and $E_{i,1} = E_{i,1}^{-1} \in GL(n, \mathbf{F}_q)$ is the appropriate permutation matrix interchanging 1 and i.

By induction on $s = s_1 + s_2 + ... + s_n$ and using the decomposition lemma we now have

$$\begin{aligned} g &= \gamma' \cdot \rho_{\underline{r}'} \cdot \sigma' \cdot E \cdot \begin{pmatrix} \pi & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \cdot E^{-1} \cdot \xi_0 \cdot \zeta_0' \quad (\gamma' \in \Gamma, E \in GL(n, \mathbf{F}_q), \xi_0, \sigma' \in GL(n, \mathcal{O}), \zeta_0' \in \mathcal{Z}) \\ &= \gamma' \cdot \rho_{\underline{r}'} \cdot \begin{pmatrix} I_i & 0 & 0 \\ 0 & \pi & a_{i+1} \ldots a_n \\ 0 & 0 & I_{n-i-1} \end{pmatrix} E^{-1} \cdot \xi_0 \cdot \zeta_0' \\ &= \gamma' \cdot \rho_{\underline{r}'} \cdot \begin{pmatrix} I_i & 0 & 0 \\ 0 & 1 & a_{i+1} \ldots a_n \\ 0 & 0 & I_{n-i-1} \end{pmatrix} \begin{pmatrix} I_i & 0 & 0 \\ 0 & \pi & 0 \ldots 0 \\ 0 & 0 & I_{n-i-1} \end{pmatrix} E^{-1} \cdot \xi_0 \cdot \zeta_0' \end{aligned}$$

But

$$\rho_{\underline{r}'} \cdot \begin{pmatrix} I_i & 0 & 0 \\ 0 & 1 & a_{i+1} \ldots a_n \\ 0 & 0 & I_{n-i-1} \end{pmatrix} = \alpha \cdot \rho_{\underline{r}'},$$

for some $\alpha \in \Gamma$. Since $\pi = \frac{1}{t}$, by multiplying by an appropriate power of t (as an element of $\mathcal{Z}$), and conjugation by an appropriate permutation matrix, we can write

$$\rho_{\underline{r}'} \cdot \begin{pmatrix} I_i & 0 & 0 \\ 0 & \pi & 0 \ldots 0 \\ 0 & 0 & I_{n-i-1} \end{pmatrix} = P \cdot \rho_{\underline{r}} \cdot P^{-1} \cdot \zeta_0''$$

for $P \in GL(n, \mathbf{F}_q)$, $\zeta_0'' \in \mathcal{Z}$ and $\underline{r}\mathcal{R}$.

Taking $\gamma = \gamma' \cdot \alpha \cdot P$, $\xi = P^{-1} \cdot E^{-1} \cdot \xi_0$, and $\zeta = \zeta_0'' \cdot \zeta_0'$, the theorem is proved.

■

To complete things, we now prove the decomposition lemma.

Proof: (of the decomposition lemma.) For $\sigma = ((\sigma_{ij})) \in GL(n, \mathcal{O})$ we have

$$\sigma \cdot \begin{pmatrix} \pi & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \pi\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \pi\sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \pi\sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}.$$

For $a \in \mathcal{O}$ let $\overline{a}$ denote its residue class in $\mathbf{F}_q$. Then let $v_i = (\sigma_{2i}, \ldots, \sigma_{ni}) \quad i = 1, \ldots, n$ and $\overline{v_i} = (\overline{\sigma_{2i}}, \ldots, \overline{\sigma_{ni}}) \quad i = 1, \ldots, n$. Note that the $\{\overline{v_i}\}_{i=1,\ldots n}$ has rank $n-1$ over $\mathbf{F}_q$.

Let i be the largest index such that

$$\overline{v_i} = \sum_{j>i} a_j \cdot \overline{v_j}.$$

Then

$$v_i = \pi w + \sum_{j>i} a_j v_j$$

for some $w \in \mathcal{O}^{n-1}$.

Suppose $w = (w_2, \ldots, w_n)$. Let $\sigma' = (\sigma'_{k,l})$ be such that

$$\sigma'_{kl} = \begin{cases} \sigma_{kl} & \text{for} \quad k \neq 1, l \neq i \\ \pi\sigma_{1l} & \text{for} \quad k = 1, l \neq i \\ w_k & \text{for} \quad k \neq 1, l = i \\ \sigma_{1i} & \text{for} \quad k = 1, l = i. \end{cases}$$

Then

$$\sigma \cdot \begin{pmatrix} \pi & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} I_i & 0 & 0 \\ 0 & \pi & a_{i+1} \ \ldots \ a_n \\ 0 & 0 & I_{n-i-1} \end{pmatrix} \cdot \sigma'.$$

■

Define now subgroups $\Gamma_{\underline{r}} \leq \Gamma$ by

$$\Gamma_{\underline{r}} = \{\gamma = (\gamma_{ij}) \in \Gamma \mid t^{r_i - r_j}\gamma_{ij} \in \mathcal{O}\}.$$

Note that $\Gamma_{\underline{r}}$ is a finite subgroup of Γ. In particular, let $\gamma \in \Gamma_{\underline{r}}$, then (1) if $r_i = r_j$, then $\gamma_{ij} \in \mathbf{F}_q$; (2) if $r_i > r_j$, then $\gamma_{ij} = 0$; (3) if $r_i < r_j$ then $deg(\gamma_{ij}) \leq r_j - r_i$. Up to conjugacy, these are the only finite subgroups which can occur.

Theorem 3.3 *Suppose $G \leq GL(n, \mathbf{F}_q(t))$ is finite. Then there exists $\Delta \in GL(n, \mathbf{F}_q(t))$ and $\underline{r} \in \mathcal{R}$ such that*

$$\Delta \cdot G \cdot \Delta^{-1} \leq \Gamma_{\underline{r}}.$$

Proof: Since G is finite, under its natural action on $\mathbf{F}_q(t)^n$, G stabilizes a rank n free $\mathbf{F}_q[t]$ submodule. This shows that G is conjugate to a subgroup of Γ. Without loss of generality, we assume that G is a subgroup of Γ. As a subgroup of $GL(n, K)$, G also acts on K^n. Again, the finiteness of G implies that there exists $g \in GL(n, K)$ such that $G \cdot g \subset g \cdot GL(n, \mathcal{O})$. By Theorem 3.1 we can write

$$g = \gamma \cdot \rho_{\underline{r}} \cdot \xi \cdot \zeta.$$

Then for $x \in G$,

$$\rho_{\underline{r}}^{-1} \cdot \gamma^{-1} \cdot x \cdot \gamma \cdot \rho_{\underline{r}} \in GL(n, \mathcal{O}) \cdot \mathcal{Z}.$$

Let $y = \gamma^{-1} \cdot x \cdot \gamma$. Then

$$\begin{pmatrix} t^{-r_1} & 0 & \cdots & 0 \\ 0 & t^{-r_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t^{-r_n} \end{pmatrix} \cdot (t^{r_i - r_j} y_{ij}) \cdot \begin{pmatrix} t^{r_1} & 0 & \cdots & 0 \\ 0 & t^{r_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t^{r_n} \end{pmatrix} = \xi' \cdot \zeta'$$

for some $\zeta' \in \mathcal{Z}$ and for some $\xi' \in GL(n, \mathcal{O})$.

Since $y \in \Gamma$, $det(\rho_{\underline{r}}^{-1} \cdot y \cdot \rho_{\underline{r}}) \in \mathbf{F}_q{}^*$, and $det(\xi') \in \mathbf{F}_q[[\frac{1}{t}]]^*$, we have $\zeta'^n = det(\zeta') \in \mathbf{F}_q[[\frac{1}{t}]]^*$. Hence, $\zeta' \in \mathbf{F}_q[[\frac{1}{t}]]^* \subset GL(n, \mathcal{O})$. Thus, $t^{r_i - r_j} y_{ij} \in \mathcal{O}$ and $y \in \Gamma_r$.

■

3.2 Bounds for finite subgroups of $GL(n, \mathbf{F}_q(t))$

In characteristic 0 finite subgroups of $GL(n, K(t))$ are bounded in size. All such bounds derive from bounds for on the size of subgroups of $GL(n, \mathbf{Z})$. For this situation, W. Feit has recently announced [5] that by using some unpublished results of B. Weisfeiler and the classification of finite simple groups (cf. [7]), except for $n = 2, 4, 6, 7, 8, 9$, and 10 the these subgroups can have size at most $2^n n!$, for which the signed permutation matrices provide an example. In the other cases, the Weyl groups of some of the exceptional groups provide larger bounds. (We also remark that some simple asymptotic analysis allows one to show that for any $\epsilon > 0$, there exists a constant $c > 2$ that depends only on ϵ such that the size of finite subgroups of $GL(n, \mathbf{Z})$ is bounded by $(n!)^{1+\epsilon} c^n$ [10].)

Other interesting results on finite subgroups of $GL(n, \mathbf{Q})$ and a wealth of references can be found in Friedland's recent paper [6].

In positive characteristic the situation is very different. Here, finite subgroups can be arbitrarily large. However, for a fixed number of generators a bound can be given. More precisely, if a finite subgroup is generated by r elements of $GL(n, \mathbf{F}_q(t))$ then a bound on the dimension of the algebra over $\mathbf{F}_q$ can be determined, and hence a bound on the size of the subgroup given. We present two different approaches towards such results. The former uses a simple analysis of the matrix multiplications which can occur (Section 3.2.1) while the latter (Section 3.2.2) introduces the notion of the diameter of a finitely generated algebra and is able to obtain bounds in terms of the diameters of certain subalgebras of $M(n_i, \mathbf{F}_q)$ for $\underline{n} = (n_1, \ldots, n_p) \models n$. We discuss the issue of the "tightness" of these bounds and show that in the case of $\underline{n} = 1^n$, the bounds obtained are indeed tight.

3.2.1 General upper bounds.

To begin, let us introduce some notation. Let $\underline{n} = (n_1, \ldots, n_d)$ be a composition of n, denoted as $\underline{n} \models n$. Let $U(\underline{n}, \mathbf{F}_q(t)) \leq GL(n, \mathbf{F}_q(t))$ be defined as the block upper triangular matrices $X = (X_{ij})$ such that X_{ij} is $n_i \times n_j$ and

$$\begin{array}{rcl} X_{ii} & \in & GL(n_i, \mathbf{F}_q) \\ X_{ij} & = & 0 \text{ if } i > j \end{array}$$

Note that $\Gamma_{\underline{r}} \leq U(\underline{n}, \mathbf{F}_q(t))$.

Theorem 3.4 *Let $\langle S \rangle = G \leq GL(n, \mathbf{F}_q(t))$ such that $|S| = r$ and G is finite. Then*

$$dim_{\mathbf{F}_q}(env_{\mathbf{F}_q}(G)) \leq \begin{cases} n & \text{if } r = 1 \text{ and} \\ \frac{1}{r}(r+1)^n & \text{if } r > 1. \end{cases}$$

Proof: First let $r = 1$, so that $G = \langle A \rangle$ and A has finite order. By Theorem 3.3, we know that $A \sim A' \in U(\underline{n}, \mathbf{F}_q(t))$ for some $\underline{n}$ a composition of n. In particular, the characteristic polynomial of A is of degree n with coefficients in $\mathbf{F}_q$. Since A will satisfy its characteristic polynomial, its powers will span a space of at most dimension n over $\mathbf{F}_q$.

Now let $r > 1$. We will repeatedly make use of the following very pessimistic bound.

Claim *Let $X_{i,i+1} \in M^{n_i \times n_{i+1}}(\mathbf{F}_q[t])$ for $i = 1, \ldots, d$ and let $G_i \leq GL(n_i, \mathbf{F}_q)$ for $i = 1, \ldots, d+1$. Let*

$$V = span_{\mathbf{F}_q}\Big(\{g_1 \cdot X_{1,2} \cdot g_2 \cdot X_{2,3} \cdot g_3 \cdots g_d \cdot X_{d,d+1} \cdot g_{d+1} \mid g_i \in G_i\}\Big).$$

Then

$$dim_{\mathbf{F}_q}(V) \leq n_1^2 n_2^2 \cdots n_{d+1}^2.$$

Proof: Notice that the entries of $X_{1,2}$ are contained in an $\mathbf{F}_q$ vector space of at most dimension $n_1 n_2$ (a set of $n_1 n_2$ polynomials over any field will span a vector space of at most dimension $n_1 n_2$ over that field). Pre- and postmultiplication by any g_1 and g_2 only effects a linear combination of these. Further postmultiplication by $X_{2,3}$ gives linear combinations of now at most $n_1 n_2 n_2 n_3 = n_1 n_2^2 n_3$ polynomials. Continuing in this fashion we see that the entries of any element in V can be linear combinations of at most $n_1 n_2^2 \cdots n_d^2 n_{d+1}$ polynomials. As there are $n_1 n_{d+1}$ entries in any element of V, we see that

$$dim_{\mathbf{F}_q}(V) \leq n_1 n_{d+1} \cdot n_1 n_2^2 \cdots n_d^2 n_{d+1} = n_1^2 n_2^2 \cdots n_d^2 n_{d+1}^2.$$

■

Using the claim we may now prove the bound for $r > 1$. Note that we may assume that G is generated by $S = \{A_1, A_2, ..., A_r\}$ such that each $A_k = (X_{ij}^{(k)}) \in U(\underline{n}, \mathbf{F}_q(t))$ and so is of a a fixed upper-triangular block-form. (As usual, we assume that $X_{ij}^{(k)}$ refers to the i,j block of size $n_i \times n_j$ in A_k.)

For any product $A_{a_1} \cdots A_{a_k}$ we have that

$$(A_{a_1} \cdots A_{a_k})_{i,j} = \left(\sum_{i \leq l_1 \leq ... \leq l_{k-2} \leq j} X_{il_1}^{(a_1)} \cdot X_{l_1 l_2}^{(a_2)} \cdot .... \cdot X_{l_{k-2} l_{k-1}}^{(k-1)} \cdot X_{l_{k-1} j}^{(k)} \right). \tag{1}$$

Since $X_{ii}^{(a_j)} \leq GL(n_i, \mathbf{F}_q)$ for each i and j, the above sum is a sum of matrices of the form

$$C_{i_0 i_0} \cdot Y_{i_0 i_1} \cdot C_{i_1 i_1} \cdot Y_{i_1 i_2} \cdot C_{i_2 i_2} \cdots Y_{i_m j} \cdot C_{jj}, \tag{2}$$

where $i = i_0 < i_1 < i_2 < ... < i_m < j$, $C_{tt} \in GL(n_t, \mathbf{F}_q)$, and $Y_{t,t'} \in \{X_{t,t'}^{(l)}, l = 1, 2, 3, ..., r\}$.

Let $E = env_{\mathbf{F}_q}(G)$. In order to bound $dim_{\mathbf{F}_q} E$ we use the simplification,

$$dim_{\mathbf{F}_q} E \leq \sum_i n_i^2 + \sum_{i<j} dim_{\mathbf{F}_q} V_{i,j} \tag{3}$$

where

$$V_{i,j} = span_{\mathbf{F}_q}\{X_{i,j} : X \in E\}. \tag{4}$$

That is, $V_{i,j}$ is the subspace of $M^{n_i \times n_j}(\mathbf{F}_q[t])$ spanned by the i,j blocks of all elements of E. The claim shows that

$$dim_{\mathbf{F}_q} V_{i,j} \leq \sum_{i<i_1<...<i_m<j} n_i^2 \cdot n_{i_1}^2 \cdots n_{i_m}^2 \cdot n_j^2 \cdot r^{m+1}, \tag{5}$$

where the factor of r^{m+1} comes from the r possible choices for each of the $Y_{t,t'}$ as defined in (2).

Thus, using (3) and (5) over all sequences $1 \leq i < i_1 < ... < i_m < j \leq d$ and reindexing we get

$$\begin{aligned} dim_{\mathbf{F}_q}(E) &\leq \sum_i n_i^2 + \sum_{1 \leq i < j \leq d} \frac{1}{r} \sum_{i=i_1<...<i_m=j} (rn_{i_1}^2) \cdots (rn_{i_m}^2) \\ &= \textstyle\sum_i n_i^2 + \frac{1}{r}(rn_1^2 + 1) \cdot ... \cdot (rn_d^2 + 1) - \frac{1}{r}(1 + rn_1^2 + rn_2^2 + ... + rn_d^2). \end{aligned} \tag{6}$$

This shows that

$$dim_{F_q} X < \frac{1}{r}(rn_1^2 + 1) \cdots (rn_d^2 + 1)$$

Finally, notice that if $r > 2$, then for any $m > 0$,

$$(rm^2 + 1) \leq (r + 1)^m$$

and thus, (as $\sum_i n_i = n$)

$$dim_{\mathbf{F}_q}(X) \leq \frac{1}{r}(r+1)^n.$$

■

Corollary 3.5 *Let $\langle S \rangle = G \leq GL(n, \mathbf{F}_q(t))$ have finite order. Then*

$$|G| \leq \begin{cases} q^n & \text{if } r = 1 \text{ and} \\ q^{\frac{1}{r}(r+1)^n} & \text{if } r > 1. \end{cases}$$

In particular, every element of finite order in $GL(n, \mathbf{F}_q(t))$ has order at most q^n.

Proof: The size of the group can be no larger than the size of the enveloping algebra over $\mathbf{F}_q$.

■

Another easy corollary also follows.

Corollary 3.6 *An element $A \in GL(n, \mathbf{F}_q(t))$ has finite order if and only if the characteristic polynomial of A is defined over $\mathbf{F}_q$.*

Theorem 3.3 shows that any finite subgroup is conjugate to a subgroup of some $U(\underline{n}, \mathbf{F}_q(t))$. With a slight modification of the above proof of Theorem 3.4, we see that in fact, any finitely generated subgroup of $U(\underline{n}, \mathbf{F}_q(t))$ is finite. We record this fact as the next theorem.

Theorem 3.7 *Any finitely generated subgroup of $U(\underline{n}, \mathbf{F}_q(t))$ is finite.*

Remark: It is of interest to investigate the strength of the bounds of Theorem 3.4 and its corollary. For example, in the Section 3.2.3 we will get some tight bounds for the case of $\underline{n} = (1, 1, \ldots, 1)$. We will then see that the bound that Corollary 3.5 yields for $n = 2$ and $\underline{n} = (1, 1)$ is tight.

3.2.2 Upper bounds and diameters for algebras

Here we introduce the notion of **diameter** for a finitely generated algebra. In so doing we are able to obtain a different upper bound on the size of finite subgroups of $GL(n, \mathbf{F}_q(t))$ generated by r elements in terms of "natural" combinatorial data derived from the generators.

Let A be an algebra over a field $\mathbf{F}$ and $S = \{X_1, \ldots, X_r\} \subset A$. As usual, let $\text{env}_{\mathbf{F}}(S)$ denote the subalgebra of A generated by S. Furthermore, for any integer $j \geq 0$, let S^j denote the subset of elements of A which can be written as products of j elements of S.

Definition 3.8 *The subalgebra $env_{\mathbf{F}}(A)$ is said to have* **diameter** *δ, written $\delta = diam(env_{\mathbf{F}}(S))$, if all elements of $env_{\mathbf{F}}(S)$ can be written as $\mathbf{F}$-linear combinations of $S^0 \cup \cdots \cup S^\delta$ and δ is the least integer such that this is true. If $X \in env_{\mathbf{F}}(S)$, define the* **length** *of X (with respect to S), denoted $len_S(X)$ to be the smallest integer j (necessarily less than $diam(env_{\mathbf{F}}(S))$) such that $X \in span_{\mathbf{F}}(S^0, \ldots, S^j)$.*

Lemma 3.9 *Let all notation be as above and let $S \subset A$ and $\delta = diam(env_{\mathbf{F}}(S))$, then*

$$\dim_{\mathbf{F}}(env_{\mathbf{F}}(S)) \leq 1 + r + \cdots + r^\delta.$$

Proof: There are at most r^j $\mathbf{F}-$linearly independent elements in S^j.

■

Remark. Notice that diameter of an algebra is related to, albeit in a seemingly loose fashion, to the concept of diameter of a finitely generated group. For example, let $X \in GL(n, \mathbf{F}_q)$. Since X satisfies its characteristic polynomial, $\text{diam}(\text{env}_{\mathbf{F}_q}(X)) \leq n-1$. However, the order of X can be at most $q^n - 1$, in which case the diameter of the cyclic group generated by X is $\frac{q^n-1}{2}$.

At the very least, it is clear that for $S \subset GL(n, \mathbf{F})$,

$$\text{diam}(\text{env}_{\mathbf{F}}(S)) \leq \text{diam}(\langle S \rangle)$$

where the righthand side denotes the diameter of the subgroup of $GL(n, \mathbf{F})$ generated by S.

Using the concept of diameter another bound for the size of finite subgroups of $GL(n, \mathbf{F}_q(t))$ can be obtained. To simplify the statement of the result, for any $X \in U(\underline{n}, \mathbf{F}_q(t))$ (for $\underline{n} = (n_1, \ldots, n_p) \models n$) let

$$X^{(i)} = \text{the } i,i - \text{block of } X. \tag{7}$$

Thus, $X^{(i)} \in GL(n_i, \mathbf{F}_q)$.

Theorem 3.10 *Let $\underline{n} \models n$ and $S = \{X_1, \ldots, X_r\} \subset U(\underline{n}, \mathbf{F}_q(t))$. Let*

$$S_i = \{X_1^{(i)}, \ldots, X_r^{(i)}\} \subset GL(n_i, \mathbf{F}_q)$$

and

$$\delta_i = diam(env_{\mathbf{F}_q}(S_i)).$$

Then

$$diam(env_{\mathbf{F}_q}(S)) \leq \delta_1 + \cdots + \delta_p + p - 1.$$

Proof: It is enough to show that if $W = Y_1 Y_2 \cdots Y_m$ is a product of $m \geq \delta_1 + \cdots + \delta_p + p$ elements in S, then another expression for W can be found which is a linear combination of products of less than m elements in S.

Since $m \geq \delta_1 + \cdots + \delta_p + p$, W can be written as

$$W = W_1 \cdots W_p \tag{8}$$

where

Consider now the $1,1$ block of W_1 or in the notation of (7), $W_1^{(1)}$. Notice that

$$W_1^{(1)} = Y_1^{(1)} \cdots Y_{\delta_1+1}^{(1)}. \tag{9}$$

Since $\text{diam}(\text{env}_{\mathbf{F}_q}(S_1)) = \delta_1$, then it must be the case that $W_1^{(1)}$ can be written as the $1,1$ block of a $\mathbf{F}_q$-linear combination of products of elements in S of length at most δ_1. Thus, let Z_1 be such an element, so that $\text{len}_S(Z_1) \leq \delta_1$ and

$$W_1^{(1)} = Z_1^{(1)}.$$

Similarly, define Z_j to be such that

- (1) $\text{len}_S(Z_j) \leq \delta_j$ and
- (2) $W_j^{(j)} = Z_j^{(j)}$.

By condition (2)

$$(W_j - Z_j)^{(j)} = 0. \tag{10}$$

Thus, using (10) it is easy to see

$$(W_1 - Z_1)(W_2 - Z_2)\cdots(W_p - Z_p) = 0. \tag{11}$$

But (11) readily implies that

$$W = W_1 \cdots W_p = Z$$

where Z is a linear combination of products of the form $A_1 \cdots A_p$ where each for each i, A_i is either equal to W_i (and hence in S^k for some $k > \delta_i$) or is a linear combination of elements which are in S^k for $k \leq \delta_i$ and furthermore **at least** one such i satisfies the latter condition. But this implies that Z is a linear combination of elements of length less than m and the theorem is proved.

■

Remark. The above definitions and results suggest several natural questions. It would be of interest to better understand the diameters of various generating sets for subalgebras of $M(n, \mathbf{F}_q)$ and perhaps investigate their relationship to diameters of corresponding subgroups of $GL(n, \mathbf{F}_q)$. Furthermore, it would be of interest to see under what conditions, if any, the bound of Theorem 3.10 is tight. To this end, in the following section we show that in the case of $\underline{n} = (1, \ldots, 1) = 1^n$, a modification of the proof of Theorem 3.10 yields a bound which is in fact tight.

3.2.3 Bounds for $U\Big((1, 1, \ldots, 1), \mathbf{F}_q(t)\Big)$.

A small modification of the proof of Theorem 3.10 gives an improved bound for the size of finitely generated subalgebras of $U\Big((1, 1, \ldots, 1), \mathbf{F}_q(t)\Big)$. We are then able to show in this case that the bound obtained is tight.

Theorem 3.11 *Let all notation be as in Theorem 3.10. Let* $S = \{X_1, \ldots, X_r\} \subset U\Big((1, 1, \ldots, 1), \mathbf{F}_q(t)\Big)$. *Then*

$$diam(S) \leq n - 1.$$

Proof: Let $W = Y_1 Y_2 \cdots Y_m$ with $Y_i \in S$ and $m \geq n$. Let

$$\begin{aligned}
W_1 &= Y_1 = \begin{pmatrix} a_1(1) & * & \cdots & * \\ 0 & a_2(1) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{(1)} \end{pmatrix} \\
&\vdots \\
W_{n-1} &= Y_{n-1} = \begin{pmatrix} a_1(n-1) & * & \cdots & * \\ 0 & a_2(n-1) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{(n-1)} \end{pmatrix} \\
W_n &= Y_n \cdots Y_m = \begin{pmatrix} a_1(n) & * & \cdots & * \\ 0 & a_2(n) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{(n)} \end{pmatrix}.
\end{aligned} \tag{12}$$

Consequently mimicking (10) consider the relation

$$(Y_1 - a_1^{(1)} I_n)(Y_2 - a_2^{(2)} I_n) \cdots (Y_n \cdots Y_m - a_n^{(n)} I_n) = 0 \tag{13}$$

where I_n is the $n \times n$ identity matrix. Thus

$$W = Y_1 \cdots Y_n \cdots Y_m = Z \tag{14}$$

where Z is a linear combination of products of at most $n-1$ of $W_1, \ldots W_n$, and hence itself has length less than m.

■

Corollary 3.12 *Let all notation be as in Theorem 3.10. Let* $S = \{X_1, \ldots, X_r\} \subset U\Big((1,1,\ldots,1), \mathbf{F}_q(t)\Big)$. *Then*

$$\dim_{\mathbf{F}_q}(env_{\mathbf{F}_q}(S)) \leq 1 + r + \cdots r^{n-1}.$$

Corollary 3.13 *Let all notation be as in Theorem 3.10. Let* $S = \{X_1, \ldots, X_r\} \subset U\Big((1,1,\ldots,1), \mathbf{F}_q(t)\Big)$. *Let* $G = \langle S \rangle$. *Then*

$$|G| \leq q^{1+r+\cdots r^{n-1}} - 1.$$

We now take up the problem of investigating if this bound is tight. Notice that if $r = 1$, the bound given by Corollary 3.13 is tight. Recall that $GL(n, \mathbf{F}_q)$ contains elements of order $q^n - 1$, so called *Singer cycles*. They are realized as follows: Consider $\mathbf{F}_{q^n}$ as an n-dimensional vector space over $\mathbf{F}_q$. This gives a representation of $\mathbf{F}_{q^n}^{\times}$ as a subgroup of $GL(n, \mathbf{F}_q)$, by considering the action of multiplication of $\mathbf{F}_{q^n}^{\times}$ on $\mathbf{F}_{q^n}$. Any generator of $\mathbf{F}_{q^n}^{\times}$ will then have order $q^n - 1$.

When $r > 1$, the situation is slightly more complicated. For example, we note that when $n = 2$ the bound is tight. In this case Theorem 3.4 yields $dim_{\mathbf{F}_q}(env_{\mathbf{F}_q}(G)) \leq \frac{1}{r}(r+1)^2 = r + 2 + \frac{1}{r}$. So for any $r > 1$, consider the generators,

$$\begin{pmatrix} a & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & t \\ 0 & a \end{pmatrix}, \ldots, \begin{pmatrix} 1 & t^{r-1} \\ 0 & 1 \end{pmatrix}$$

where a generates ${\mathbf{F}_q}^{\times}$.

It is not hard to check that these elements generate the subgroup G,

$$G = \left\{ \begin{pmatrix} x & h(t) \\ 0 & y \end{pmatrix} \mid x, y \in {\mathbf{F}_q}^{\times}, h(t) \in \mathbf{F}_q[t], deg(h) < r \right\}.$$

Notice that $dim_{\mathbf{F}_q}(env_{\mathbf{F}_q}(G)) = r + 1$.

In fact, this sort of construction can be generalized to any $\underline{n} = 1^n$ and $r \geq 0$.

Theorem 3.14 *There exist subsets* $S = \{X_1, \ldots, X_r\} \subset U\Big((1,1,\ldots,1), \mathbf{F}_q(t)\Big)$ *such that*

$$\dim_{\mathbf{F}_q}(env_{\mathbf{F}_q}(S)) = 1 + r + \cdots r^{n-1}.$$

The proof of Theorem 3.14 is quite technical and we postpone it to an appendix.

3.3 Algorithm 1

Theorem 3.4 yields immediately a simple exponential (in n) algorithm for deciding finiteness for $GL(n, \mathbf{F}_q(t))$.

That is, given the generators $\{S_1, \ldots, S_r\}$, attempt to construct a basis for $env_{\mathbf{F}_q}(G)$. In a given round, we test if the products of the current independent set by the generators contain any new independent elements. Thus any given round takes at most $O(rm \cdot (n^2)^2)$ operations where m is the current number of independent elements. By Theorem 3.4 in at most $\frac{1}{r}(r+1)^n + 1$ rounds we will either have found a complete basis for $env_{\mathbf{F}_q}(G)$ or generated too many independent elements. We record this simple algorithm as

Theorem 3.15 *In at most* $O(r \cdot (\frac{1}{r}(r+1)^n)^2 n^4) = O(n^4 \frac{1}{r}(r+1)^{2n})$ *operations we can decide finiteness for a subgroup of* $GL(n, \mathbf{F}_q(t))$ *generated by* r *elements.*

3.4 Algorithm 2

As a next step towards giving an efficient algorithm (deterministic or randomized) we outline another exponential algorithm. Although still exponential we believe that this different framework may be more amenable to improvement.

Theorem 3.7 indicates a possibly fruitful approach towards deciding finiteness. Letting V_0 denote the vector space $\mathbf{F}_q(t)^n$, Theorem 3.7 indicates that we should be searching for a flag of V_0,

$$V_0 > V_1 \cdots > V_d > V_{d+1} = \{0\},$$

with subspaces V_i such that the successive quotients V_i/V_{i+1} are (1) G-invariant and (2) such that the induced action of G on V_i/V_{i+1} gives a representation of G in $GL(n_i, \mathbf{F}_q)$ where $n_i = dim(V_i/V_{i+1})$. By Theorem 3.7 it is enough to check (1) and (2) on the generators.

Direct implementation of this idea still seems to give an exponential upper bound.

Theorem 3.16 *Let* $S = \{S_1, \ldots, S_r\} \subset GL(n, \mathbf{F}_q(t))$. *Then in at most* $O(n^4 \cdot [(\frac{1}{r}(r+1)^n)^8])$ *operations we can decide finiteness for* $G = \langle S \rangle$.

Proof: The idea of the algorithm is as follows: Suppose $\langle S \rangle = G \sim G' \leq U(\underline{n}, \mathbf{F}_q(t))$. Then we would have a homomorphism

$$\begin{array}{ccc} G & \longrightarrow & GL(n_1, q) \times \cdots \times GL(n_d, q) = D(\underline{n}), \\ A & \mapsto & \overline{A} \end{array}$$

given by projection onto the block-diagonal subgroup.

Suppose $\mathcal{A} = \{A_1, \ldots, A_m\} \subset G$ were such that it could be guaranteed that $\overline{\mathcal{A}} = \{\overline{A_1}, \ldots, \overline{A_m}\}$ span $env_{\mathbf{F}_q}(\overline{G})$. Then, for every $A \in G$, either (1) $A \in span_{\mathbf{F}_q}(\mathcal{A})$, or (2) $A - B$ is nilpotent for some $B \in span_{\mathbf{F}_q}(\mathcal{A})$. Furthermore, if (2) holds, then the kernel of $A - B$ will contain some nonzero invariant subspace W, and V/W will also be G-invariant.

Thus, our method of attack is to attempt to successively apply the above idea until we arrive at a subspace W which is invariant and supports a representation over F_q. If G is finite this will be possible. This subspace W will then serve as V_d. Having done this, we then apply the algorithm to the quotient V/V_d and so on, ultimately arriving at a change of basis for G to a subgroup of $U(\underline{n}, \mathbf{F}_q(t))$. If G is infinite, at some point this algorithm will fail.

As usual, we make the inductive definition of $S^1 = S$ and for $k > 1$,

$$S^k = \bigcup_{A \in S} S^{k-1} A.$$

Since $D(\underline{n})$ can have at most dimension n^2 over $\mathbf{F}_q$ we have the following lemma.

Lemma 3.17 *Let G be finite, with all notation as above. Then*

$$env_{\mathbf{F}_q}(\overline{S^{n^2}}) = env_{\mathbf{F}_q}(\overline{G}).$$

Now, suppose that $\mathcal{A} = \{A_1, \ldots, A_m\}$ is an orthogonal basis for $X = span_{\mathbf{F}_q}(S^{n^2})$ over $\mathbf{F}_q$. Notice that by Theorem 3.4 it will take at most $n^4 \cdot \frac{1}{r}(r+1)^n$ operations to compute $\mathcal{A}$. We can and do assume that $A_i \in G$ for all i. If each of the products $A_i S_j$ is in the span of $\mathcal{A}$, then $env_{\mathbf{F}_q}(G)$ is finite dimensional and G is finite. Otherwise, some product of this form is not in $span_{\mathbf{F}_q}(\mathcal{A})$. Let A be such an element.

Claim *Suppose G is finite. Let A and $\mathcal{A}$ be as above. If G is finite then we can compute elements $\alpha_i \in \mathbf{F}_q$ such that*

$$A' = A - \sum_{i=1}^{m} \alpha_i A_i$$

is nilpotent and nonzero.

Proof: Consider the $d \times d$ matrix T with i,j entry given by $trace(A_i A_j)$. Note that if G is finite then T is defined over $\mathbf{F}_q$. At most $nm^2 < n\left(\frac{1}{r}(r+1)^n\right)^2$ operations are needed to form T. Also, $trace(A_i A_j) = trace(\overline{A_i A_j})$ and as a bilinear map from $\overline{X} \times \overline{X} \longrightarrow \mathbf{F}_q$ it is nondegnerate.

Consequently, we can assume that after reordering, the first $k \leq m$ rows of T are a basis for the span over $\mathbf{F}_q$ of all the rows of T. At most m^3 operations are required here.

Consider the new row vector v_A with i^{th} entry given by $trace(AA_i)$. Since $\overline{A} \in span_{\mathbf{F}_q}(\mathcal{A})$, it must be that v_A is in the span of the first k rows of T so that there exist $\alpha_i \in \mathbf{F}_q$ such that

$$trace(AA_j) = \sum_{i=1}^{k} \alpha_i \; trace(A_i A_j).$$

But this then implies that $trace((A - \sum_i \alpha_i A_i)A_j) = 0$ for all A_j. Since *trace* is nondegenerate, this can only mean that $\overline{A - \sum_i \alpha_i A_i} = 0$. But since the A_i span the subalgebra of the block-diagonal entries, this must mean that $A' = A - \sum_i \alpha_i A_i$ is equivalent to some matrix contained in the span of the strictly upper triangular block of $U(\underline{n}, \mathbf{F}_q(t))$, and thus, if nonzero, must be nilpotent.

Let $W' = kernel(A') < \mathbf{F}_q(t)^n$. Then $W = W' \cap S_1 W' \cap \cdots \cap S_r W'$ will be G-invariant and nonzero, assuming G is finite. We can then iterate the above on W.

If G is infinite, it will be detected at one of several places:

(1) More than $\frac{1}{r}(r+1)^n$ independent elements will be generated to span S^{n^2}, contradicting Theorem 3.4.

(2) For some i,j, $trace(A_i A_j) \notin \mathbf{F}_q$.

(3) $W = 0$.

If G is finite then the above procedure will need to be executed at most n^2 times. This yields an upper bound of on the total number of operations required of

$$O\Big(n^2 \cdot [n(\frac{1}{r}(r+1)^n)^3] \cdot [n(\frac{1}{r}(r+1)^n)^2] \cdot [(\frac{1}{r}(r+1)^n)^3]\Big) \leq O(n^4 \cdot [(\frac{1}{r}(r+1)^n)^8].$$

■

3.5 Remarks on Algorithms 1 and 2

Remark 1: Perhaps simplifications could be found if one makes the assumption that G has some extra structure (eg. solvable or nilpotent).

Remark 2: The main goal of Algorithm 2 is the detection of a nontrivial invariant subspace. If this could be detected with high probability in some randomized fashion the efficiency of the algorithm could be increased tremendously. Perhaps a nontrivial nilpotent element could be detected easily in some randomized way.

3.6 A Randomized Algorithm

In this section we present a simple randomized algorithm whose motivation owes much to the approach taken by Parker's "Meat-Axe", an algorithm which decomposes modular representations of finite groups. We are at present unable to give a proof of reliability here, appealing only to the success of the Meat-Axe as an indication that this idea may prove useful for implementation.

The main tool we use is a result of S. P. Norton, which is also a theoretical lynchpin in the Meat-Axe.

Theorem *(Due to S. P. Norton, c.f. [9], Section 5) Let* $\mathbf{F}$ *denote any field and* $\mathcal{S} = \{S_1, \ldots, S_r\} \subset M_n(F)$. *Then for any* $B \in env_{\mathbf{F}}(\mathcal{S})$, *at least one of the following must hold:*

(1) B *is non-singular;*

(2) At least one non-zero null vector of B *lies in a proper subspace invariant under* $\mathcal{S}$*;*

(3) Every non-zero null vector of B^T *lies in a proper subspace invariant under* $\mathcal{S}^T = \{S_1^T, \ldots, S_r^T\}$*;*

(4) There is no proper subspace invariant under $\mathcal{S}$.

Thus, let $\mathcal{S} = \{S_1, \ldots, S_n\} \subset GL(n, \mathbf{F}_q(t))$. Norton's Theorem indicates the following algorithm for deciding finiteness for $\mathcal{S}$.

Randomized Algorithm

Step 1: As in the description of Algorithm 1 (c.f. Section 3.3), attempt to generate n^2+1 independent elements over $\mathbf{F}_q$. If this is not possible (this can be determined in polynomial time), then $\langle \mathcal{S} \rangle$ is finite. Otherwise, proceed to Step 2.

Step 2: Generate a singular element $B \in env_{\mathbf{F}_{q(t)}}(\mathcal{S})$. Check if the translates of one of its null-vectors generate an invariant subspace. If one does, then perform a change of basis, thereby simultaneously rewriting the generators in some block upper triangular form,

$$S_i \sim \begin{pmatrix} A^i_{1,1} & * \\ 0 & A^i_{2,2} \end{pmatrix},$$

and $A^i_{j,j}$ is $d_j \times d_j$. Now return to Step 1, successively using as input the sets $\{A^1_{j,j}, \ldots, A^r_{j,j}\}$.

If no null-vector generates a nontrivial invariant space, then proceed to Step 3.

Step 3: Take any non-zero null vector of B^T. If this does not generate a nontrivial invariant subspace for $\mathcal{S}^T$, then G is infinite. Otherwise, we may now find a change of basis given a block upper

triangularization of the a group isomorphic to the group generated by $\mathcal{S}^T$. Note that the "transpose group" is finite if and only if the original group is finite. Return now to Step 1 with the blocks for the transposed group and continue.

The difficulty with this algorithm lies in Step 2. If we could guarantee that B has rank $n-1$, then up to scalar multiples there would be a unique null vector to test. Otherwise there are an infinite number. Thus, it is here that we would have to apply randomization. We would construct B in some randomized fashion. Parker points out that almost immediately elements of rank $n-1$ are found. If in fact elements of rank $n-1$ are not constructed, further randomization could then be applied and a null vector could be chosen randomly. The hope again is that (assuming invariant subspaces exist) with high probability a vector is found generating an invariant subspace.

References

[1] L. Babai, R. Beals and D. Rockmore. Deciding finiteness for matrix groups in deterministic polynomial time. in *Proc. ISSAC '93*, pp. 117-126.

[2] K. Friedl and L. Ronyai. Polynomial time solutions of some problems in computational algebra, in *Proc.* 17^{th} *ACM STOC*, 1985, pp. 153-162.

[3] C. W. Curtis and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras.* Wiley, NY, 1966.

[4] L. E. Dickson. *Algebras and Their Arithmetics.* University of Chicago Press, Chicago, 1923.

[5] W. Feit. Private correspondence. September 5, 1994.

[6] S. Friedland. Discrete groups of unitary isometries and balls in hyperbolic manifolds. Preprint. (1993).

[7] D. Gorenstein. *The Classification of Finite Simple Groups.* Vol. 1, Plenum Press, NY, 1982.

[8] N. Jacobson. *Basic Algebra I.* Freeman and Co., San Francisco, 1974.

[9] R. A. Parker. The computer calculation of modular characters (the Meat-Axe). in (M. Atkinson, ed.) *Computational Group Theory*, Academic Press, London, 1984, pp. 267-274.

[10] D. Rockmore and K.-S. Tan. A note on the order of finite subgroups of $GL(n, \mathbf{Z})$. *Arch. Math.* To appear.

[11] K.-S. Tan and D. Rockmore. Computation of L-series for elliptic curves over function fields. *J. reine angew. Math.* **424** (1992), 107-135.

[12] A. Weil. On the analogue of the modular group in characteristic p. In *Functional Analysis and Related Fields, Proc. Conf. in honor of M. Stone, U. of Chicago.* 1968, Springer-Verlag.