Can. J. Math., Vol. XXXVII, No. 2, 1985, pp. 238-259

THE ORBIT-STABILIZER PROBLEM FOR LINEAR GROUPS

JOHN D. DIXON

1. Introduction. Let G be a subgroup of the general linear group $GL(n, \mathbf{Q})$ over the rational field $\mathbf{Q}$, and consider its action by right multiplication on the vector space $\mathbf{Q}^n$ of n-tuples over $\mathbf{Q}$. The present paper investigates the question of how we may constructively determine the orbits and stabilizers of this action for suitable classes of groups. We suppose that G is specified by a finite set $\{x_1, \ldots, x_r\}$ of generators, and investigate whether there exist algorithms to solve the two problems:

(Orbit Problem) Given $u, v \in \mathbf{Q}^n$, does there exist $x \in G$ such that $ux = v$; if so, find such an element x as a word in $x_1, \ldots, x_r$ and their inverses.

(Stabilizer Problem) Given $u \in \mathbf{Q}^n$, describe all words in $x_1, \ldots, x_r$ and their inverses which lie in the stabilizer

$$G_u := \{x \in G \mid ux = u\}.$$

These two problems will be called jointly the *orbit-stabilizer problem.* A solution to the orbit-stabilizer problem for G, u, v will include a description of the set of all $x \in G$ for which $ux = v$ since the latter set equals $G_u x_0$ where x_0 is any element of G such that $ux_0 = v$.

As we shall see below, for some linear groups there are no algorithms to solve these problems. On the other hand, for some restricted classes of groups, we shall describe algorithms which could be quite practical for solving the orbit-stabilizer problem for moderate sizes of n (see Sections 4, 5 and 7).

Remark. We have rather uninteresting special cases when either u or v equals 0. In what follows we shall tacitly assume that both u and v are nonzero.

2. Known results and examples. 1. In the case where $n = 2$ and G is cyclic it is shown by elementary methods in [**23**] that the orbit problem is solvable and an explicit algorithm is given (compare with Section 4 below).

Received July 26, 1983 and in revised form February 13, 1984. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (Grant No. A7171).

2. In [**10**] Grunewald and Segal show that the orbit-stabilizer problem is solvable for "explicitly given" arithmetic subgroups of "explicitly given" algebraic $\mathbf{Q}$-groups in $GL(n, \mathbf{C})$, but they make no attempt to present computationally efficient algorithms (see Algorithms A and B of [**10**] noting that G_u is Zariski-closed). Similar results are obtained in [**22**].

3. The classical membership and conjugacy problems for groups are related to the orbit problem as follows. Let $Mat(n, \mathbf{Q})$ denote the vector space of all $n \times n$ matrices over $\mathbf{Q}$, and consider the two representations ρ and σ of $GL(n, \mathbf{Q})$ into $GL(Mat(n, \mathbf{Q})) \simeq GL(n^2, \mathbf{Q})$ given by:

$$a\rho(x) := ax \quad \text{and}$$

$$a\sigma(x) := x^{-1}ax \quad \text{for all } a \in Mat(n, \mathbf{Q}).$$

Then the special case of the membership problem, "Given $x \in GL(n, \mathbf{Q})$, is $x \in G$?" is equivalent to the orbit problem, "Do x and 1 lie in the same $\rho(G)$-orbit?" Similarly, the conjugacy problem, "Given $x, y \in G$, are they conjugate in G?" is equivalent to "Do x and y lie in the same $\sigma(G)$-orbit?"

4. Mihailova [**19**] (see also [**20**, p. 42]) has shown that there exists a finitely generated subgroup L of $SL(4, \mathbf{Z})$ for which the conjugacy problem is unsolvable. Hence, it follows from what we have just seen, there is a finitely generated subgroup of $GL(16, \mathbf{Q})$ for which the orbit problem is unsolvable.

5. Kopytov [**14**] shows that the membership problem is solvable for a finitely generated solvable subgroup G of $GL(n, \mathbf{Q})$, and his results may be used to show that the orbit problem is solvable for a completely reducible finitely generated solvable group G. Dan Segal has also pointed out to me that the fact that a polycylic-by-finite group has a solvable conjugacy problem can be used to show that the orbit problem is solvable for any polycyclic-by-finite subgroup of $GL(n, \mathbf{Z})$. In neither case, however, does the implied algorithm appear computationally practical. The related results in [**1**] should also be mentioned.

6. In the sections which follow we shall be concerned exclusively with the case where a solution to the stabilizer problem produces a finite set of generators for G_u. This considerably restricts the class of groups which we can consider since, even for finitely generated solvable groups, the stabilizers need not be finitely generated. For example, if

$$G = \left\langle \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\rangle \quad \text{and } u = (1\ 0)$$

then

$$G_u = \left\{ \begin{bmatrix} 1 & 0 \\ r/2^s & 1 \end{bmatrix} \middle| r, s \in \mathbf{Z} \right\}$$

which is not finitely generated.

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

7. Roger Lyndon has pointed out to me that the problem of deciding whether two matrices generate a free group may be rephrased as a stabilizer problem. Extensive work done on this particular problem (see, for example [**17**] and [**11**]) indicates that it is very difficult to decide freeness. This suggests that the stabilizer problem may be very difficult (perhaps unsolvable) for a wide class of groups.

Remark. In this paper we only consider linear groups over $\mathbf{Q}$, although similar questions could be asked for linear groups over any constructible field. Extension of the results to an algebraic number field K can be obtained immediately by treating K^n as a vector space of dimension $n[K:\mathbf{Q}]$ over $\mathbf{Q}$.

3. Estimates with valuations. This section is devoted to proving a series of estimates which will be used in the next two sections to deal with the orbit-stabilizer problem for abelian groups.

Let K be an algebraic number field and consider a set Λ of additive valuations on K which includes not only the usual nonarchimedean valuations but also additive versions of the archimedean norms. Specifically, $\Lambda = \Lambda_1 \cup \Lambda_0$ is a union of nonarchimedean and archimedean valuations. By Λ_1 we denote the set of all nonarchimedean valuations

$$\lambda: K \to \mathbf{R} \cup \{\infty\}$$

which are normalized so that λ restricted to $\mathbf{Q}$ is equal to the usual p-adic valuation λ_p for some prime p (so $\lambda(p) = \lambda_p(p) = 1$). On the other hand, if $K = \mathbf{Q}(\theta)$, and $\theta_1, \ldots, \theta_d$ are the algebraic conjugates of θ in $\mathbf{C}$, then Λ_0 consists of all $\lambda: K \to \mathbf{R} \cup \{\infty\}$ given by

$$\lambda(h(\theta)) := -\log(|h(\theta_i)|)$$

for a fixed i when $h(X) \in \mathbf{Q}[X]$. Some of these latter valuations may coincide, and their restrictions to $\mathbf{Q}$ are all equal to λ_0 which is the negative of the logarithm of the absolute value (see, for example, [**15**] for elementary properties of valuations).

We summarize some properties of Λ which we shall need.

LEMMA 1. *Let* $f(X) = f_0 + f_1X + \ldots + f_sX^s \in \mathbf{Z}[X]$ *be a primitive polynomial with* $f_0f_s \neq 0$ *which has a root* $\gamma \in K$.

(a) *If* $\lambda \in \Lambda_1$ *and* $\lambda(f_0f_s) = 0$, *then* $\lambda(\gamma) = 0$. *Conversely, if a prime* $p|f_0f_s$, *then there exists* $\lambda \in \Lambda_1$ *with* $\lambda(p) = 1$ *and* $\lambda(\gamma) \neq 0$. *Moreover, whenever* $\lambda \in \Lambda_1$ *and* $\lambda(\gamma) \neq 0$, *then* $|\lambda(\gamma)| \geqq 1/s$.

(b) *Suppose that* $|\lambda(\gamma)| < 1/s$ *for all* $\lambda \in \Lambda_1$ *and* $|\lambda(\gamma)| < \log s/7s^2$ *for all* $\lambda \in \Lambda_0$. *Then* γ *is a root of unity.*

Proof. (a) This follows easily from elementary properties of nonarchimedean valuations (see, for example, [**3** Section 3.4]).

(b) Part (a) shows that the first condition of γ implies that $\lambda(\gamma) = 0$ for

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

all $\lambda \in \Lambda_1$, and hence that γ is a unit in the ring of algebraic integers in K. The second condition shows that for each algebraic conjugate γ' of γ over $\mathbf{Q}$:

$$|\gamma'| < \exp\{\log s/7s^2\} \leqq 1 + \log s/6s^2.$$

Hence by [**9**] γ is a root of unity, and the lemma is proved.

Remark. It is known that the bound $\log s/7s^2$ in (b) can be replaced by one of the form $cs^{-1}(\log \log s/\log s)^3$ for some constant $c > 0$. A longstanding conjecture of D. H. Lehmer implies that a bound of the form cs^{-1} should suffice (see [**4**] for further references).

The following result is a variant of Theorem 1 of [**21**].

LEMMA 2. *Let T be the subgroup of the roots of unity in K^*, the multiplicative group of nonzero elements of K. Suppose that $\beta_1, \ldots, \beta_r \in K^*$ are multiplicatively dependent. Let $d := [K:\mathbf{Q}]$. Then there exist integers $s_1, \ldots, s_r$ not all zero such that*

$$\prod \beta_i^{s_i} \in T$$

and, for each i,

$$|s_i| < (B^r + 1)/B$$

provided

$$B \geqq d \sum_i |\lambda(\beta_i)| \quad \textit{for all } \lambda \in \Lambda_1, \quad \textit{and}$$

$$B \geqq 7d^2(\log d)^{-1} \sum_i |\lambda(\beta_i)| \textit{ for all } \lambda \in \Lambda_0.$$

Proof. First recall Minkowski's lemma (see, for example, [**6**, p. 71]). This states that if $M = [\mu_{ij}]$ is a real $r \times r$ matrix and $\nu_1, \ldots, \nu_r$ are positive real numbers such that

$$\prod \nu_i \geqq |\det M|,$$

and k is an integer, $1 \leqq k \leqq r$, then there exist integers $s_1, \ldots, s_r$ not all zero such that

$$\left|\sum_j \mu_{ij} s_j\right| \leqq \nu_i$$

for all i with strict inequality when $i \neq k$.

Now, by hypothesis, there exist integers $t_1, \ldots, t_r$ not all zero such that

$$\prod \beta_i^{t_i} = 1;$$

choose k so that

$$|t_k| = \max |t_i|.$$

Take M to be the negative of the identity matrix with its kth column replaced by $(t_1/t_k, t_2/t_k, \ldots, t_r/t_k)$. Then $|\det M| = 1$, so Minkowski's lemma shows that there exist integers $s_1, \ldots, s_r$ not all zero such that

$$|s_k t_i/t_k - s_i| < B^{-1} \quad \text{for } i \neq k \text{ and } |s_k| \leqq B^{r-1}. \tag{3.1}$$

We set

$$\eta := \prod \beta_i^{s_i}$$

and claim that $\eta \in T$; this will prove the lemma since the inequalities (3.1) show that

$$|s_i| < B^{r-1} + B^{-1} \quad \text{for all } i.$$

However, the hypotheses on the t_i show that

$$\sum_i t_i \lambda(\beta_i) = 0 \quad \text{for all } \lambda \in \Lambda,$$

and so

$$\begin{aligned} |\lambda(\eta)| &= \left| \sum_i s_i \lambda(\beta_i) \right| \\ &= \left| \sum_i (s_i - s_k t_i/t_k) \lambda(\beta_i) \right| \\ &\leqq B^{-1} \sum_i |\lambda(\beta_i)| \end{aligned}$$

with strict inequality unless $\lambda(\beta_i) = 0$ for each i. Thus, the hypothesis on B shows that

$$|\lambda(\eta)| < 1/d \text{ if } \lambda \in \Lambda_1 \quad \text{and}$$

$$|\lambda(\eta)| < \log d/7d^2 \quad \text{if } \lambda \in \Lambda_0.$$

Hence Lemma 1(b) shows that η is a root of unity as required, and Lemma 2 is proved.

If λ is any (additive) valuation defined on $\mathbf{Q}$ or some finite extension of $\mathbf{Q}$ we shall extend the definition of λ (as a function but no longer as a valuation) to $\mathbf{Q}[X]$ by:

$$\lambda\left(\sum_i h_i X^i\right) := \begin{cases} \min_i \lambda(h_i) & \text{if } \lambda \text{ is nonarchimedean} \\ \lambda\left(\sum_i |h_i|\right) & \text{if } \lambda \text{ is archimedean.} \end{cases}$$

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

We conclude this section with some elementary estimates on the values of polynomials.

LEMMA 3. *Let* $f(X) = \sum f_i X^i \in \mathbf{Z}[X]$ *be a primitive polynomial of degree* s *with* $f_0 f_s \neq 0$, *and let* $g(X) \in \mathbf{Q}[X]$ *be a polynomial of degree* $< s$ *which is relatively prime to* $f(X)$. *Define* $g^*(X)$ *as the unique polynomial of degree* $< s$ *such that*

$$g(X)g^*(X) \equiv 1 \pmod{f(X)}.$$

For each $\lambda \in \Lambda$ *define*

$$C_\lambda := \max\{\, |\lambda(g(X))|, |\lambda(g^*(X))| \,\}.$$

If γ *is a root of* $f(X)$ *lying in* K, *then*:

(a) $|\lambda(g(\gamma))| \leqq (s-1)|\lambda(\gamma)| + C_\lambda$;

(b) $|\lambda(\gamma)| \leqq \max\{\lambda(f_0), \lambda(f_s)\}$ *if* $\lambda \in \Lambda_1$ *and*

$|\lambda(\gamma)| \leqq \max_i |\lambda(1 + |f_i|)|$ *if* $\lambda \in \Lambda_0$.

Remark. (b) gives only a rough estimate for $\lambda(\gamma)$. Much more precise results are known. In the case $\lambda \in \Lambda_1$, an exact bound may be obtained easily using Newton polygons (see, for example, [**13**, p. 19]), and [**18**] deals at length with the case where $\lambda \in \Lambda_0$.

Proof. (a) Since $g(\gamma)g^*(\gamma) = 1$,

$$\lambda(g(\gamma)) = -\lambda(g^*(\gamma)).$$

Thus, if we prove that

$$\lambda(g(\gamma)) \geqq -(s-1)|\lambda(\gamma)| - C_\lambda \quad \text{whenever } \lambda(g(\gamma)) < 0,$$

then this, and the corresponding inequality for $\lambda(g^*(\gamma))$ will prove (a). However, writing $g(X) = \sum g_i X^i$ we have for all $\lambda \in \Lambda_1$ that

$$\lambda(g(\gamma)) = \lambda(\textstyle\sum g_i \gamma^i) \geqq \min_i\{\lambda(g_i) + i\lambda(\gamma)\}$$

$$\geqq -C_\lambda - (s-1)|\lambda(\gamma)|$$

because $g(X)$ has degree $< s$. Thus (a) follows in the case $\lambda \in \Lambda_1$, and a similar argument starting with the inequality

$$|\textstyle\sum g_i \gamma^i| \leqq \sum |g_i| \max\{|\gamma|^{s-1}, 1\}$$

gives the corresponding inequality when $\lambda \in \Lambda_0$.

(b) First suppose that $\lambda \in \Lambda_1$. Then, if $\lambda(\gamma) \geqq 0$, we have

$$\lambda(f_0) = \lambda\left(-\sum_{i>0} f_i \gamma^i\right) \geqq \min_{i>0}\{\lambda(f_i) + i\lambda(\gamma)\} \geqq \lambda(\gamma)$$

because each $f_i \in \mathbf{Z}$. Similarly, if $\lambda(\gamma) < 0$, then

$$\lambda(f_s) \geqq \lambda(\gamma^{-1}) = -\lambda(\gamma).$$

This proves (b) in the case $\lambda \in \Lambda_1$. The proof when $\lambda \in \Lambda_0$ is similar, based on the inequality

$$1 \leqq |f_0| = \left| \sum_{i>0} f_i \gamma^i \right| < |\gamma| (1 - |\gamma|)^{-1} \max |f_i|$$

when $|\gamma| < 1$ and a similar inequality involving powers of γ^{-1} when $|\gamma| > 1$.

4. The orbit-stabilizer problem for cyclic groups. In the case that $G = \langle b \rangle$ is a cyclic subgroup of $GL(n, \mathbf{Q})$ we can give a quite precise solution to the orbit-stabilizer problem. Since this seems an especially interesting case, we deal with it separately from the more general cases dealt with in Sections 5 and 7.

Let u and v be nonzero vectors in $\mathbf{Q}^n$, and let U be the subspace spanned by u, ub, $ub^2, \ldots$. Then $Ub = U$ and we can compute (by elementary linear algebra) the minimal polynomial of b acting on U. Suppose that the primitive polynomial

$$(4.1) \quad f(X) = f_0 + f_1 X + \ldots + f_d X^d \in \mathbf{Z}[X]$$

(where $d = \dim U$) is the minimal polynomial for b acting on U. Clearly, u and v can only lie in the same G-orbit if there exist polynomials $h(X)$, $h^*(X) \in \mathbf{Q}[X]$ such that $v = uh(b)$ and $u = vh^*(b)$. In the latter case we can choose the degrees of these polynomials to be less than d, and since

$$u\{h(b)h^*(b) - 1\} = 0$$

we have

$$h(X)h^*(X) \equiv 1 \pmod{f(X)}.$$

Again the coefficients of $h(X)$ and $h^*(X)$ can be easily computed from the data b, u and v. Now, by the definition of $f(X)$,

$$(4.2) \quad ub^t = u \text{ if and only if } X^t \equiv 1 \pmod{f(X)}$$

$$(4.3) \quad ub^t = v \text{ if and only if } X^t \equiv h(X) \pmod{f(X)}.$$

There are well-established factorization algorithms for polynomials in $\mathbf{Z}[X]$ (see, for example, [**12**]), and indeed recently it has been proved that factorization can be carried out in polynomial-time (see [**16**]). Thus we can assume that we have available the factorization of $f(X)$. Moreover, for each irreducible factor $g(X)$ of $f(X)$, we can decide quickly whether or not $g(X)$ is a cyclotomic polynomial (and, if it is, find its order). Indeed, a

monic irreducible polynomial $g(X) \in \mathbf{Z}[X]$ is a cyclotomic polynomial of order r if and only if

$$g(X) \mid X^r - 1 \quad \text{and} \quad g(X) \nmid X^s - 1$$

for any proper divisor s of r. Moreover, in this case $\deg g(X) = \varphi(r)$ (the Euler phi-function). This criterion is easily checked since, for any $r > 1$,

$$4\varphi(r) \log \log 6\varphi(r) > r$$

(see [**21**, Lemma 6]).

With the notation above, we have the following result.

THEOREM 1. (Orbit-stabilizer problem for cyclic groups).

(a) *Suppose that $f(X)$ is a product of distinct cyclotomic polynomials $g_i(X)$ of order e_i $(i = 1, \ldots, r)$, say. Then $ub^t = v$ if and only if for each i there exists an integer $t_i \in [0, e_i - 1]$ such that*

$$X^{t_i} \equiv h(X) \pmod{g_i(X)} \quad \text{and} \quad t \equiv t_i \pmod{e_i}.$$

The stabilizer $G_u = \langle b^e \rangle$ where

$$e := \mathrm{LCM}\{e_i \mid i = 1, \ldots, r\}.$$

(b) *Suppose that $f(X)$ is divisible by the square $g(X)^2$ of some irreducible polynomial. Then there is at most one integer t such that $ub^t = v$ which (if it exists) satisfies the congruence*

$$th(X) \equiv Xh'(X) \pmod{g(X)}$$

where $h'(X)$ denotes the derivative of $h(X)$. In this case, $G_u = 1$.

(c) *Suppose that $f(X)$ is divisible by a (primitive) irreducible polynomial $g(X) = \sum g_i X^i \in \mathbf{Z}[X]$ of degree s, which is not a cyclotomic polynomial. Then there is at most one integer t such that $ub^t = v$. When such a t exists it satisfies $|t| \leqq C + s - 1$ where, if $g_0 g_s \neq \pm 1$ we can take*

$$C = s \max\{ |\lambda_p(h(X))|, |\lambda_p(h^*(X))| \}$$

for any prime $p \mid g_0 g_s$ and, if $g_0 g_s = \pm 1$ we can take

$$C = 7s^2(\log s)^{-1} \max\{ |\lambda_0(h(X))|, |\lambda_0(h^*(X))| \}.$$

In this case, $G_u = 1$.

Remark. In every case at least one of (a), (b) and (c) will apply. The bound for t in (c) (the "generic" case) is polynomial in d and in the (logarithmic) size of the integers required to specify the coefficients of $h(X)$ and $h^*(X)$. The values for t in cases (a) and (b) may be much larger; but in all cases the obvious algorithms to compute t (and in case (a), to compute e as well) are polynomial-time in d and the logarithmic sizes of the integers required to specify $f(X)$, $h(X)$ and $h^*(X)$.

Proof. (a) $ub^t = v$ if and only if

$$X^t \equiv h(X) \pmod{f(X)}$$

if and only if

$$X^t \equiv h(X) \pmod{g_i(X)} \quad \text{for } i = 1, \ldots, r.$$

Since $g_i(X)$ is the cyclotomic polynomial of order e_i,

$$X^t \equiv X^{t'} \pmod{g_i(X)} \text{ if and only if } t \equiv t' \pmod{e_i};$$

and so $X^t \equiv h(X) \pmod{g_i(X)}$ if and only if there exists an integer $t_i \in [0, e_i - 1]$ such that

$$X^{t_i} \equiv h(X) \pmod{g_i(X)} \quad \text{and} \quad t \equiv t_i \pmod{e_i}.$$

This proves the first assertion. The proof that $G_u = \langle b^e \rangle$ now follows from the special case where $v = u$ and $h(X) = 1$.

(b) Since $g(X)^2$ divides $f(X)$, $ub^t = v$ implies that

$$X^t \equiv h(X) \pmod{g(X)^2}.$$

But the latter condition holds if and only if

$$X^t \equiv h(X) \pmod{g(X)} \quad \text{and} \quad tX^{t-1} \equiv h'(X) \pmod{g(X)};$$

and so implies that

$$th(X) \equiv Xh'(X) \pmod{g(X)}.$$

Since the hypotheses imply that $f(X)$ and $h(X)$ are relatively prime, $g(X)$ does not divide $h(X)$, and so there is at most one t satisfying

$$th(X) \equiv Xh'(X) \pmod{g(X)}.$$

This proves the first assertion. The fact that $G_u = 1$ follows by taking $v = u$.

(c) Let $\gamma \in \mathbf{C}$ be a root of $g(X)$ and put $K := \mathbf{Q}(\gamma)$. By hypothesis, γ is not a root of unity, and so Lemma 1(b) shows that either there is a valuation $\lambda \in \Lambda_1$ such that $|\lambda(\gamma)| \geqq 1/s$ or a valuation $\lambda \in \Lambda_0$ such that

$$|\lambda(\gamma)| \geqq \log s/7s^2.$$

On the other hand, (4.3) shows that $ub^t = v$ implies that $\gamma^t = h(\gamma)$ and hence

$$t\lambda(\gamma) = \lambda(h(\gamma)).$$

If $g_0 g_s$ is divisible by a prime p, then Lemma 1 (a) shows that we may choose λ so that its restriction to $\mathbf{Q}$ equals λ_p, whilst if $g_0 g_s = \pm 1$ then necessarily $\lambda \in \Lambda_0$ and its restriction to $\mathbf{Q}$ equals λ_0. In either case, Lemma 4(a) shows that

$$|t| = |\lambda(h(\gamma))|/|\lambda(\gamma)| \leqq s - 1 + C_\lambda/|\lambda(\gamma)|$$

and so the first assertion follows. Finally, taking $v = u$, the uniqueness shows that $G_u = 1$.

5. The orbit-stabilizer problem for abelian groups. The present section deals with the orbit-stabilizer problem for a restricted class of abelian groups (this will be extended to include all abelian groups in Section 7). Specifically, we suppose that $G = \langle b_1, \ldots, b_r \rangle$ is an abelian subgroup of $GL(n, \mathbf{Q})$ satisfying the conditions:

(P1) the group is contained in the polynomial algebra $\mathbf{Q}[b]$ for some specified $b \in GL(n, \mathbf{Q})$;

(P2) no element of the group has an eigenvalue which is a nontrivial root of unity (in particular, the group is torsionfree).

Let u and v be nonzero vectors in $\mathbf{Q}^n$ and define U to be the subspace spanned by $u, ub, ub^2, \ldots$. The vectors u and v can only lie in the same G-orbit if $v \in U$ and so we shall restrict ourselves to this case. The space U is G-invariant by (P1), so there is no loss in generality in supposing that $U = \mathbf{Q}^n$.

We shall first consider the stabilizer problem. As in Section 4 we can compute from $b_1, \ldots, b_r$ and u the following polynomials:

(i) a primitive polynomial

$$f(X) = f_0 + f_1 X + \ldots + f_n X^n \in \mathbf{Z}[X]$$

which is the minimal polynomial for b;

(ii) for $i = 1, \ldots, r$, polynomials $h_i(X), h_i^*(X) \in \mathbf{Q}[X]$ of degrees $< n$ such that.

$$b_i = h_i(b) \quad \text{and} \quad b_i^{-1} = h_i^*(b).$$

Then, since we are supposing that $U = \mathbf{Q}^n$, $G_u = 1$ and so

$$\text{(5.1)} \quad \prod b_i^{t_i} \in G_u \text{ if and only if } \prod h_i(X)^{t_i} \equiv 1 \pmod{f(X)}.$$

We shall put

$$\text{(5.2)} \quad T: = \{(t_1, \ldots, t_r) \in \mathbf{Z}^r | \prod b_i^{t_i} = 1\}.$$

Then T is a submodule of the free $\mathbf{Z}$-module $\mathbf{Z}^r$, and if we obtain a set of generators for T then we shall have a corresponding set of generators for G_u expressed as words in $b_1, \ldots, b_r$.

Recall a few facts about submodules of $\mathbf{Z}^r$. If S is any subset of $\mathbf{Z}^r$, then we denote the orthogonal complement (relative to the usual dot product) by $S^\perp$. The set $S^\perp$ is always a submodule, and

$$\operatorname{rank}(S^\perp) = r - k$$

where k is the dimension of the $\mathbf{Q}$-space $\mathbf{Q}S$ spanned by S. We always

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

have $S \subseteq S^{\perp\perp}$ and, if S itself is a submodule, then S and $S^{\perp\perp}$ have the same rank. The structure theorem for finitely generated $\mathbf{Z}$-modules shows that if S_1 and S_2 are two submodules of $\mathbf{Z}^r$ and have equal ranks, and $S_1 \subseteq S_2$, then $lS_2 \subseteq S_1$ for some integer $l \geqq 1$. Finally, we remark that the classical problem of computing a $\mathbf{Z}$-basis for $S^{\perp}$ for a finite set S has been well studied, and efficient algorithms have been developed (see, for example, [7]).

The following lemma is essentially equivalent to a particular case of our problem.

LEMMA 4. *Let* $\gamma \in \mathbf{C}$ *be algebraic and put* $K = \mathbf{Q}(\gamma)$. *Suppose that* $\langle \beta_1, \ldots, \beta_r \rangle$ *is a torsion-free subgroup of* K^* *and define the constant* B *as in Lemma* 2. *Let*

$$T := \{\mathbf{t} = (t_1, \ldots, t_r) \in \mathbf{Z}^r \mid \prod \beta_i^{t_i} = 1\}$$

$$T_0 := \{\mathbf{t} = (t_1, \ldots, t_r) \in T \mid |t_i| \leqq (B^r + 1)/B \text{ for all } i\}.$$

Then $T = T_0^{\perp\perp}$.

Remark. It remains an open question as to whether T_0 actually generates T.

Proof. We shall first show that it is enough to prove that $T^{\perp} = T_0^{\perp}$. Indeed, if the latter holds, then $T \subseteq T^{\perp\perp} = T_0^{\perp\perp}$. From above, we know that T and $T^{\perp\perp}$ have the same rank and that hence for some integer $l \geqq 1$,

$$lT_0^{\perp\perp} \subseteq T.$$

But $\langle \beta_1, \ldots, \beta_r \rangle$ is torsion-free, and so $l\mathbf{t} \in T$ implies $\mathbf{t} \in T$, and hence $T = T_0^{\perp\perp}$ as required.

We now prove by induction on r that $T^{\perp} = T_0^{\perp}$. The cases where $r \leqq 1$ or $T = 0$ are clearly true, so suppose $r > 1$ and $T \neq 0$. By Lemma 2 there exists a nonzero $\mathbf{s} = (s_1, \ldots, s_r) \in T_0$, and without loss in generality we may assume $s_r \neq 0$. Let S and S_0 be the subsets of T and T_0, respectively, which consist of all elements $(t_1, \ldots, t_r)$ with $t_r = 0$. Induction shows that $S^{\perp} = S_0^{\perp}$. Moreover, since $S \cup \{\mathbf{s}\}$ spans the same $\mathbf{Q}$-space as T does,

$$T^{\perp} = (S \cup \{\mathbf{s}\})^{\perp} = S^{\perp} \cap \{\mathbf{s}\}^{\perp} = S_0^{\perp} \cap \{\mathbf{s}\}^{\perp} \supseteq T_0^{\perp}.$$

On the other hand, $T_0 \subseteq T$ implies $T_0^{\perp} \supseteq T^{\perp}$. Thus $T^{\perp} = T_0^{\perp}$ and the lemma is proved.

Remarks. 1. The proof of Lemma 4 shows that in actual computations one can replace T_0 by a set R_0 such that $T_0^{\perp} = R_0^{\perp}$ and $|R_0| = \text{rank } T$. The first step is to search for a nonzero $\mathbf{s}$ in T_0. If none is found, then $T = 0$ by Lemma 2 and $R_0 = \emptyset$. Otherwise $\mathbf{s}$ is an element of R_0, and T_0 is

replaced by a smaller set (denoted S_0 in the proof) by restricting to vectors with a specified component equal to 0 (the bound B may also be reduced). Further elements of R_0 may then be obtained by repeating this basic step. This process eventually leads to a linearly independent set R_0 such that $R_0{}^{\perp} = T^{\perp}$. The construction of R_0 will require at most $O(B^{r(r-1)})$ verifications of a condition of the form: is $t \in T$?

2. In doing such computations it is necessary to compute an appropriate value for B. In the situation in which we shall be interested, γ will be a root of a known primitive polynomial $g(X) \in \mathbf{Z}[X]$ of degree s, and

$$\beta_i = h_i(\gamma) \quad \text{for } i = 1, \ldots, r$$

(see (ii) above). Then Lemma 3 shows that if $\lambda \in \Lambda_1$ has its restriction to $\mathbf{Q}$ equal to λ_p, then there is a uniform bound on $|\lambda(\beta_i)|$ depending only on p and i (and is equal to 0 if p does not divide $g_0 g_s$ or the numerator or denominator of any coefficient of $h_i(X)$ or $h_i^*(X)$). Similarly Lemma 3 gives a uniform bound on $|\lambda(\beta_i)|$ for all $\lambda \in \Lambda_0$. Thus computing a suitable value of B is a strictly finite problem and in general quite easy. In particular cases we may be able to obtain smaller values for B using direct information, but even from Lemma 3 we have a value of

$$B = O(s^3(\log s)^{-1} r(L + 1))$$

where L is the logarithm of the largest integer required to specify the coefficients of $g(X)$ and $h_i(X)$ and $h_i^*(X)(i = 1, \ldots, r)$.

We now turn to a solution of the stabilizer problem for the abelian group $G = \langle b_1, \ldots, b_r \rangle$ satisfying the conditions (P1) and (P2). It is enough to construct a generating set for the $\mathbf{Z}$-module T defined by (5.2). First note that $T = T^{\perp\perp}$. Indeed, $T \subseteq T^{\perp\perp}$ and both have the same rank, so $lT^{\perp\perp} \subseteq T$ for some integer $l \geqq 1$. However, since G is torsion-free by (P2), $l\mathbf{t} \in T$ implies that $\mathbf{t} \in T$, and so $T = T^{\perp\perp}$. We consider the stabilizer problem in three cases.

Case 1. ($f(X) = g(X)$ is irreducible). In this case let $\gamma \in \mathbf{C}$ be a root of $g(X)$ and put $\beta_i := h_i(\gamma)(i = 1, \ldots, r)$. Then Lemma 4 (and the remarks following it) forms a basis for an algorithm to compute a finite set of generators for T using standard techniques.

Case 2. ($f(X) = g(X)^k$ where $g(X)$ is irreducible and $k > 1$). In this case (5.1) shows that $(t_1, \ldots, t_r) \in T$ if and only if

$$\prod h_i(X)^{t_i} \equiv 1 \pmod{g(X)^k}.$$

Since

$$\{\prod h_i(X)^{t_i} - 1\}' = \prod h_i(X)^{t_i} \sum t_i h_i'(X)/h_i(X)$$

we conclude that $(t_1, \ldots, t_r) \in T$ if and only if

$$\prod h_i(X)^{t_i} \equiv 1 \pmod{g(X)} \text{ and } \sum t_i \bar{h}_i(X) = 0$$

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

where $\bar{h}_i(X)$ is the polynomial of degree less than deg $g(X)^{k-1}$ satisfying

$$h_i'(X) \equiv h_i(X)\bar{h}_i(X) \pmod{g(X)^{k-1}}.$$

Thus with the notation of Case 1, $T = T_1 \cap T_2$ where

$$T_1 := \{ (t_1, \ldots, t_r) \mid \prod \beta_i^{t_i} = 1 \} \text{ and}$$

$$T_2 := \{ (t_1, \ldots, t_r) \mid \sum t_i \bar{h}_i(X) = 0 \}.$$

It is evident that $T_2 = R^{\perp}$ where R is a finite set of vectors in $\mathbf{Z}^r$ which can be written down from a knowledge of the coefficients of the $\bar{h}_i(X)$. A basis for T_1 can be computed as in Case 1, and so we can compute a basis for T via

$$T = T^{\perp\perp} = (T_1^{\perp} \cup R)^{\perp}.$$

Case 3. (General case: $f(X)$ has the canonical factorization $\prod g_j(X)^{k_j}$). In this case we can compute bases for each of the $\mathbf{Z}$-modules

$$S_j := \{ t_1, \ldots, t_r) \mid \prod h_i(X)^{t_i} \equiv 1 \pmod{g_j(X)^{k_j}} \}$$

using Cases 1 and 2. Then

$$T = \bigcap_j S_j$$

by (5.1) and so a basis for T can be computed via

$$T = T^{\perp\perp} = \left\{ \bigcup_j S_j^{\perp} \right\}^{\perp}.$$

This completes the solution to the stabilizer problem.

The solution of the orbit problem for G can be reduced to the solution of the stabilizer problem. If v is in the same G-orbit as u, then necessarily we can compute

(iii) two polynomials $h_0(X)$, $h_0^*(X) \in \mathbf{Q}[X]$ of degrees $< n$ such that

$$v = uh_0(b) \quad \text{and} \quad u = vh_0^*(b).$$

We can restrict ourselves to this case, and note that since $U = \mathbf{Q}^n$ we have $b_0 := h_0(b)$ is invertible with $b_0^{-1} = h_0^*(b)$, and that $b_0 \in G$ if and only if u and v lie in the same G-orbit. On the other hand, if $b_0 \in G$, then the following method will find $(s_1, \ldots, s_r)$ such that $b_0 = \prod b_i^{s_i}$, whilst otherwise it will break down at some stage and hence show that u and v lie in distinct orbits.

Set $\tilde{G} := \langle b_0, b_1, \ldots, b_r \rangle$ and attempt to apply the algorithm above to find the stabilizer for u under $\tilde{G}$. Since $\tilde{G}$ may not satisfy the condition (P2), it is possible that the algorithm may break down; we then know that u and v are in distinct G-orbits. If it does not break down we obtain a putative basis R for the $\mathbf{Z}$-module

$$\tilde{T}: = \left\{ (t_0, \ldots, t_r) \in \mathbf{Z}^{r+1} \,\middle|\, \prod_{i=0}^{r} b_i^{t_i} = 1 \right\}.$$

If $b_0 \in G$, then $\tilde{T}$ contains a vector whose initial component is 1, and that implies that the greatest common divisor of the initial components of the vectors in R must be 1. On the other hand, if the latter condition holds, then we can compute an integral linear combination of the vectors in R equal to a vector of the form $(1, -s_1, \ldots, -s_r)$, and then $b_0 \in G$ only if

$$b_0 = \prod b_i^{s_i}.$$

This completes the solution to the orbit problem.

In general, abelian linear groups do not satisfy (P1). For example, if $n = 2m$, then $GL(n, \mathbf{Q})$ contains a finitely generated group of unipotent upper-triangular matrices which is abelian and generates a $\mathbf{Q}$-algebra of dimension $m^2 + 1$. On the other hand every $\mathbf{Q}$-algebra of the form $\mathbf{Q}[b]$ has dimension at most n. The following lemma is therefore of interest. It shows that any abelian group generated by a finite number of semisimple elements satisfies (P1), and "almost all" linear combinations of the generators give a suitable b. This can be used as a basis for an algorithm to find such a matrix (the condition that $b \in GL(n, \mathbf{Q})$ can be satisfied by adding a suitable scalar multiple of 1).

LEMMA 5. *Let $G = \langle b_1, \ldots, b_r \rangle$ be an abelian subgroup of $GL(n, \mathbf{Q})$ whose generators b_i are all semisimple. Let Δ be a subset of $\mathbf{Q}$ containing more than $n(n-1)/2$ elements and put*

$$m: = |\Delta| - n(n-1)/2.$$

Then for at least m^r of the r-tuples $(\delta_1, \ldots, \delta_r) \in \Delta^r$ the element $b: = \sum \delta_i b_i$ has the property that $G \subseteq \mathbf{Q}[b]$.

Proof. An obvious induction on r shows that it is enough to prove:

(5.3) if x and y are commuting semisimple $n \times n$ matrices over $\mathbf{Q}$, then for at least m values of $\delta \in \Delta$ the matrix $x + \delta y$ is semisimple and $x, y \in \mathbf{Q}[x + \delta y]$.

To prove (5.3) we first note that because x and y commute and are semisimple, a well known theorem of Schur shows that they are simultaneously diagonalizable over $\mathbf{C}$. Thus, for some $c \in GL(n, \mathbf{C})$,

$$c^{-1}xc = \operatorname{diag}(\xi_1, \ldots, \xi_n) \quad \text{and}$$

$$c^{-1}yc = \operatorname{diag}(\eta_1, \ldots, \eta_n),$$

say, and so

$$c^{-1}(x + \delta y)c = \operatorname{diag}(\xi_1 + \delta\eta_1, \ldots, \xi_n + \delta\eta_n).$$

This shows that $x + \delta y$ is semisimple, and Lagrange's interpolation formula shows that $x, y \in \mathbf{C}[x + \delta y]$ (and so $x, y \in \mathbf{Q}[x + \delta y]$ because $\delta \in \mathbf{Q}$) provided:

(5.4) $\xi_i + \delta\eta_i \neq \xi_j + \delta\eta_j$ whenever $(\xi_i, \eta_i) \neq (\xi_j, \eta_j)$.

But the condition (5.4) can be violated by at most $n(n-1)/2$ scalars δ, and so (5.3) follows. This proves the lemma.

6. Connectedness of principal congruence subgroups. To study the orbit-stabilizer problem for a more general class of groups we shall need some properties about Zariski-connectedness in linear groups which may be of independent interest. For elementary properties of the Zariski topology in linear groups see [**8**] or [**24**].

LEMMA 6. *Let F be an arbitrary field, G be any subgroup of $GL(n, F)$, and $\bar{G}$ be the Zariski-closure of G in $GL(n, F)$. Then G is connected if and only if $\bar{G}$ is connected.*

Proof. It is well known, and easily proved, that under any topology the closure of a connected set is connected. It is the proof of the reverse implication, namely G must be connected if $\bar{G}$ is, which requires special properties of the Zariski topology. However, in this topology we know that the connected component of 1 in G is a normal, open and closed subgroup G^0 of finite index. Hence $\bar{G}$ is the union of the closures $\overline{G^0 a}$ of finitely many cosets $G^0 a$ of G^0 in G. Moreover, in the Zariski topology a connected group is irreducible as a topological space (see [**24**, Lemma 14.3]), and so $\bar{G} = \overline{G^0 a}$ for some $a \in G$. Thus $G = G^0 a$ because $G^0 a$ is closed in G, and so $G = G^0$ is connected. This proves the lemma.

We now restrict F to be a finite extension of the field $\mathbf{Q}_p$ of the p-adic numbers for some prime p. Let D denote the local ring in F and πD be the maximal ideal in D.

LEMMA 7. *Let $x \in GL(n, F)$ and suppose that the eigenvalues $\xi_1, \ldots, \xi_n$ of x all lie in D and that $\xi_i - 1 \in pD$ for each i. If $p > 2$, then $\langle x \rangle$ is connected in the Zariski topology.*

Remark. Suppose that F, and hence D, contains a primitive pth root ζ of 1. Put $x = \mathrm{diag}(1, \zeta)$. Then $\langle x \rangle$ is a nontrivial finite group and hence not connected. In particular, this shows that the hypothesis "$p > 2$" in Lemma 7 cannot be dropped. Moreover, if $p > 2$, then

$$(\zeta - 1)^p \equiv \zeta^p - 1 = 0 \pmod{p},$$

and so $\zeta - 1 \in \pi D$. This shows that the hypothesis "$\xi_i - 1 \in pD$" cannot be weakened to "$\xi_i - 1 \in \pi D$".

Proof. Firstly, consider the case where x is semisimple; we can assume without loss of generality that

$$x = \operatorname{diag}(\xi_1, \ldots, \xi_n).$$

By [**8**, Lemma 8.8 B] or [**2**, Proposition 7.2] we know that the ideal I consisting of all polynomials

$$w(X_1, \ldots, X_n) \in F[X_1, \ldots, X_n]$$

such that $w(\xi_1, \ldots, \xi_n) = 0$ has a basis consisting of polynomials of the form

$$X_1^{i_1} \ldots X_n^{i_n} - X_1^{j_1} \ldots X_n^{j_n}.$$

If $\langle x \rangle$ were not connected, then its connected component of 1 would be a proper subgroup of finite index, and then for some integer $h \geqq 1$, $\langle x^h \rangle$ would have its connected component of 1 with prime index. Thus, it is enough to show that the connected component of 1 in $\langle x \rangle$ does not have prime index.

Suppose the contrary. Then $\langle x^q \rangle$ is a closed connected subgroup of index q in $\langle x \rangle$ for some prime q. Thus there exists a polynomial of the form

$$w(X_1, \ldots, X_n) = X_1^{i_1} \ldots X_n^{i_n} - X_1^{j_1} \ldots X_n^{j_n}$$

such that

$$w(\xi_1^q, \ldots, \xi_n^q) = 0 \text{ but } w(\xi_1, \ldots, \xi_n) \neq 0.$$

Putting $k_t := i_t - j_t$, this implies that

$$\prod \xi_i^{k_i} = \zeta,$$

say, where ζ is a primitive qth root of unity. By hypothesis,

$$\xi_i \equiv 1 \pmod{p} \quad \text{for each } i$$

and so

$$\zeta^j \equiv 1 \pmod{p} \text{ for all } j.$$

Thus

$$q = \prod_{j=1}^{q-1} (1 - \zeta^j) \equiv 0 \pmod{p^{q-1}}.$$

Hence $p = q$ and $q = 2$ contrary to the hypothesis that $p > 2$. This proves the lemma in the case that x is semisimple.

In the general case we can write $x = x_s x_u = x_u x_s$ where x_s and x_u are the semisimple and unipotent parts of x and both lie in $GL(n, F)$ (see [**24**, Theorem 7.2]). Since F has characteristic 0, $\langle x_u \rangle$ is connected (see [**2**,

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

Proposition 8.1]), and $\langle x_s \rangle$ is connected by what we have just proved. Thus $H := \langle x_s \rangle \times \langle x_u \rangle$ is also connected. However, it is known that x_s and x_u both lie in the Zariski-closure $\overline{\langle x \rangle}$ of $\langle x \rangle$, and so $H \subseteq \overline{\langle x \rangle} \subseteq \bar{H}$. Lemma 6 now shows that the connectedness of H implies successively that $\bar{H}$, $\overline{\langle x \rangle}$ and $\langle x \rangle$ are all connected. This proves the lemma.

LEMMA 8. *With the notation above, define*

$$C := \{x \in GL(n, D) \mid x - 1 \in pMat(n, D)\}$$

(the principal p-congruence subgroup of $GL(n, D)$*). If* $p > 2$ *then every subgroup* G *of* C *is connected.*

Proof. If E is the algebraic closure of F, then the Zariski-closure of G in $GL(n, E)$ is an algebraic group. Hence Lemma 6 and [**24**, Corollary 14.15] show that G is connected provided every cyclic subgroup of G is connected. Thus it is enough to show that every $x \in C$ generates a connected group.

Let $x \in C$. Replacing F by a finite extension of itself if necessary, we may assume that the eigenvalues $\xi_1, \ldots, \xi_n$ of x lie in F. We want to show that the hypotheses of Lemma 7 hold for x. Since $x = 1 + py$ for some $y \in Mat(n, D)$, we have

$$(6.1) \quad 0 = \det(x - \xi 1)$$
$$= (1 - \xi)^n + p \operatorname{tr} y(1 - \xi)^{n-1} + \ldots + p^n \det y$$

for each eigenvalue $\xi = \xi_i$. Since $\xi \in F$ and (6.1) shows that ξ is integral over D, $\xi \in D$. If $\xi = 1$, then $\xi - 1 \in pD$ certainly holds. If $\xi \neq 1$, then there is an integer $s \geqq 0$ and a unit $\beta \in D$ such that $\xi - 1 = \beta\pi^s$. We also know that $p = \alpha\pi^e$ for some integer $e \geqq 1$ and some unit $\alpha \in D$. This, together with (6.1), shows that

$$sn \geqq \min\{ie + (n - i)s \mid i = 1, \ldots, n\}$$

and hence $e \leqq s$. Thus $\xi - 1 \in pD$ in all cases, and Lemma 7 now completes the proof.

LEMMA 9. *Let* G *be a subgroup of* C *where* C *is defined in Lemma* 8. *Then*:

(a) *no element of* G *has a nontrivial root of unity as an eigenvalue*;

(b) *if* G *has a nilpotent subgroup of finite index then* G *is nilpotent*;

(c) *if* G *has a solvable subgroup of finite index, then* G' *is a group of unipotent matrices.*

Proof. (a) If ξ is an eigenvalue of $x \in G$, then $\xi \in D$ and $\xi - 1 \in pD$ by the proof of Lemma 8. Thus Lemma 7 shows that $\langle \xi \rangle \subseteq GL(1, D)$ is torsionfree.

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

(b) Let N be a nilpotent subgroup of finite index in G. Then the closure $\bar{N} \cap G$ in G is also nilpotent (see [**8**, Lemma 8.6 B]), and is a closed subgroup of finite index in G. Hence $\bar{N} \cap G \supseteq G^0$ and so $\bar{N} \cap G = G$ because G is connected by Lemma 8.

(c) This follows from [**24**, Theorem 5.8] since G is connected (and remains connected under extensions of the ground field).

7. The orbit-stabilizer problem for nilpotent-by-finite groups. We finally consider the case where $G = \langle x_1, \ldots, x_r \rangle \subseteq GL(n, \mathbf{Q})$ is known to be nilpotent-by-finite (that is, have a nilpotent subgroup of finite index), and u and v are nonzero vectors in $\mathbf{Q}^n$. We shall show how to reduce the orbit-stabilizer problem for G, u, v to corresponding problems for abelian groups of the type considered in Section 5.

Firstly, let R denote the subring of $\mathbf{Q}$ generated by the entries of the matrices $x_1, \ldots, x_r, x_1^{-1}, \ldots, x_r^{-1}$, so $G \subseteq GL(n, R)$. If p is the smallest odd prime which does not divide the denominator of any of these entries, then R has a natural embedding into $\mathbf{Z}_p$, the p-adic integers, and we may consider $G \subseteq GL(n, \mathbf{Z}_p)$. Since $\mathbf{Z}_p/p\mathbf{Z}_p \simeq \mathbf{Z}/p\mathbf{Z}$, there is a group homomorphism

$$\psi: G \to GL(n, \mathbf{Z}/p\mathbf{Z})$$

given by $\psi(x) := x \bmod p$; and the kernel H is contained in the principal p-congruence subgroup of $GL(n, \mathbf{Z}_p)$. In particular, Lemma 9 shows that H satisfies condition (P2) of Section 5 and also:

(P3) the group is nilpotent; and

(P4) the derived group is unipotent.

The constructive nature of ψ permits us to obtain a set T of coset representatives for H in G, and then the Schreier construction enables us to write down a set of generators for H, namely the elements $x_i t s^{-1}$ with $i = 1, \ldots, r$ and $s, t \in T$ such that $Hx_i t = Hs$ (in practice, we may expect a much smaller generating set to be obtained; compare [**5**]). If the orbit-stabilizer problem is now solved for H, then the solution to the orbit-stabilizer problem for G, u, v can be obtained as follows. Compute

$$S := \{s \in T \mid u, us \text{ lie in the same } H\text{-orbit}\}$$

and for each $s \in S$ find $h(s) \in H$ such that $u = ush(s)$. Then u and v lie in the same G-orbit if and only if for some $t \in T$, ut and v lie in the same H-orbit; and

$$G_u = \bigcup_{s \in S} sh(s)H_u.$$

Thus we shall specialize to the case where $G = \langle x_1, \ldots, x_r \rangle$ satisfies the conditions (P2), (P3) and (P4), and proceed inductively on n to show how to solve the orbit-stabilizer problem for G, u, v.

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

Using elementary linear algebra, we can compute a basis for the subspace $W \subseteq \mathbf{Q}^n$ consisting of all $w \in \mathbf{Q}^n$ such that

$$w(x_i x_j - x_j x_i) = 0 \quad \text{for } 1 \leqq i < j \leqq r.$$

Clearly W is a G-subspace and $w \in W$ if and only if $wx = w$ for all $x \in G'$, so G acts as an abelian group on W.

Set $W_0 := W$. For $i = 1, \ldots, r$ we can compute the minimal polynomial of x_i acting on W_{i-1}, find an irreducible factor, say $g_i(X)$, of this polynomial, and set

$$W_i := \{w \in W_{i-1} | wg_i(x_i) = 0\}.$$

At each step W_i is a nonzero subspace of W_{i-1}, and it is G-invariant because G acts as an abelian group on W. In particular, W_r is a nonzero G-subspace on which each x_i has an irreducible minimal polynomial $g_i(X)$ and so is semisimple. Thus (see Lemma 5) we can compute $b \in GL(W_r)$ such that the $\mathbf{Q}$-space spanned by the restriction of G to W_r is equal to $\mathbf{Q}[b]$. We can then compute the minimal polynomial for b on W_r, choose an irreducible factor, say $g(X)$, of this polynomial, and find $w \neq 0$ such that $wg(b) = 0$. Since b commutes with the action of G on W_r, the subspace V spanned by w, wb, $wb^2, \ldots$ is a nonzero G-subspace. Since $g(X)$ is irreducible, $\langle b \rangle$ acts irreducibly on V, and so V is an irreducible G-subspace. Let

$$d := \dim V = \deg g(X).$$

Remark. It seems to be an open problem as to whether there is a good constructive way of finding an irreducible subspace for an arbitrary finitely generated linear group.

We now consider three cases.

Case 1. ($d = n$). In this case G is abelian and satisfies conditions (P1) and (P2) and so the orbit-stabilizer problem can be solved using the methods of Section 5.

Case 2. ($d = n - 1$). In this case $\dim(\mathbf{Q}^n/V) = 1$, so G acts trivially on $\mathbf{Q}^n/V$. Using a basis for V extended to a basis for $\mathbf{Q}^n$ we can compute $c \in GL(n, \mathbf{Q})$ such that, for all $x \in G$,

$$(7.1) \quad c^{-1}xc = \begin{bmatrix} \rho(x) & 0 \\ \tau(x) & 1 \end{bmatrix}$$

where ρ is an irreducible representation of G of degree d with $G' \subseteq \ker \rho$, and $\tau(x)$ is a $1 \times d$ matrix block. Suppose, firstly, that there exists $z \neq 1$ in the centre of G such that

$$(7.2) \quad c^{-1}zc = \begin{bmatrix} 1 & 0 \\ w & 1 \end{bmatrix}$$

https://doi.org/10.4153/CJM-1985-015-4 Published online by Cambridge University Press

for some $w \neq 0$. We claim that then $G \subseteq \mathbf{Q}[z]$ and, in particular, G is abelian. Indeed, $xz = zx$ implies that $w\rho(x) = w$ for all $x \in G$, and so the irreducibility of ρ implies Im $\rho = 1$ and $d = 1$. Now it is easy to verify that $G \subseteq \mathbf{Q}[z]$. Note that the existence of such z is easily recognized; such a z exists if and only if $d = 1$ and $G \neq 1$, and in these circumstances any nontrivial element of G is suitable. Now suppose that there is no central element $z \in G$ for which $c^{-1}zc$ has the form (7.2). This implies that ker $\rho = 1$ and, in particular, the (irreducible) minimal polynomial $g_i(X)$ of x_i acting on V is equal to $X - 1$ only if $x_i = 1$. Thus, for each nontrivial x_i, the minimal polynomial for x_i on $\mathbf{Q}^n$ has distinct roots, and so x_i is semisimple. Hence in either case G is an abelian group which satisfies conditions (P1) and (P2), and so the orbit-stabilizer problem can be solved using the methods of Section 5.

Case 3. ($1 \leqq d < n - 1$). Using a basis for V extended to a basis for $\mathbf{Q}^n$ we can find $c \in GL(n, \mathbf{Q})$ such that, for all $x \in G$,

$$(7.3) \quad c^{-1}xc = \begin{bmatrix} \rho(x) & 0 \\ \tau(x) & \sigma(x) \end{bmatrix}$$

where ρ is an irreducible representation of G of degree d with $G' \subseteq \ker \rho$, σ is a representation of G of degree $n - d > 1$, and $\tau(x)$ is an $(n - d) \times d$ block. We write

$$uc = (u_1 \quad u_2) \quad \text{and} \quad vc = (v_1 \quad v_2),$$

partitioned into vectors of lengths d and $n - d$. If $x \in G$, then

$$(7.4) \quad ux = u \Leftrightarrow u_1\rho(x) + u_2\tau(x) = u_1 \text{ and } u_2\sigma(x) = u_2$$

$$(7.5) \quad ux = v \Leftrightarrow u_1\rho(x) + u_2\tau(x) = v_1 \text{ and } u_2\sigma(x) = v_2.$$

Since Im σ evidently satisfies (P2), (P3) and (P4) and deg $\sigma < n$, we may assume that we can solve the orbit-stabilizer problem for Im σ, u_2, v_2. Thus we can find a generating set for the set of all words in $x_1, \ldots, x_r$, $x_1^{-1} \ldots, x_r^{-1}$ which lie in

$$G^* := \{x \in G | u_2\sigma(x) = u_2\},$$

and determine whether u_2 and v_2 lie in the same orbit under Im σ and, if so, find $x_0 \in G$ such that

$$u_2\sigma(x_0) = v_2.$$

In what follows we shall assume the existence of x_0 since otherwise we know that u and v cannot lie in the same G-orbit.

Now define $\psi: G^* \to GL(d + 1, \mathbf{Q})$ by

$$(7.6) \quad \psi(x) := \begin{bmatrix} \rho(x) & 0 \\ u_2\tau(x) & 1 \end{bmatrix}.$$

It is easily verified that ψ is a representation of G^* and that the conditions

(P2), (P3) and (P4) hold for Im ψ. Moreover, (7.4) and (7.5) show that, if $x \in G$, then

$$(7.7)\quad ux = u \Leftrightarrow x \in G^* \text{ and } (u_1 \quad 1)\psi(x) = (u_1 \quad 1)$$

$$(7.8)\quad ux = v \Leftrightarrow xx_0^{-1} \in G^* \text{ and } (u_1 \quad 1)\psi(xx_0^{-1})$$
$$= (v_1\rho(x_0^{-1}) + v_2\tau(x_0^{-1}) \quad 1).$$

Since $d < n - 1$, $\deg \psi < n$ and so we may assume that we can solve the orbit-stabilizer problem for Im ψ. Then (7.7) and (7.8) show that we can solve the problem for G.

This completes the solution of the orbit-stabilizer problem for nilpotent-by-finite groups.

Remark. It remains an open question as to whether similar methods might be used to solve the orbit-stabilizer problem for the class of polycyclic-by-finite linear groups. In one sense the latter class is a natural limit since it can be shown using Tits' theorem ([**24**, Theorem 10.16]) that a finitely generated linear group is polycyclic-by-finite if and only if each subgroup is finitely generated (compare with Example 6 of Section 2).

Acknowledgement. The author wishes to acknowledge the kind hospitality of the Mathematics Institute at the University of Warwick during the period when this paper was written.

References

1. G. Baumslag, F. B. Cannonito and C. F. Miller III, *Computable algebra and group embeddings*, J. Algebra *69* (1981), 186-212.
2. A. Borel, *Groupes linéaires algebraiques*, Ann. of Math. *64* (1956), 20-82.
3. Z. I. Borevich and I. R. Shafarevich, *Number theory* (Academic Press, New York, 1966).
4. D. W. Boyd, *Speculations concerning the range of Mahler's measure*, Can. Math. Bull. *24* (1981), 453-469.
5. G. Butler and C. C. Cannon, *Computing in permutation and matrix groups I*, Math. Comp. *37* (1982), 663-670.
6. J. W. S. Cassels, *An introduction to the geometry of numbers* (Springer-Verlag, Berlin, 1971).
7. T.-W. J. Chou and G. E. Collins, *Algorithms for the solution of systems of linear diophantine equations*, SIAM J. Comput. *11* (1982), 687-708.
8. J. D. Dixon, *The structure of linear groups* (Van Nostrand Reinhold, London, 1971).
9. E. Dobrowolski, *On the maximal modulus of conjugates of an algebraic integer*, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. *26* (1978), 291-292.
10. F. Grunewald and D. Segal, *Some general algorithms I: arithmetic groups*, Ann. of Math. *112* (1980), 531-583.
11. Ju. A. Ignatov, *Free and nonfree subgroups of $PSL_2(C)$ that are generated by two parabolic elements*, Mat. Sb. (NS) *106* (148) (1978), 372-379 = Math. USSR-Sb. *35* (1978), 49-56.
12. E. Kaltofen, *Factorization of polynomials*, in *Computer algebra, symbolic and algebraic computation* (Springer-Verlag, New York, 1982), 95-114.

13. N. Koblitz, *P-adic analysis: a short course on recent work* (LMS Lecture Notes *46*, C.U.P., Cambridge, 1980).
14. V. M. Kopytov, *Solvability of the occurence problem in finitely generated solvable groups of matrices over an algebraic number field*, Algebra i Logika *7* (1968), 53-63 (Russian).
15. S. Lang, *Algebra* (Addison-Wesley, Massachusetts, 1967).
16. A. K. Lenstra, H. W. Lenstra Jr. and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Annalen *261* (1982), 515-534.
17. R. C. Lyndon and J. L. Ullman, *Groups generated by two parabolic linear fractional transformations*, Can. J. Math. *21* (1969), 1388-1403.
18. M. Marden, *The geometry of zeros of a polynomial in a complex variable* (Amer. Math. Soc., New York, 1949).
19. K. A. Mihailova, *The occurence problem for direct products of groups*, Dokl. Akad. Nauk SSSR *119* (1958), 1103-1105 (Russian).
20. C. F. Miller III, *On group-theoretic decision problems and their classification* (Princeton U. P., Princeton, 1971).
21. A. J. van der Poorten and J. H. Loxton, *Multiplicative relations in number fields*, Bull. Austral. Math. Soc. *16* (1977), 83-98; errata, ibid *17* (1977), 151-155.
22. R. A. Sarkisjan, *Algorithmic problems for linear algebraic groups I & II*, Mat. Sb. (NS) *113* (*155*) (1980), 179-216 and 400-436.
23. H. S. Shanks, *The rational case of a matrix problem of Harrison*, Discrete Math. *28* (1979), 207-212.
24. B. A. F. Wehrfritz, *Infinite linear groups* (Springer-Verlag, New York, 1973).

Carleton University,
Ottawa, Ontario