JOURNAL OF ALGEBRA 34, 292–299 (1975)

The Burnside Problem for Semigroups*

ROBERT MCNAUGHTON†

Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, New York 12181

AND

YECHEZKEL ZALCSTEIN

Department of Computer Science, State University of New York, Stony Brook, New York 11790

Communicated by I. N. Herstein

Received December 1, 1973

We show that the strong Burnside problem has an affirmative answer for semigroups of finite dimensional matrices over a field. As a corollary of this result and the proof of a theorem of Procesi, it follows that a torsion semigroup embeddable in the multiplicative semigroup of an algebra over a field satisfying a polynomial identity is locally finite. We prove, more generally, that a torsion semigroup of matrices over a skew field all of whose subgroups are locally finite is locally finite.

1. INTRODUCTION

The strong Burnside Problem is the question whether every torsion group is locally finite. The weak Burnside Problem is the question whether every torsion group, in which the order of every element is less than a fixed integer, is locally finite. Both questions have obvious generalizations to semigroups.

The weak Burnside Problem has been shown to have a negative answer by Novikov and Adjan [10]. The corresponding problem for semigroups had been answered in the negative many years earlier by Morse and Hedlund [9], who exhibited an infinite semigroup S having three generators and satisfying $x^2 = 0$ for all x in S (see also [20]).

* The results of this paper were presented at the Cambridge (Massachusetts) meeting of the American Mathematical Society, October 27, 1973.

† Supported by the National Science Foundation, under NSF GJ 35293.

For groups having faithful representation by matrices over a field, however, the strong Burnside Problem has an affirmative answer. This has been shown many years ago by Schur [13], who used a result of Burnside. Schur's theorem, for matrices over the field of the complex numbers, was extended by Kaplansky [4] to matrices over arbitrary fields. Zalcstein [17] has proved a partial generalization of Schur's theorem to semigroups. In this paper we remove the restriction imposed in [17] and prove the full generalization of Kaplansky's extension to semigroups.

Our proof is combinatorial and thus can be extended slightly to apply to semigroups of matrices over skew fields, provided that all subgroups are locally finite. Processi [11] and Herstein (unpublished) have generalized the Burnside–Schur Theorem to groups that are embeddable in a ring that satisfies a polynomial identity. As a corollary of our theorem and of Procesi's proof of a special case of the Procesi–Herstein Theorem, it follows that a torsion semigroup embeddable in an algebra over a field satisfying a polynomial identity is locally finite.

2. The Theorem

A semigroup S is *torsion* (or *periodic*) if, for all x in S, the subsemigroup generated by x is finite, or, equivalently, there are positive integers m, r, such that $x^{m+r} = x^m$. S is *locally finite* if all its finitely generated subsemigroups are finite.

Burnside–Schur Theorem. *Any torsion group of n-by-n matrices over a field is locally finite.*

For a proof see [2, p. 66] or [4, p. 105]. It is to be noted that we are using the version of the theorem as strengthened by Kaplansky to arbitrary fields. Our theorem is the following generalization of this theorem to semigroups:

Theorem. *Any torsion semigroup of n-by-n matrices over a field is locally finite.*

Proof. We will need some word-set notation, more familiar to automata theorists than to algebraists (see McNaughton and Yamada [8]). Let X be a finite nonempty set and let X^* denote the free monoid generated by X. We consider the elements of X^* as words over X. The identity element of X^* is the null word and is denoted by λ. Let A and B be subsets of X^*; then

$$AB = \{vw \colon v \in A,\, w \in B\},$$

where juxtaposition denotes concatenation of words. A^k is defined recursively as

$$A^0 = \{\lambda\}, \qquad A^{k+1} = AA^k,$$

$$A^* = \bigcup_{k=0}^{\infty} A^k.$$

Let S be a monoid generated by X. Then there is a surjective homomorphism $\phi: X^* \to S$. Note that $\phi(A \cup B) = \phi(A) \cup \phi(B)$; $\phi(AB) = \phi(A)\,\phi(B)$; if A is not the empty set then $\phi(A^*)$ is the submonoid of S generated by $\phi(A)$; and $\phi(\lambda)$ is the identity element. An element of X^* will be denoted by w. If $w = w_1w_2$ then w_1 is a *prefix* of w and w_2 is a *suffix* of w.

LEMMA. *A torsion semigroup of invertible linear transformations on a vector space is a group.*

Proof. If $t^{m+r} = t^m$ then t^q is an idempotent for some q. If t is invertible then t^q is invertible and must be the identity. t^{q-1} is the inverse of t (which is t itself if $q = 1$), and so the semigroup is a group.

(We note that this lemma is valid if the vector space is over a skew field.)

We can now proceed to prove the theorem. Let S be the given torsion semigroup and let T be a subsemigroup of S having a finite set of generators, X. By adjoining an identity element if necessary, we may assume that T is a monoid.

We prove that T is finite by proving the following assertion by induction on h: T has finitely many distinct matrices of rank $n - h$.

The case $h = 0$ follows easily from the Burnside–Schur Theorem: Let T' be the subset of T consisting of the matrices of rank n, i.e., the invertible matrices. Clearly T' is a subsemigroup of T, and since T is torsion so is T'. By virtue of the one-to-one correspondence between linear transformations and matrices, the Lemma applies and T' is a group. Furthermore, T' is generated by the subset of X consisting of the invertible matrices (since a product of matrices is invertible if and only if each of the factors is invertible). Thus T' is finitely generated as a semigroup and, a fortiori, also as a group, and is therefore finite by the Burnside–Schur Theorem.

Assume now that the set of matrices of rank greater than $n - h$ in T is finite. We must prove that there are only finitely many matrices in T of rank $n - h$. We prove first that among the elements of T there are only finitely many distinct ranges of dimension $n - h$. (The argument holds more generally for matrices over skew fields.) Let us say that a word w in X^* has ***rank r*** if $\phi(w)$ has rank r, where ϕ is the surjective homomorphism onto T. Let w be a word of rank $n - h$. Factor w as $w_1w_2w_3$, where w_2w_3 is the shortest suffix of w of rank $n - h$ and w_2 is the shortest prefix of w_2w_3 of rank $n - h$. Note that w_2 cannot be the null word, but w_1 or w_3 or both may be null.

The range of $w = w_1w_2w_3$ is included in the range of w_2w_3, by the nature of the corresponding linear transformations. But since $\text{rank}(w) = \text{rank}(w_2w_3)$, $\text{range}(w) = \text{range}(w_2w_3)$. Thus w_1 is immaterial as far as the range is concerned. Since no proper prefix of w_2 is of rank $n - h$, $w_2 = w_2'x$, where $x \in X$ and w_2' (possibly null) is of rank greater than $n - h$. By the induction hypothesis, as we vary w over all words of rank $n - h$, there are only finitely many distinct $\phi(w_2')$. Since X is also finite, it follows that there are only finitely many $\phi(w_2)$. Furthermore, since no proper suffix of w_2w_3 is of rank $n - h$, w_3 has a rank greater than $n - h$. By the induction hypothesis, as w varies over all words of rank $n - h$, there are only finitely many distinct $\phi(w_3)$. It follows that there are only finitely many distinct $\phi(w_2w_3)$ and, consequently, there are only finitely many ranges of dimension $n - h$.

Let $V_1,..., V_p$ be all the ranges of dimension $n - h$. Define a partial function $f: \{V_1,..., V_p\} \times X \to \{V_1,..., V_p\}$ by $f(V_i, x) = V_j$ if $\phi(x)$ maps V_i onto V_j, for some j; otherwise let $f(V_i, x)$ be undefined. Note that if $w = x_1 \cdots x_q$, each $x_k \in X$ and $\phi(w)$ maps V_i onto V_j, then there must be a sequence $V_{k_0}, V_{k_1},..., V_{k_q}$ such that $V_{k_0} = V_i$, $V_{k_q} = V_j$ and for all m, $0 \leqslant m \leqslant q - 1, f(V_{k_m}, x_{m+1}) = V_{k_{m+1}}$.

For the remainder of the argument, it will be convenient to think of $V_1,..., V_p$ as nodes of a graph, with a directed branch from V_i to V_j labeled x whenever $f(V_i, x) = V_j$, for x in X. If $w = x_1 \cdots x_q$, and there is a sequence $V_{k_0},..., V_{k_q}$ such that $V_{k_{m+1}} = f(V_{k_m}, x_{m+1})$ for $0 \leqslant m \leqslant q - 1$, then we say that w takes V_{k_0} *to* V_{k_q} and *through* $V_{k_1}, V_{k_2},..., V_{k_{q-1}}$. Note that going through a node involves both "entering" and "leaving" that node; thus w does not take V_{k_0} "through" V_{k_0} or V_{k_q} unless one of these is equal to some V_{k_m}, $1 \leqslant m \leqslant q - 1$.

We define α_{ij}^k to be the set of all words over X that take V_i to V_j but not through any V_m, for $m > k$. Thus α_{ij}^0 is a (possibly empty) subset of X if $i \neq j$, and is a subset of $X \cup \{\lambda\}$ if $i = j$. Furthermore, for all i, j, k such that $1 \leqslant i, j, k \leqslant p$,

$$\alpha_{ij}^k = \alpha_{ij}^{k-1} \cup \alpha_{ik}^{k-1}(\alpha_{kk}^{k-1})^* \alpha_{kj}^{k-1}. \tag{*}$$

(See Section 2 of McNaughton–Yamada [8] or Chapter 2, Section 36 of Takahashi [19]).

Let w take V_i to V_j and define $\phi_i(w)$ to be the linear transformation from V_i to V_j induced by $\phi(w)$. Note that if $i = j$, $\phi_i(w)$ is an invertible linear transformation on V_i.

Let w be a word over X of rank $n - h$ and write w as w_1w_2, where w_1 is the shortest prefix of w of rank $n - h$. Then the range of w_1 is some V_i and w_2 takes V_i to some V_j, V_j being the range of $\phi(w)$. Recalling the proof above that there are finitely many ranges of dimension $n - h$, we see that as w varies over all words of rank $n - h$, there are only finitely many $\phi(w_1)$. We note that if w_1' is the shortest prefix of rank $n - h$ of $w' = w_1'w_2'$ which is of rank $n - h$, if $\phi(w_1') = \phi(w_1)$ and if $\phi_i(w_2') = \phi_i(w_2)$ then $\phi(w') = \phi(w)$. Our objective of proving that there are finitely many $\phi(w)$ will be achieved, therefore, once we have succeeded in proving that there are finitely many $\phi_i(w_2)$.

To this end we observe that, for a fixed i, the set of words w_2 which take V_i to some V_j is, $\bigcup_j \alpha_{ij}^p$, so

$$\{\phi_i(w_2): w_2 \text{ takes } V_i \text{ to } V_j \text{, for some } j\} = \phi_i\left(\bigcup_j \alpha_{ij}^p\right).$$

Since ϕ_i preserves unions, our proof will be complete when we show that, for each i and j, $\phi_i(\alpha_{ij}^p)$ is finite. We do this by showing that $\phi_i(\alpha_{ij}^k)$ is finite for all i and j, by induction on k.

Since $\alpha_{ij}^0 \subseteq X \cup \{\lambda\}$, α_{ij}^0 is finite. As an induction hypothesis, assume that $\phi_i(\alpha_{ij}^{k-1})$ is finite, for all i and j. By (*), we obtain

$$\phi_i(\alpha_{ij}^k) = \phi_i(\alpha_{ij}^{k-1}) \cup \phi_i(\alpha_{ik}^{k-1})\, \phi_k((\alpha_{kk}^{k-1})^*)\, \phi_k(\alpha_{kj}^{k-1}).$$

Thus it will suffice to prove that $Q_k = \phi_k((\alpha_{kk}^{k-1})^*)$ is finite. Put $Q = \phi((\alpha_{kk}^{k-1})^*)$.

Both Q and Q_k are semigroups, being closed under the product operation. For $w_1, w_2 \in (\alpha_{kk}^{k-1})^*$, $\phi(w_1) = \phi(w_2)$ implies $\phi_k(w_1) = \phi_k(w_2)$. Thus, for $t \in Q$, $\psi(t) = \phi_k(\phi^{-1}(t))$ is a singleton of Q_k. Furthermore, since ϕ and ϕ_k are homomorphisms, so is ψ.

The semigroup Q_k, being the homomorphic image of the subsemigroup Q of the torsion semigroup T, is torsion. But, since each of its elements is an invertible linear transformation on V_k, by the Lemma Q_k is a group. Also it is generated by $\phi_k(\alpha_{kk}^{k-1})$ which is finite by the induction hypothesis. So, by the Burnside–Schur Theorem, Q_k is finite, concluding the proof of the theorem.

3. Corollaries of the Theorem

To our knowledge, the extendability of the Burnside–Schur Theorem to matrices over a skew field is an open question. The following (which, strictly speaking, is a corollary to the proof rather than to the theorem) shows that if the Burnside–Schur Theorem can be so extended then our theorem can also be extended to matrices over a skew field.

Corollary 1. *Let S be a torsion semigroup of n-by-n matrices over a skew field. If all subgroups of S are locally finite then S is locally finite.*

Proof. Following the proof of our theorem in Section 2: The case $h = 0$ follows from the hypothesis that all subgroups of S are locally finite. The only other modification is in the last part of the proof. Since the homomorphic image of a locally finite group is locally finite and since Q_k is finitely generated, we can prove Q_k finite if we can prove it is the homomorphic image of a subgroup of T. Thus it will suffice to prove the following:

LEMMA. *Let S be a torsion semigroup of n-by-n matrices over a skew field. Consider the matrices in S as linear transformations of an n-dimensional vector space V. Let U be a subspace of V. Let R be a subsemigroup of S all of whose elements M are linear transformations having U as an invariant subspace and such that M restricted to U is invertible. Let g be the homomorphism mapping each M in R to its restriction to U. $H = g(R)$ is a group and there is a subgroup G of R such that $H = g(G)$.*

Proof of Lemma. Since R is nonempty and torsion, at least one of its elements is idempotent and effects the identity transformation on U. Hence H has an identity 1. Choose an idempotent e such that $g(e) = 1$ and such that e has smallest rank among all idempotents which map to 1. Let $G = eRe$. Then $g(G) = g(e)\,g(R)\,g(e) = g(R) = H$. G is a monoid with identity e.

Claim. e is the only idempotent in G. Let f be another idempotent in G; then $g(f)$ is an idempotent in H. But H has only one idempotent, namely 1, so $g(f) = 1$. Now, $f = exe$, so $\text{rank}(f) \leqslant \text{rank}(e)$, and by the minimality of e, $\text{rank}(f) = \text{rank}(e)$; therefore, $\text{range}(f) = \text{range}(e)$. Since an idempotent transformation is the identity on its range, $vef = ve$ for all vectors v in V. So $ef = e$. But since $ef = eexe = exe = f$, $e = f$, proving the claim.

Now, since G is torsion, for each $y \in G$, $y^n = e$, for some n and y^{n-1} is the inverse of y. Thus G is a group, completing the proof of both the Lemma and Corollary 1.

A class of semigroups of particular interest in automata theory is the class of semigroups having only trivial (i.e., one-element) subgroups (see [5], [7], [14]). Corollary 1 clearly implies that our theorem can be generalized to this class of semigroups of matrices over a skew field.

Another class of semigroups of interest in automata theory is the class of locally testable semigroups, which is a subclass of the class of semigroups having only trivial subgroups. Locally testable languages were introduced by McNaughton and Papert [7], and locally testable semigroups by Zalcstein [15] (see also [1], [6]). The definition of "k-testable" that follows is another slight variation of others appearing in the literature, although the resulting concept of local testability coincides with the concept as used in the other references.

Let k be a positive integer. For a word $w = x_1 \cdots x_n$, with $n \geqslant k$, let

$L_k(w)$ and $R_k(w)$ be, respectively, the prefix of w of length k and the suffix of w of length k, and let $S_k(w) = \{x_i \cdots x_{i+k-1}: 1 \leqslant i \leqslant n - k + 1\}$. Let S be a semigroup and let S^+ be the free semigroup generated by the *set* S. S is 1-*testable* if for any two words w_1, w_2 in S^+, $L_1(w_1) = L_1(w_2)$, $R_1(w_1) = R_1(w_2)$ and $S_1(w_1) = S_1(w_2)$ imply that w_1 and w_2 multiply in S to the same element. Let $k > 1$; then S is k-*testable* if for any two words w_1, w_2 of length $\geqslant k$, $L_{k-1}(w_1) = L_{k-1}(w_2)$, $R_{k-1}(w_1) = R_{k-1}(w_2)$ and $S_k(w_1) = S_k(w_2)$ imply that w_1 and w_2 multiply in S to the same element. S is *locally testable* if it is k-testable, for some positive integer k.

Corollary 1 specializes to yield an affirmative answer to the question left open in [16] that was the original motivation for our research.

COROLLARY 2. *Let S be a torsion semigroup of n-by-n matrices over a skew field. Then S is locally testable if and only if, for every idempotent e in S, eSe is idempotent and commutative.*

Proof. The "only if" part is proved in [16, Proposition 2]. To prove the "if" part it suffices, by Theorem 3 of [16], to show that S is locally finite, which follows from Corollary 1.

As is pointed out in [16], local testability can be viewed as a nontrivial generalization of the notion of nilpotence, while the condition of Corollary 2 (for torsion semigroups) is a generalization of the notion of a nil semigroup. Thus, Corollary 2 can be viewed as a generalization of the following well known result (which can also be easily deduced from Corollary 1).

COROLLARY 3. *Any nil semigroup of n-by-n matrices over a skew field is nilpotent of degree at most n.*

Finally, from our Theorem and the proof of Corollary 1 of Procesi [11], we obtain the following generalization of Procesi's Theorem to semigroups:

COROLLARY 4. *Any torsion semigroup that can be embedded in the multiplicative semigroup of an algebra over a field satisfying a polynomial identity is locally finite.*

Proof. Procesi's proof carries over almost verbatim; the only exceptions are that, first, an element x of a torsion semigroup satisfies $x^{m+r} = x^m$ rather than $x^m = 1$, and, second, that our theorem is used in place of the Burnside–Schur Theorem.

Note added in proof. C. Procesi has informed us that Corollary 4 can be extended to PI rings that are not necessarily algebras over a field.

References

1. J. A. Brzozowski and I. Simon, Characterization of locally testable events, *Discrete Math.* **4** (1973), 243–271.
2. I. N. Herstein, "Noncommutative Rings," Carus Monographs No. 15, Mathematical Association of America, Buffalo, New York, 1968.
3. N. Jacobson, "Structure of Rings," American Mathematical Society, Providence, RI, 1968.
4. I. Kaplansky, "Fields and Rings," 2nd ed., The University of Chicago Press, Chicago and London, 1972.
5. K. Krohn and J. Rhodes, Complexity of finite semigroups, *Ann. of Math.* **88** (1968), 128–160.
6. R. McNaughton, Algebraic decision procedures for local testability, *Math. Systems Theory* **8** (1974), 60–76.
7. R. McNaughton and S. Papert, "Counter-Free Automata," M.I.T. Press, Cambridge, MA, 1971.
8. R. McNaughton and H. Yamada, Regular expressions and state graphs for automata, *IRE Trans. Electronic Computers* **9** (1960), 39–47; reprinted in "Sequential Machines–Selected Papers," (E. F. Moore, Ed.), pp. 157–174, Addison-Wesley, Reading, MA, 1964.
9. M. Morse and G. Hedlund, Unending chess, symbolic dynamics and a problem in semigroups, *Duke Math. J.* **11** (1944), 1–15.
10. P. S. Novikov and S. I. Adjan, Infinite periodic groups, I, II, III, *Izv. Akad. Nauk SSSR Ser. Mat.* **32** (1968), 212–244, 251–524, and 709–731; translations *Math. USSR-Izv.* **2** (1968), 209–236, 241–279, and 665–685.
11. C. Procesi, The Burnside Problem, *J. Algebra* **4** (1966), 421–425.
12. J. Rhodes and B. Tilson, Local structure theorems for finite semigroups, *in* "Algebraic Theory of Machines, Languages and Semigroups" (M. Arbib, Ed.), pp. 147–189, Academic Press, New York, 1968.
13. I. Schur, Über Gruppen periodischer Substitutionen, *Sitzungsbericht Preuss. Akad. Wiss.* (1911), 619–627.
14. M. P. Schützenberger, On finite monoids having only trivial subgroups, *Information and Control* **8** (1965), 190–194.
15. Y. Zalcstein, Locally testable languages, *J. Comput. System Sci.* **6** (1972), 151–167.
16. Y. Zalcstein, Locally testable semigroups, *Semigroup Forum* **5** (1973), 216–227.
17. Y. Zalcstein, Finiteness conditions for matrix semigroups, *Proc. Amer. Math. Soc.* **38** (1973), 247–249.
18. Y. Zalcstein, Representations of semigroups by matrices over skew fields, in preparation.
19. H. Takahashi, "Keisankikai," Vol. 2, Iwanami Shoten, Tokyo, 1958 (in Japanese).
20. A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Skr. Vid. Kristiana I. Mat. Naturv. Klasse* **1** (1912), 1–67.