ON THE KROHN-RHODES COMPLEXITY OF SEMIGROUPS OF UPPER TRIANGULAR MATRICES

Mark Kambites
School of Mathematics and Statistics, Carleton University,
Herzberg Laboratories, 1125 Colonel By Drive,
Ottawa, Ontario K1S 5B6, Canada.
mkambite@math.carleton.ca

Abstract. We consider the Krohn-Rhodes complexity of certain semigroups of upper triangular matrices over finite fields. We show that for any $n > 1$ and finite field k, the semigroups of all $n \times n$ upper triangular matrices over k and of all $n \times n$ unitriangular matrices over k have complexity $n - 1$. A consequence is that the complexity $c > 1$ of a finite semigroup places a lower bound of $c + 1$ on the dimension of any faithful triangular representation of that semigroup over a finite field.

1. Introduction

Among the most natural and frequently occurring finite semigroups are those of upper triangular matrices over some finite field. For example, such semigroups appear in the study of algebraic semigroups, where it is known that an algebraic monoid has a faithful triangular representation if and only if its group of units is solvable [7]. As such, triangularizable monoids can be thought of as a natural generalisation of solvable groups. More recently, Almeida, Margolis and Volkov [2] have shown that semigroups of matrices over finite fields generate natural pseudovarieties, and hence have language-theoretic interpretations. More properties of these semigroups can be found in [6].

In their groundbreaking work in the 1960s, Krohn and Rhodes showed that every finite semigroup can be expressed as *divisor* (a homomorphic image of a subsemigroup) of a wreath product of finite groups and finite aperiodic semigroups [3, 5]. The *Krohn-Rhodes complexity* (also called the *group complexity* or just *complexity*) of a finite semigroup S is the least number of group terms in such a wreath product of which S is a divisor, and is a key concept in the theory of finite semigroups. A major open question is that of whether there is an algorithm which, given as input the multiplication table of a finite semigroup S, determines the complexity of S

2000 *Mathematics Subject Classification.* 20M99.
Key words and phrases. finite semigroups, upper triangular matrices, Krohn-Rhodes complexity.

(see, for example, [9]). In the case where each regular $\mathscr{J}$-class of S forms a subsemigroup, however, Rhodes and Tilson [8] demonstrated that there is an algorithm to perform this task.

In this paper, we consider the Krohn-Rhodes complexity of the semigroups of all upper triangular matrices and of all upper unitriangular matrices over a given finite field. In particular, we prove the following.

Theorem 1. *Let S be the semigroup of all $n \times n$ upper unitriangular matrices over a finite field k. Then S has complexity $n-1$.*

Theorem 2. *Let S be the semigroup of all $n \times n$ upper triangular matrices over a finite field k. Then S has complexity 1 if $n = 1$ and k has more than two elements, or complexity $n - 1$ otherwise.*

In addition to this introduction, this paper comprises five sections. In Section 2, we analyse the structure of the semigroups of all upper triangular and all unitriangular matrices of a given size over a given finite field, establishing some facts which will be useful in the sections that follow. In Section 3 we recall a number of constructions and results of Krohn and Rhodes, and analyse how these apply to the semigroups in which we are interested. Sections 4 employs this analysis to prove Theorems 1 and 2. Finally, in Section 5, we describe a consequence of our results for the theory of triangularizable semigroups.

Throughout this paper we assume a familiarity with the standard terminology of semigroup theory, including Green's relations $\mathscr{L}$, $\mathscr{R}$, $\mathscr{D}$ and $\mathscr{J}$ (where, of course, $\mathscr{D}$ and $\mathscr{J}$ coincide in a finite semigroup) and the natural partial order $\leq_{\mathscr{J}}$ on $\mathscr{J}$-classes. The reader not acquainted with these is referred to a text such as [3].

2. Structure of Upper Triangular Matrix Semigroups

Let k be a finite field with identity 1 and zero 0. Let n be a positive integer. If x is an $n \times n$ matrix then for $1 \leq i, j \leq n$ we denote by x_{ij} the entry of x in position (i, j), that is, in the ith row and jth column, of x. Recall that the matrix x is *(upper) triangular* if $x_{ij} = 0$ whenever $1 \leq j < i \leq n$. Following [2], we called an upper triangular matrix *(upper) unitriangular* if, in addition, $x_{ii} = 0$ or $x_{ii} = 1$ for $1 \leq i \leq n$. We call x a *subidentity* if it is unitriangular and $x_{ij} = 0$ whenever $i \neq j$. We denote by $T(n, k)$ and $UT(n, k)$ the semigroups of all $n \times n$ upper triangular matrices and of all $n \times n$ unitriangular matrices respectively, with entries drawn from k, the operation in both cases being usual matrix multiplication.

We define a relation σ on each semigroup $T(n, k)$ by $x \mathrel{\sigma} y$ if and only $x = \lambda y$ for some non-zero field element λ. This relation is easily verified to be a congruence on $T(n, k)$. We shall be interested in the quotient semigroups of the form $T(n, k)/\sigma$ which we call *projective triangular semigroups* and denote by $PT(n, k)$. We denote by $\overline{x}$ the element of $PT(n, k)$ which is the σ-equivalence class of a matrix $x \in T(n, k)$.

By a *row operation* on an upper triangular matrix we shall mean:

(i) adding a multiple of one row to a row above; or
(ii) scaling a row by a non-zero field element; or
(iii) scaling a row by zero.

A row operation is said to be *invertible* if it is of type (i) or (ii), or *unitriangular* if it is of type (i) or (iii). *Column operations* of different types are defined analogously, a type (i) operation being adding a multiple of one column to a row to the right, and so forth.

Let $s, x \in T(n, k)$ [respectively, $s, x \in UT(n, k)$]. It is easy to see that xs can be obtained from s by a certain sequence of [unitriangular] row operations determined by the matrix x. Conversely, every [unitriangular] row operation can be represented as left-multiplication by a certain triangular [unitriangular] matrix. There is an analogous relationship between right-multiplication and [unitriangular] column operations. An immediate consequence of these observations is the following characterization of Green's relations $\mathscr{L}$, $\mathscr{R}$ and $\mathscr{J}$ on the semigroups $T(n, k)$ and $UT(n, k)$.

Proposition 3. *Two matrices in $T(n, k)$ [respectively, $UT(n, k)$] are:*

(i) *$\mathscr{L}$-related exactly if each can be obtained from the other by [unitriangular] row operations;*
(ii) *$\mathscr{R}$-related exactly if each can be obtained from the other by [unitriangular] column operations;*
(iii) *$\mathscr{J}$-related exactly if each can be obtained from the other by [unitriangular] row and column operations;*

We shall use the following straightforward but helpful reformulation of a known result characterising the regular elements of the semigroups $T(n, k)$ and $UT(n, k)$.

Proposition 4. *Let $x \in T(n, k)$ or $x \in UT(n, k)$. Then the following are equivalent:*

(i) *x is regular;*
(ii) *every row in x is a linear combination of rows in x with non-zero diagonal entries;*
(iii) *every column in x is a linear combination of columns in x with non-zero diagonal entries;*
(iv) *x is $\mathscr{J}$-related to a subidentity.*

Proof. We prove the result first for matrices in $T(n, k)$. In [2, Proposition 2.1], the regular elements of $T(n, k)$ are characterized as those matrices whose rank is the equal to the number of their non-zero diagonal entries. The equivalence of (i), (ii) and (iii) in the upper triangular case is a consequence of this, together with the observation that sets of rows (or columns) with non-zero diagonal entries in an upper triangular matrix are necessarily linearly independent. Now if $x \in UT(n, k)$ and $y \in T(n, k)$ are such that

$xyx = x$ and $yxy = y$ then we must have $y \in UT(n,k)$. Thus, a unitriangular matrix is regular in $UT(n,k)$ exactly if it is regular in $T(n,k)$, and (i), (ii) and (iii) are also equivalent in the unitriangular case.

Since subidentities are idempotent and hence regular, and regularity is a property of $\mathscr{J}$-classes, it is immediate that (iv) implies (i). Finally, if x satisfies (ii) and (iii), it is easy to see that one can use row operations to make zero all entries in rows with diagonal zeros, column operations to make zero all entries in columns with diagonal zeros, row operations to make all the non-zero diagonal entries 1, and then column operations to remove the remaining non-diagonal entries, so that, by Proposition 3, x is $\mathscr{J}$-related to a subidentity in $T(n,k)$. Moreover, if x is unitriangular, then this can be accomplished with unitriangular operations so that x is $\mathscr{J}$-related to a subidentity in $UT(n,k)$. Hence, (iv) holds. □

We now introduce some terminology and associated lemmas which will be helpful in analysing the structure of certain quotients of upper triangular matrix semigroups.

Let $x \in UT(n,k)$ or $x \in T(n,k)$. The *diagonal shape* of x is the set

$$\mathrm{Shape}(x) = \{i \in \mathbb{Z} \mid 1 \leq i \leq n, s_{ii} \neq 0\}.$$

Thus, two matrices have the same diagonal shape if they have zeros in exactly the same positions on the main diagonal. Note that diagonal shape is preserved by invertible row and column operations, and that for upper triangular matrices s and t we always have $\mathrm{Shape}(st) = \mathrm{Shape}(s) \cap \mathrm{Shape}(t)$.

Corollary 5. *Regular elements of $T(n,k)$ or $UT(n,k)$ are $\mathscr{J}$-related if and only if they have the same diagonal shape.*

Proof. It follows easily from Proposition 3 together with our observations above concerning diagonal shape that $\mathscr{J}$-related elements must have the same diagonal shape. Conversely, if x and y are regular and have the same diagonal shape, then by Proposition 4 they must be $\mathscr{J}$-related to the unique subidentity with that diagonal shape, and hence to each other. □

Proposition 6. *Let $x, s \in T(n,k)$ or $x, s \in UT(n,k)$ with x regular. Then the following are equivalent:*

(i) $\mathrm{Shape}(x) \subseteq \mathrm{Shape}(s)$;
(ii) $x \leq_{\mathscr{J}} s$
(iii) $xs \mathrel{\mathscr{J}} x$;
(iv) $xs \mathrel{\mathscr{R}} x$;
(v) $sx \mathrel{\mathscr{J}} x$;
(vi) $sx \mathrel{\mathscr{L}} x$;
(vii) $xsx \mathrel{\mathscr{J}} x$.

In particular, each regular $\mathscr{J}$-class of $T(n,k)$ or $UT(n,k)$ is a subsemigroup.

Proof. We show first that (i) implies (iv). If (i) holds, then by our observations above, xs can be obtained from x by applying [unitriangular] column operations to x. By Lemma 3, every column of x is a linear combination of columns with non-zero diagonal entries. But since $\mathrm{Shape}(x) \subseteq \mathrm{Shape}(s)$, it follows that every column of x will be a combination of columns of xs with non-zero diagonal entries. Hence, $xst = x$ for some t, which suffices to show that $xs \mathscr{R} x$ and (iv) holds.

Now if x and s are regular and $\mathscr{J}$-related, then by Corollary 5, it follows that $\mathrm{Shape}(x) = \mathrm{Shape}(s)$, so that $xs \mathscr{J} x$. Hence, the regular $\mathscr{J}$ classes of $T(n,k)$ and $UT(n,k)$ form subsemigroups. The equivalence of (ii), (iii), (iv), (v), (vi) and (vii) is well-known to hold in any finite semigroup whose regular $\mathscr{J}$-classes are subsemigroups (see, for example, [3] and [1]), so it will suffice to show that (iii) implies (i). But if (iii) holds then by Proposition 3, $\mathrm{Shape}(x) = \mathrm{Shape}(xs) = \mathrm{Shape}(x) \cap \mathrm{Shape}(s)$ so we must have $\mathrm{Shape}(x) \subseteq \mathrm{Shape}(s)$, and (i) holds. □

By an n-*interval*, we mean a subset

$$\alpha \subseteq \{1, \ldots, n\}$$

with the property that for any $i \leq j \leq k \in \{1, \ldots, n\}$ with $i, k \in \alpha$ we have also $j \in \alpha$.

Now let x be an $n \times n$ matrix over a field k and α an n-interval. The *restriction* $x|_\alpha$ of x to α is the matrix $(x_{ij})_{i,j\in\alpha}$ with rows and columns indexed by α. We say that matrices x and y *agree on* α if $x_{ij} = y_{ij}$ for all $i, j \in \alpha$. We say that x and y *are scalar multiples on* α if there exists a non-zero field element $\lambda \in k$ such that $x_{ij} = \lambda y_{ij}$ for all $i, j \in \alpha$, that is, such that x and λy agree on α. Notice that, if $x, y \in UT(n,k)$ are such that $x_{ii} \neq 0$ for some $i \in \alpha$ then x and y are scalar multiples on α exactly if they agree on α.

Row and column operations can be applied to submatrices in a natural way. If α is an n-interval, we say that two $n \times n$ matrices x and y are:

(i) *[unitriangular] row-related within* α if each of $x|_\alpha$ and $y|_\alpha$ can be obtained from the other by applying successive [unitriangular] row operations;
(ii) *[unitriangular] column-related within* α if each of $x|_\alpha$ and $y|_\alpha$ can be obtained from the other by applying successive [unitriangular] column operations; and
(iii) *[unitriangular] related within* α if each of $x|_\alpha$ and $y|_\alpha$ can be obtained form the other by applying successive [unitriangular] row and/or column operations.

Notice that two matrices which are scalar multiples on α are both row-related and column-related within α.

For any n-interval α, any $x, y \in T(n,k)$ and any $i, j \in \alpha$ we have

$$(xy)_{ij} = \sum_{i \leq p \leq j} x_{ip} y_{pj}$$

where the fact that α is an n-interval ensures that each $p \in \alpha$. Thus, we have

$$(x|_\alpha)(y|_\alpha) = (xy)|_\alpha,$$

that is, restriction to a given interval is a semigroup homomorphism. The following simple consequences of this observation are now clear.

Lemma 7. *Suppose $s, t \in T(n,k)$ agree [respectively, are scalar multiples on] an n-interval α. Then for any $x \in T(n,k)$, we have that xs and xt agree [are scalar multiples] on α and sx and tx agree [are scalar multiples] on α.*

Lemma 8. *Let α be an n-interval. If $p, q \in T(n,k)$ are related by a sequence of row operations [column operations, row and column operations] then $p|_\alpha$ and $q|_\alpha$ are related by a sequence of row operations [column operations, row and column operations]. Likewise for unitriangular operations.*

Next, we prove a technical lemma which characterizes Green's relations $\mathscr{L}$, $\mathscr{R}$ and $\mathscr{J}$ in certain quotients associated with intervals.

Lemma 9. *Let n be an integer, and*

$$f : \mathscr{P}(\{1, \dots, n\}) \to \mathscr{P}(\{1, \dots, n\})$$

a function which takes shapes to n-intervals. Define a relation ρ on $T(n,k)$ [or on $UT(n,k)$] by $x \ \rho \ y$ if and only if x and y have the same shape X and agree on the interval $f(X)$. Suppose the relation ρ is a congruence on $T(n,k)$ [or on $UT(n,k)$]. Let $[x]$ denote the ρ-equivalence class of a matrix x. Then for any s, t we have

- (i) *$[s] \ \mathscr{L} \ [t]$ if and only if s and t have the same diagonal shape and are [unitriangular] row-related within $f(\mathrm{Shape}(s))$;*
- (ii) *$[s] \ \mathscr{R} \ [t]$ if and only if s and t have the same diagonal shape and are [unitriangular] column-related within $f(\mathrm{Shape}(t))$; and*
- (iii) *$[s] \ \mathscr{J} \ [t]$ if and only if s and t have the same diagonal shape and are [unitriangular] related within $f(\mathrm{Shape}(s))$.*

Proof. We prove the result for $\mathscr{L}$. The result for $\mathscr{R}$ is dual, and that for $\mathscr{J}$ is a consequence of the other two.

Suppose first that s and t have the same diagonal shape and are [unitriangular] row-related within $f(\mathrm{Shape}(s)) = f(\mathrm{Shape}(t))$. Let p and q be the matrices which agree with s and t respectively on $f(\mathrm{Shape}(s))$ and are zero everywhere else. Then p and q are easily seen to be related by [unitriangular] row operations, so by Proposition 3, we have $p \ \mathscr{L} \ q$. It follows that $[s] = [p] \ \mathscr{L} \ [q] = [t]$, as required.

Conversely, suppose that $[s] \ \mathscr{L} \ [t]$. Then by [3, Chapter 7, Fact 2.1(d)], there exist $p, q \in T(n,k)$ [$p, q \in UT(n,k)$] such that $[p] = [s]$, $[q] = [t]$ and $p \ \mathscr{L} \ q$. Now p and q are related by a sequence of [unitriangular] row operations, and it follows by Lemma 8 that p and q are [unitriangular] row-related within $f(\mathrm{Shape}(s))$. But p and q must agree with s and t respectively within $f(\mathrm{Shape}(s))$, so it follows that s and t are row-related within $f(\mathrm{Shape}(s))$. □

Lemma 10. *Let ρ be a congruence as in the statement of Lemma 9. Let $x, s \in T(n,k)$ or $x, s \in UT(n,k)$ with $[x]$ regular. Then the following are equivalent:*

(i) $\mathrm{Shape}(x) \subseteq \mathrm{Shape}(s)$;
(ii) $[x] \leq_{\mathscr{J}} [s]$;
(iii) $[xs] \mathrel{\mathscr{J}} [x]$;
(iv) $[xs] \mathrel{\mathscr{R}} [x]$;
(v) $[sx] \mathrel{\mathscr{J}} [x]$;
(vi) $[sx] \mathrel{\mathscr{L}} [x]$;
(vii) $[xsx] \mathrel{\mathscr{J}} [x]$.

Proof. We have already observed that the equivalence of (ii), (iii), (iv), (v), (vi) and (vii) holds in any finite semigroup whose regular $\mathscr{J}$-classes are subsemigroups, so it will suffice to show that (i) implies (iii), and that (iii) implies (i). Suppose (i) holds. Then by [3, Chapter 7, Fact 2.1(e)] we may choose $y \in T(n,k)$ (or $y \in UT(n,k)$) such that y is regular and $[x] = [y]$. In particular,

$$\mathrm{Shape}(y) = \mathrm{Shape}(x) \subseteq \mathrm{Shape}(s),$$

so by Proposition 6 we have $ys \mathrel{\mathscr{J}} y$, from which it follows that

$$[xs] = [x][s] = [y][s] = [ys] \mathrel{\mathscr{J}} [y] = [x],$$

so that (iii) holds.

Conversely, if (iii) holds then using Lemma 9 we have

$$\mathrm{Shape}(x) \cap \mathrm{Shape}(s) = \mathrm{Shape}(xs) = \mathrm{Shape}(x),$$

from which it follows that $\mathrm{Shape}(x) \subseteq \mathrm{Shape}(s)$ and (i) holds. □

3. Calculating Complexity

Let S be a finite semigroup. We define a relation $\equiv$ on S by $s \equiv t$ if and only if for all regular elements x and y in the same $\mathscr{J}$-class, we have $xsy \mathrel{\mathscr{J}} x \iff xty \mathrel{\mathscr{J}} x$, and if $xsy, xty \mathrel{\mathscr{J}} x$ then $xsy = xty$. Throughout this paper, we will denote the equivalence class of an element x under this relation by $[x]$.

We define another relation $\sim$ on S by $s \sim t$ if and only if for every regular element x we have $xs \mathrel{\mathscr{R}} x \iff xt \mathrel{\mathscr{R}} t$, and if $xs, xt \mathrel{\mathscr{R}} t$ then $xs \mathrel{\mathscr{L}} xt$. Throughout this paper, we will denote the equivalence class of an element x under this relation by $\langle x \rangle$.

The relations $\equiv$ and $\sim$ are easily verified to be congruences on any semigroups whose regular $\mathscr{J}$-classes are subsemigroups. The following restatement of a result of Rhodes and Tilson [8] provides the connection with complexity.

Theorem 11 ([8])**.** *Let S be a finite semigroup in which each regular $\mathscr{J}$-class is a subsemigroup. Then $S/{\equiv}$ has the same complexity as S, and $(S/{\equiv})/{\sim}$ has complexity one less than that of S (or 0 if S has complexity 0).*

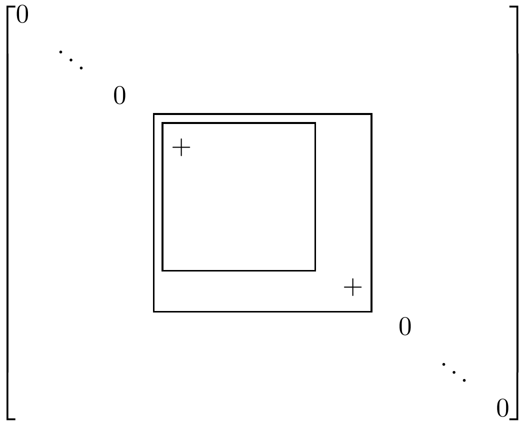

FIGURE 1. The first and second inner squares of an upper triangular matrix. The symbol $+$ denotes a non-zero entry, while omitted entries may take any value.

Since one can test a semigroup for aperiodicity, since the congruences $\equiv$ and $\sim$ on a given finite semigroup are computable, and since the property of the regular $\mathscr{J}$-classes being subsemigroups is preserved under homomorphism, it follows that one can calculate the complexity of a semigroup S satisfying this property by repeatedly calculating $(S/\equiv)/\sim$ until one obtains an aperiodic semigroup.

In the rest of this section, we consider these relations as applied to semigroups of upper triangular matrices. We define

$$T^1(n,k) \;=\; T(n,k)/\equiv \;=\; \{[x] \mid x \in T(n,k)\};\ \text{and}$$

$$T^2(n,k) \;=\; T^1(n,k)/\sim \;=\; \{\langle [x] \rangle \mid x \in T(n,k)\};$$

and similarly

$$UT^1(n,k) \;=\; UT(n,k)/\equiv \;=\; \{[x] \mid x \in UT(n,k)\};\ \text{and}$$

$$UT^2(n,k) \;=\; UT^1(n,k)/\sim \;=\; \{\langle [x] \rangle \mid x \in UT(n,k)\}.$$

Our objective is to obtain concrete characterizations of these semigroups. Give a matrix $s \in T(n,k)$ or $s \in UT(n,k)$, we associate to it two particular n-intervals, called respectively the *first* and *second* inner squares of s:

$$\text{Square}^1(s) = \{i \in \mathbb{Z} \mid p \leq i \leq q \text{ for some } p, q \in \text{Shape}(s)\}$$

$$\text{Square}^2(s) = \{i \in \mathbb{Z} \mid p \leq i < q \text{ for some } p, q \in \text{Shape}(s)\}.$$

See Figure 1 for an illustration. Notice that these squares are a function only of the diagonal shape of a matrix, so that two matrices with the same diagonal shape have the same first and second inner squares.

We now consider the relation $\equiv$ on the semigroups $T(n,k)$ and $UT(n,k)$.

Proposition 12. *For any $s, t \in T(n,k)$ or $s, t \in UT(n,k)$ we have $s \equiv t$ if and only if s and t have the same diagonal shape and agree on their first inner square.*

Proof. Suppose first that $s \equiv t$. Suppose for a contradiction that s and t have different diagonal shapes, say s has a diagonal zero where t does not. Let x be the subidentity with the same diagonal shape as t. Now by Proposition 6, we have $xtx \mathscr{J} x$ but xsx is not $\mathscr{J}$-related to x. Hence, s and t must have the same diagonal shape.

Now suppose that s and t do not agree on their first inner square. Then certainly the first inner square is non-empty, so the shape is non-empty. Let $p, q \in \mathrm{Shape}(s)$ be respectively the least and greatest elements of $\mathrm{Shape}(s)$. Let $i \in \mathrm{Square}^1(s)$ be minimal such that $s_{ij} \neq t_{ij}$ for some $j \in \mathrm{Square}^1(s)$. Let $j \in \mathrm{Square}^1(s)$ be maximal such that $s_{ij} \neq t_{ij}$. Notice that, from the definition of the first inner square, we have $p \leq i \leq j \leq q$

Now let x be the matrix with 1s in position (p,p), (p,i) and (q,q) and 0s elsewhere. Let y be the matrix with 1s in position (p,p), (j,q) and (q,q) and 0s elsewhere. By Proposition 4, x and y are both regular. By Propositions 4 and 6, x, y, xsy and xty are all $\mathscr{J}$-related. Now if $p \neq i$ and $j \neq q$, then a simple calculation shows that

$$(xsy)_{pq} = s_{pj} + s_{pq} + s_{ij} + s_{iq}$$

and

$$(xty)_{pq} = t_{pj} + t_{pq} + t_{ij} + t_{iq}.$$

Now by the minimality assumption on i and maximality assumption on j, we have $s_{pj} = t_{pj}$, $s_{pq} = t_{pq}$ and $s_{iq} = t_{iq}$, so that $(xsy)_{pq} \neq (xty)_{pq}$, and hence $xsy \neq xty$. Similar arguments apply in the cases where $p = i$ and/or $j = q$, so that we always have $xsy \neq xty$, as required.

Conversely, suppose s and t have the same diagonal shape and agree on their first inner square, and let x and y be regular with $x \mathscr{J} y$. Then by Proposition 4, x and y are both $\mathscr{J}$-related to the same subidentity e, so $x = aeb$, $y = ced$ for some a, b, c, d. Now if

$$\mathrm{Shape}(x) = \mathrm{Shape}(y) \nsubseteq \mathrm{Shape}(s) = \mathrm{Shape}(t)$$

then by Proposition 6, both xsy and xty will not be $\mathscr{J}$-related to x and y.

Otherwise, the shape of x and y is contained within that of s and t, so the first inner square of x and y is contained within that of s and t. Now e has 0s everywhere outside the first inner square of x and y, and hence everywhere outside the first inner square of s and t. By Lemma 7, bsc and btc agree on the first inner square of s and t. It follows that $ebsce = ebtce$ so that $xsy = aebxced = aebtced = xty$ as required. ∎

A consequence is that the relation $\equiv$ on $T(n,k)$ and $UT(n,k)$ satisfies the conditions of Lemma 9 and Lemma 10. This gives us a characterization of Green's relations $\mathscr{L}$, $\mathscr{R}$ and $\mathscr{J}$ in the semigroups $T^1(n,k)$ and $UT^1(n,k)$.

We now consider the congruence $\sim$ on $UT^1(n,k)$ and $T^1(n,k)$.

Proposition 13. *For any $s, t \in UT(n,k)$ or $s, t \in T(n,k)$, we have $[s] \sim [t]$ if and only if s and t have the same diagonal shape and are scalar multiples on their second inner square.*

Proof. Suppose first that $[s] \sim [t]$. Suppose for a contradiction that s and t have different diagonal shape, say s has a diagonal zero where t does not. Let x be the subidentity with the same shape as t. Now using Lemma 10 it is easy to see that $[xt] \mathscr{R} [x]$ but $[xs]$ and $[x]$ are not $\mathscr{R}$-related. Hence, s and t must have the same diagonal shape.

Now suppose for another contradiction that s and t are not scalar multiples on their second inner square. Then the second inner square of s and t is non-empty, from which it follows that the shape of s and t has more than one element. Let $p, q \in \mathrm{Shape}(s)$ be respectively the least and greatest elements of $\mathrm{Shape}(s)$, noting that $p \neq q$.

We define a matrix y differently in two cases.

(i) We consider first the case in which row p of s is not a scalar multiple of row p of t on the second inner square. In this case, we define y to be the matrix with 1s in positions (p,p) and (q,q) and 0s elsewhere.

(ii) Suppose now that row p of s is a scalar multiple of row p of t, say row p of s is λ times row p of t, on the second inner square. Then we can choose a row i with $p < i < q$ such that row i of s is not λ times row i of t on the second inner square. In this cases, we let y be the matrix with 1s in positions (p,p), (q,q) and (p,i) and 0s elsewhere.

Now by Proposition 4, y is regular and so, since images of regular elements are regular, $[y]$ is regular. Now y, ys and yt have non-zero entries only in rows p and q, and also have the same first inner square as s and t. Also, $\mathrm{Shape}(y) \subseteq \mathrm{Shape}(s)$, so by Lemma 10, we have $[y][s] \mathscr{R} [y] \mathscr{R} [y][t]$.

We seek now to show that $[ys]$ is not $\mathscr{L}$-related to $[yt]$. To understand the following argument, the reader may find it helpful to refer to Figure 2. In fact, row p of ys is the sum of rows p and i in s (or just row p in case (i)), while row p of yt is the sum of rows p and i in t (or just row p in case (i)). It is readily seen, in both cases, that row p of ys is not a multiple of row p of yt. Moreover, the only changes which can be effected to row p of ys within the second inner square by row operations are multiplications by scalars. It follows that they cannot be made into multiples of each other by row operations. Thus, one cannot get from the first inner square of ys to that of yt and back again by row operations, so by Lemma 9 $[y][s] = [ys]$ is not $\mathscr{L}$-related to $[yt] = [y][t]$. This contradicts the assumption that $[s] \sim [t]$.

Conversely, suppose that s and t have the same diagonal shape and are scalar multiples on their second inner square. Let $[x]$ be regular. Then by Lemma 10,

$$[xs] \mathscr{R} [x] \iff \mathrm{Shape}(x) \subseteq \mathrm{Shape}(s) = \mathrm{Shape}(t) \iff [xt] \mathscr{R} [x].$$

$$\begin{bmatrix}
0 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 0 & & & & & & & & \\
 & & & s_{pp} & (ys)_{p,p+1} & \cdots & (ys)_{p,q-1} & (ys)_{p,q} & (ys)_{p,q+1} & \cdots & (ys)_{p,n} \\
 & & & & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
 & & & & & \ddots & \vdots & \vdots & \vdots & & \vdots \\
 & & & & & & 0 & 0 & 0 & \cdots & 0 \\
 & & & & & & & s_{qq} & s_{q,q+1} & \cdots & s_{q,n} \\
 & & & & & & & & 0 & \cdots & 0 \\
 & & & & & & & & & \ddots & \vdots \\
 & & & & & & & & & & 0
\end{bmatrix}$$

FIGURE 2. The form of the matrix ys showing the first and second inner squares of s.

Now assume further that $\text{Shape}(x) \subseteq \text{Shape}(s)$. By Lemma 7, xs and xt are scalar multiples on the second inner square of s and t. We consider first the case in which x has a zero in the bottom right position of the first inner square of s. In this case, the first inner squares of xs and xt are contained within the second inner square of s and t, so xs and xt are scalar multiples on their first inner square. In the unitriangular case, since the first inner square is either empty or contains a non-zero diagonal entry, this means that xs and xt agree on their first inner square, so that $[xs] = [xt]$. In the general triangular case, scalar multiplication can be done by row operations, so it follows that $[xs] \mathscr{L} [xt]$.

Now consider the case in which x does not have a zero in the bottom right position of the first inner square of s. Otherwise, there exists a non-zero field element λ such that the first inner squares of λxs and xt differ only in the rightmost column, and the bottom row has a non-zero entry in this column, so one can get from the first inner square of λxs that of xt by adding multiples of the bottom row to other rows. It follows by Lemma 9 that $[x][s] = [xs] \mathscr{L} [xt] = [x][t]$, and so $[s] \sim [t]$ as required. $\square$

4. PROOFS OF THE MAIN THEOREMS

In this section, we apply the results of the previous section to the proof of Theorems 1 and 2.

Proposition 14. *For any finite field k and $n \geq 2$, the semigroup $UT^2(n,k)$ has the same complexity as $UT(n-1,k)$.*

Proof. We shall show that $UT^2(n,k)$ has a subsemigroup with the same complexity as $UT(n-1,k)$, and divides the direct product of $UT(n-1,k)$ with the 2-element semilattice.

$$\left(\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,n-1} \\ & x_{22} & \cdots & x_{2,n-1} \\ & & \ddots & \vdots \\ & & & x_{n-1,n-1} \end{bmatrix}, \epsilon \right) \mapsto \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,n-1} & 0 \\ & x_{22} & \cdots & x_{2,n-1} & 0 \\ & & \ddots & \vdots & \vdots \\ & & & x_{n-1,n-1} & 0 \\ & & & & \epsilon \end{bmatrix}$$

FIGURE 3. The action of the function $g : UT(n-1,k) \times L \to UT(n,k)$.

Let $L = \{0, 1\}$ be the two element semilattice, viewed as a subset of the field k. We define a function $g : UT(n-1,k) \times L \to UT(n,k)$ by

$$g(x,\epsilon)_{ij} = \begin{cases} x_{ij} & \text{if } 1 \le i,j < n \\ \epsilon & \text{if } i = j = n \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 3 for an illustration. We claim that this induces a surjective homomorphism $g' : UT(n-1,k) \times L \to UT^2(n,k)$ given by $g'(x,\epsilon) = \langle [g(x,\epsilon)] \rangle$.

The function g is a standard faithful representation of the direct product of matrix groups, and in particular is a homomorphism. It follows immediately that g' is a homomorphism.

Next, we show that g' is surjective. Given $\langle [x] \rangle \in UT^2(n,k)$, we define $y \in UT(n-1,k)$ by $y_{ij} = x_{ij}$ for $1 \le i,j < n$. Now it is easily verified that $g(y, x_{nn})$ has the same diagonal shape as x and agrees with x on the second inner square, so that by Proposition 13, $g'(y, x_{nn}) = \langle [x] \rangle$ as required.

We have shown that $UT^2(n,k)$ is a homomorphic image of the direct product of $UT(n-1,k)$ with an aperiodic semigroup. It follows that $UT^2(n,k)$ has complexity no greater than that of $UT(n-1,k)$.

Next, we claim that the restriction of the map g' to the set $UT(n-1,k) \times \{1\}$ is *aperiodic*, that is, injective on every subgroup.

To show this, let $(x,1)$ and $(y,1)$ be distinct elements of the same subgroup of $UT(n-1,k) \times \{1\}$. Then x and y are distinct elements of the same subgroup of $UT(n-1,k)$. Now since the congruence $\equiv$ is aperiodic [8], it follows that $[x] \neq [y]$, so by Proposition 12, x and y do not agree on the first inner square. But the first inner square of x and y must be contained within the second inner square of $g(x,1)$ and $g(y,1)$, so that $g(x,1)$ and $g(y,1)$ differ in the second inner square. Hence, by Proposition 13, $g'(x,1)$ is distinct from $g'(y,1)$.

Since aperiodic maps preserve complexity (see, for example, [4]), it follows that $UT^2(n,k)$ has a subsemigroup of the same complexity as $UT(n-1,k)$, and hence has complexity no less than that of $UT(n-1,k)$. This completes the proof. □

We are now ready to prove Theorem 1.

Theorem 1. *The semigroup $UT(n,k)$ of all $n \times n$ upper unitriangular matrices over a finite field k has complexity $n-1$.*

Proof. First, we observe that $UT(1,k)$ is always the two-element semilattice, and so has complexity 0. On the other hand, $UT(2,k)$ is easily seen to have group of units isomorphic to the (necessarily non-trivial) additive group of k, and hence has complexity at least 1.

To complete the proof, we use induction on the size n of the matrices. Assume that $n \geq 2$ and that $UT(m,k)$ has complexity $m-1$ for all $m < n$. Then by Proposition 14, $UT^2(n,k)$ has complexity $n-2$. Also, $UT(n,k)$ contains a subsemigroup isomorphic to $UT(2,k)$, and hence has strictly positive complexity. Now by Theorem 11, $UT(n,k)$ has complexity $n-1$ as required. □

Our next objective is to determine the Krohn-Rhodes complexity of the semigroup of all upper triangular square matrices of a given size over a finite field, by proving Theorem 2. We begin by proving a weak analogue of Proposition 14 which holds in the more general upper triangular case.

Proposition 15. *For any finite field k and $n \geq 2$, the complexity of the semigroup $T^2(n,k)$ does not exceed that of the projective triangular matrix semigroup $PT(n-1,k)$.*

Proof. We shall show that $T^2(n,k)$ divides the direct product of $PT(n-1,k)$ with the 2-element semilattice.

As in the proof of Proposition 14, let $L = \{0,1\}$ be the two element semilattice, viewed as a subset of the field k. We define a function $g : T(n-1,k) \times L \to T(n,k)$ much as before, by

$$g(x,\epsilon)_{ij} = \begin{cases} x_{ij} & \text{if } 1 \leq i,j < n \\ \epsilon & \text{if } i = j = n \\ 0 & \text{otherwise.} \end{cases} .$$

We claim that now this induces a surjective homomorphism $g' : PT(n-1,k) \times L \to T^2(n,k)$ given by $g'(\overline{x},\epsilon) = \langle [g(x,\epsilon)] \rangle$.

First, we show that g' is well-defined. Suppose $x, y \in T(n-1,k)$ are such $\overline{x} = \overline{y}$ in $PT(n-1,k)$. Then $x = \lambda y$ for some non-zero field element λ. Let $\epsilon \in L$. Then for any $i,j < n$ we have

$$g(x,\epsilon)_{ij} = x_{ij} = \lambda y_{ij} = \lambda g(y,\epsilon)_{ij}$$

and furthermore

$$g(x,\epsilon)_{nn} = \epsilon = g(y,\epsilon)_{nn}.$$

Clearly, then, $g(x,\epsilon)$ and $g(y,\epsilon)$ have the same diagonal shape. Moreover, since their second inner square cannot contain n, they are scalar multiples on the second inner square. Thus, by Proposition 13, we have $\langle [g(x,\epsilon)] \rangle = \langle [g(y,\epsilon)] \rangle$ as required.

It follows by the same argument as in Proposition 14 that g is a homomorphism, and again by general principles that g' is a homomorphism. The same proof used in Proposition 15 suffices to show that g' is surjective. Hence, $T^2(n,k)$ is a homomorphic image of the direct product of $PT(n-1,k)$ with an aperiodic semigroup. It follows that $T^2(n,k)$ has complexity no greater than that of $PT(n-1,k)$, as required. □

We are now ready to prove Theorem 2.

Theorem 2. *The semigroup $T(n,k)$ of all $n \times n$ upper triangular matrices over a finite field k has complexity 1 if $n = 1$ and k has more than two elements, or complexity $n-1$ otherwise.*

Proof. Clearly, $T(1,k)$ is isomorphic to the multiplicative group of the field k with an adjoined zero. Hence, $T(1,k)$ has complexity 0 if k has two elements (and hence trivial multiplicative group), and 1 otherwise.

Next, we consider the complexity of $T(2,k)$. Certainly $T(2,k)$ contains $UT(2,k)$ as a subsemigroup, so by Theorem 1, $T(2,k)$ has complexity at least 1. Hence, by Theorem 11, $T(2,k)$ has complexity exactly 1 more than that of $T^2(2,k)$. It follows easily from Proposition 13 that for any matrix $x \in T^2(2,k)$ we have $\langle [x] \rangle = \langle [e] \rangle$ where e is the unique subidentity with the same diagonal shape as x. Thus, every element of $T^2(2,k)$ is idempotent, and so $T^2(2,k)$ has no non-trivial subgroups. We conclude that $T^2(2,k)$ has complexity 0, and so $T(2,k)$ has complexity 1.

Now let $n \geq 3$ and assume for induction that $T(m,k)$ has complexity $m-1$ for $2 \leq m < n$. By Proposition 15, the complexity of $T^2(n,k)$ is no more than that of $PT(n-1,k)$. But the complexity of $T^2(n,k)$ is one less than that of $T(n,k)$, so it follows that the complexity of $T(n,k)$ exceeds that of $PT(n-1,k)$ by at most one. Now $PT(n-1,k)$ is a quotient of $T(n-1,k)$, which by the inductive hypothesis has complexity $n-2$. Thus, $PT(n-1,k)$ has complexity at most $n-2$, while $T(n,k)$ has complexity at most $n-1$. But $T(n,k)$ contains $UT(n,k)$ as a subsemigroup, so by Theorem 1, $T(n,k)$ must have complexity at least $n-1$. Hence, $T(n,k)$ has complexity $n-1$, as required. □

Theorem 2 and Proposition 15 together describe the complexity of the projective triangular semigroups.

Corollary 16. *For any positive integer n and finite field k, the projective triangular semigroup $PT(n,k)$ has complexity $n-1$.*

5. Complexity of Triangularizable Semigroups

Recall that a semigroup S is *triangularizable* over a field k if S embeds in $T(n,k)$ for some n. An immediate consequence of Theorem 2 is that the complexity of a finite semigroup S places a lower bound on the dimension of any faithful triangular representation of S over a finite field.

Corollary 17. *Let S be a finite semigroup of Krohn-Rhodes complexity $c > 1$. Then S does not embed into $T(n,k)$ for any finite field k and $n < c+1$.*

It seems natural to ask whether there is a converse to this result, that is, whether the minimum dimension of a faithful triangular representation of a triangularizable semigroup is always exactly one more than the Krohn-Rhodes complexity. In fact, the answer to this question is negative. To see this, let k be any finite field and consider the four subidentity matrices in the semigroup $T(2, k)$. These form a semilattice which necessarily has complexity 0. However, they cannot embed into $T(1, k)$ (or indeed any $T(1, f)$), which we saw in the proof Theorem 2 is a zero-group, and hence has only two idempotents.

However, one might still ask the following rather general question.

Question 18. *Let S be a subsemigroup of $T(n, k)$ for some integer n and finite field k. Is there a straightforward characterization of the complexity of S?*

Acknowledgements

This research was funded by the Leverhulme Trust. The author would like to thank Benjamin Steinberg for many helpful discussions, and Kirsty for all her support and encouragement.

References

[1] J. Almeida. *Finite semigroups and universal algebra*, volume 3 of *Series in Algebra*. World Scientific, 1994.

[2] J. Almeida, S. W. Margolis, and M. V. Volkov. The pseudovariety of semigroups of triangular matrices over a finite field. preprint, 2004.

[3] M. A. Arbib, K. Krohn, and J. L. Rhodes. *Algebraic theory of machines, languages, and semigroups*. Academic Press, New York, 1968.

[4] S. Eilenberg. *Automata, languages, and machines. Vol. B*. Academic Press, New York, 1976. With two chapters ("Depth decomposition theorem" and "Complexity of semigroups and morphisms") by B. R. Tilson, Pure and Applied Mathematics, Vol. 59.

[5] K. Krohn and J. L. Rhodes. Complexity of finite semigroups. *Ann. of Math. (2)*, 88:128–160, 1968.

[6] J. Okniński. *Semigroups of matrices*, volume 6 of *Series in Algebra*. World Scientific Publishing Co. Inc., River Edge, NJ, 1998.

[7] M. S. Putcha. *Linear algebraic monoids*, volume 133 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1988.

[8] J. Rhodes. Algebraic theory of finite semigroups. Structure numbers and structure theorems for finite semigroups (with an appendix by B. R. Tilson). In *Semigroups (Proc. Sympos., Wayne State Univ., Detroit, Mich., 1968)*, pages 125–208. Academic Press, New York, 1969.

[9] P. Weil. Profinite methods in semigroup theory. *Internat. J. Algebra Comput.*, 12(1-2):137–178, 2002. International Conference on Geometric and Combinatorial Methods in Group Theory and Semigroup Theory (Lincoln, NE, 2000).