Semigroup Forum (2019) 99:655–678
https://doi.org/10.1007/s00233-018-09992-7

RESEARCH ARTICLE

Random ubiquitous transformation semigroups

Julius Jonušas[1] · **Sascha Troscheit**[2]

Received: 7 February 2018 / Accepted: 17 December 2018 / Published online: 22 January 2019

Abstract
A *smallest generating set* of a semigroup is a generating set of the smallest cardinality. Similarly, an *irredundant generating set* X is a generating set such that no proper subset of X is also a generating set. A semigroup S is *ubiquitous* if every irredundant generating set of S is of the same cardinality. We are motivated by a naïve algorithm to find a small generating set for a semigroup, which in practice often outputs a smallest generating set. We give a sufficient condition for a transformation semigroup to be ubiquitous and show that a transformation semigroup generated by k randomly chosen transformations asymptotically satisfies the sufficient condition. Finally, we show that under this condition the output of the previously mentioned naïve algorithm is irredundant.

Keywords Computational semigroup theory · Transformation monoids · Asymptotic behavior

1 Introduction

A generating set X of a semigroup S is a *smallest generating set*, also known as *minimum generating set*, if every subset of S with cardinality strictly smaller than $|X|$ does not generate S. The size of a smallest generating set is known as the *rank of* S. Similarly, an *irredundant generating set for* S is a generating set X such that no proper

Communicated by Benjamin Steinberg.

We want to thank J. D. Mitchell for providing the experimental data.

ST was initially supported by EPSRC DTG EP/K503162/1.

✉ Julius Jonušas
j.jonusas@gmail.com

Sascha Troscheit
sascha.troscheit@univie.ac.at

[1] Institut für Diskrete Mathematik und Geometrie, FG Algebra, TU Wien, Vienna, Austria

[2] Faculty of Mathematics, University of Vienna, Vienna, Austria

subset of X is a generating set for S. Of course, the notions of a smallest generating set, irredundant generating set, and the rank have a natural interpretation for groups and other algebraic objects. The question of finding a smallest generating set or a rank is a classical one, see for example [1,13] in the case of quasigroups and [7–9] in the case of semigroups. However, from a computational perspective this is, in general, not an easy problem. In particular, there is no known efficient algorithm to find the rank of a given S, besides examining most of its subsets. As such, fast naïve algorithms are sometimes used to obtain small, but not necessarily smallest, generating sets. The simplest of them is Algorithm 1.

Algorithm 1: Greedy

Input : A list S of all the elements of a semigroup
Output: A generating set X

```
1 X ⟵ ∅;
2 while |⟨X⟩| ≠ |S| do
3     for s ∈ S do
4         if s ∉ ⟨X⟩ then
5             X ⟵ X ∪ {s} ;
```

The algorithm applies to both groups and semigroups. The advantages of the Greedy algorithm are its speed and that it requires no a priori knowledge about the object. The latter might be seen as a drawback if some structural information is known. For semigroups this algorithm can be improved by taking into account its $\mathcal{J}$-class structure.

In order to define the next algorithm, we require some notation. Let S be a semigroup and let 1 be a symbol which is not in S. Define $S^1 = S \cup \{1\}$ to be a semigroup such that for all $x, y \in S$ the product $x \cdot y$ in S^1 is the same as the product in S, and $x \cdot 1 = 1 \cdot x = x$ for all $x \in S^1$. It is routine to verify that the operation $[\cdot]$ on S^1 is associative. If $A, B \subseteq S^1$ define $Ax = \{a \cdot x : a \in A\}$ and similarly define xA and AxB. Define a relation on S by

$$x \leq y \quad \text{if and only if} \quad S^1 x S^1 \subseteq S^1 y S^1.$$

Then $\leq$ is reflexive and transitive, however it might fail to be antisymmetric. In other words, $\leq$ is a *preorder* on S. Clearly, the relation

$$a \mathcal{J} b \quad \text{if and only if} \quad a \geq b \quad \text{and} \quad a \leq b$$

is an equivalence relation on S and the preorder $\leq$ induces a partial order on the equivalence classes of $\mathcal{J}$, which we will also denote by $\leq$ if the distinction is clear from the context. We say that the list $s_{1,1}, \ldots, s_{1,n_1}, s_{2,1}, \ldots, s_{2,n_2} \ldots s_{k,n_k}$ of all elements of S is ordered according to the preorder $\leq$ if $\{s_{i,1}, \ldots, s_{i,n_i}\}$ is an equivalence class of $\mathcal{J}$ for all i and if $s_{i,k} \geq s_{j,m}$ then $i \leq j$. Using this idea we can state Algorithm 2.

If S is a group, then $S^1 x S^1 = S$ for every $x \in S$. Hence there is a single $\mathcal{J}$-class in S and so any permutation of elements of S is ordered according to the preorder $\leq$, and so the SmallGeneratingSet algorithm does not perform any better than Greedy. For proper semigroups the algorithm is particularly useful if the $\mathcal{J}$-class structure is

Algorithm 2: SmallGeneratingSet, (implemented in *Semigroups* [12] for GAP [14])

Input : A semigroup S
Output: A generating set X

1 $L \longleftarrow$ order elements of S according to the preorder $\leq$;
2 $X \longleftarrow Greedy(L)$;

Table 1 Subsemigroups of $\mathcal{T}_3$

Rank	Size of the output						
	1	2	3	4	5	6	7
1	7	3	1	0	0	0	0
2	–	32	25	11	3	1	0
3	–	–	38	50	23	9	2
4	–	–	–	23	28	6	6
5	–	–	–	–	5	7	2

known in advance, for example if the semigroup was enumerated using the *Froidure–Pin algorithm* [6,11] or algorithms appearing in [5]. It is easy to come up with examples for which SmallGeneratingSet might return a generating set which is not a smallest generating set or even an irredundant generating set, for example any non-trivial finite group G. Even though the algorithm is naïve, it performs surprisingly well in practice. For instance, we ran the SmallGeneratingSet algorithm on all 836 021 semigroups (up to (anti-)isomorphism) of size 7, available in *SmallSemi* [4]. In all cases the generating set found was a smallest generating set.

Let $n \in \mathbb{N}$ and let $\mathcal{T}_n$ be the *transformation monoid* on n points, that is the set of all functions from $\{1, \ldots, n\}$ to itself. The set $\mathcal{T}_n$ forms a semigroup under the composition of functions. In the following table, we consider every subgroup of $\mathcal{T}_3$ up to conjugation. Observe that—for most of them—the size of the generating set output by SmallGeneratingSet is equal to the rank or is one greater (Table 1).

The main motivation for this paper is to provide mathematical justification as to why SmallGeneratingSet algorithms often returns a smallest generating set. In order to do so, we consider properties of transformation semigroups picked at random, in a certain way. We say that a semigroup S is *ubiquitous* if every irredundant generating of S is also a smallest generating set. Alternatively, if r is the rank of S, then S is ubiquitous if every irredundant generating set is of size r.

First we will provide a sufficient condition for a transformation semigroup to be ubiquitous.

Theorem 1.1 *Let $S \leq \mathcal{T}_n$ and suppose that X is a generating set for S such that $\operatorname{rank}(xyz) < \operatorname{rank}(y)$ for all $x, y, z \in X$. Then S is ubiquitous.*

Even though we restrict our attention to transformation semigroups in this paper, Theorem 1.1 can be generalised to include semigroups of partial bijections as well. We follow the approach of Cameron [2] of choosing a random transformation semigroup. That is for some $k \geq 1$ we choose k transformations of degree n with uniform probabil-

ity and consider the semigroup generated by them. We show that most transformation semigroups are ubiquitous.

Theorem 1.2 *Let $k \geq 1$, and let $\mathbb{P}_k(n)$ be the probability that for $x_1, \ldots, x_k \in \mathcal{T}_n$, chosen with uniform probability, the semigroup $\langle x_1, \ldots, x_k \rangle$ is ubiquitous. Then $\mathbb{P}_k(n) \to 1$ as $n \to \infty$ exponentially fast.*

Even though SmallGeneratingSet does not return an irredundant generating set in general, we show that under the assumptions of Theorem 1.1 the output is irredundant. Hence the final result of the paper is as follows.

Theorem 1.3 *Let $k \geq 1$, and let $\mathbb{W}_k(n)$ be the probability that for $x_1, \ldots, x_k \in \mathcal{T}_n$, chosen with uniform probability, SmallGeneratingSet returns a smallest generating set for a semigroup $\langle x_1, \ldots, x_k \rangle$. Then $\mathbb{W}_k(n) \to 1$ as $n \to \infty$ exponentially fast.*

Here we only look at the asymptotic behaviour of transformation semigroups, however the same question can be investigated for any other infinite family of semigroups, for example symmetric inverse monoids on $\{1, \ldots, n\}$, or binary relations on n points.

2 Preliminaries

In this section we give the definitions and notation needed in the remainder of the paper.

Definition 2.1 Let S be a semigroup and let $x, y \in S$. The *Green's relations* $\mathcal{L}, \mathcal{R}, \mathcal{J}$, and $\mathcal{D}$ are the following equivalence relations on S:

$$\begin{aligned} x \mathcal{L} y &\quad \text{if and only if} \quad S^1 x = S^1 y \\ x \mathcal{R} y &\quad \text{if and only if} \quad x S^1 = y S^1 \\ x \mathcal{J} y &\quad \text{if and only if} \quad S^1 x S^1 = S^1 y S^1 \end{aligned}$$

and $\mathcal{D}$ is the smallest equivalence relation containing both $\mathcal{L}$ and $\mathcal{R}$.

Let $x \in S$. Then L_x, R_x, and D_x denote the equivalences classes of $\mathcal{L}$, $\mathcal{R}$, and $\mathcal{D}$, respectively, containing x. If S is finite, then $\mathcal{D} = \mathcal{J}$, for a proof see [10]. Since we are only interested in finite semigroups we will not make any distinction between the $\mathcal{D}$ and $\mathcal{J}$ relations.

Throughout the paper, we write elements of $\mathcal{T}_n$ on the right of their argument and we write functions from a subset of $\mathbb{R}^n$ to $\mathbb{R}$ on the left. This is done in agreement with two different notations prevalent in algebra and analysis.

Let $f \in \mathcal{T}_n$, and let $A \subseteq \{1, \ldots, n\}$. Then

$$(A)f = \{(a)f : a \in A\}$$

and the *image of* f is the set $\operatorname{im}(f) = (\{1, \ldots, n\})f$. A *transversal of* f is a set $\mathfrak{T} \subseteq \{1, \ldots, n\}$ such that f is injective on $\mathfrak{T}$ and $(\mathfrak{T})f = \operatorname{im}(f)$. The *rank of* f is

$\operatorname{rank}(f) = |\operatorname{im}(f)| = |\mathfrak{T}|$, where $\mathfrak{T}$ is a transversal of f. The *kernel of* f, denoted by $\ker(f)$, is the equivalence relation defined by

$$(x, y) \in \ker(f) \quad \text{if and only if} \quad (x)f = (y)f.$$

Hence a *kernel class of* f containing $x \in \{1, \ldots, n\}$ is the set

$$\{y \in \{1, \ldots, n\} : (y)f = (x)f\}.$$

Using the above definition we can state a classical result describing Green's classes of transformation semigroups. The proof is easy and thus omitted.

Lemma 2.2 *Let $S \leq \mathcal{T}_n$, and let $f, g \in S$. Then*

(i) *if $f \mathcal{L} g$ then* $\operatorname{im}(f) = \operatorname{im}(g)$*;*
(ii) *if $f \mathcal{R} g$ then* $\ker(f) = \ker(g)$*;*
(iii) *if $f \mathcal{D} g$ then* $\operatorname{rank}(f) = \operatorname{rank}(g)$.

3 Sufficient condition for ubiquitous semigroups

In this section we prove Theorem 1.1. We will do so in a series of lemmas. The first of which is the following easy observation about products in the $\mathcal{D}$-classes. Recall that if $x, y \in S$, then by D_x we denote the $\mathcal{D}$-class containing x and $x \leq y$ if and only if $S^1 x S^1 \subseteq S^1 y S^1$.

Lemma 3.1 *Let S be a semigroup, and let $z_1 \cdots z_m \in D_x$ where $x, z_1, \ldots, z_m \in S$. Then $x \leq z_i \cdots z_j$ under the preorder on S for all $i, j \in \{1, \ldots, m\}$ with $i \leq j$.*

Proof Let $x, z_1, \ldots, z_m \in S$ be such that $z_1 \cdots z_m \in D_x$. Then

$$S^1 x S^1 = S^1 z_1 \cdots z_m S^1 \subseteq S^1 z_i \cdots z_j S^1,$$

and so $x \leq z_i \cdots z_j$ by definition for all $i, j \in \{1, \ldots, m\}$ with $i \leq j$. □

Next we give a condition for a semigroup S which restricts allowed products in a given $\mathcal{D}$-class.

Lemma 3.2 *Let $S \leq \mathcal{T}_n$, let X be a generating set for S, and let $x \in X$ be such that* $\operatorname{rank}(y_1 x y_2) < \operatorname{rank}(x)$ *for all $y_1, y_2 \in X$ where $y_1, y_2 \geq x$. Then only the following products*

$$x, \quad y_1 \cdots y_m, \quad x y_1 \cdots y_m, \quad \text{or} \quad y_1 \cdots y_m x$$

where $m \geq 1$, $y_1, \ldots, y_m \in X \backslash \{x\}$ and $y_1, \ldots, y_m \geq x$ can be in D_x.

Proof First observe that if $x^2 \in D_x$, then both x and x^2 have the same rank by Lemma 2.2, in other words $|\mathrm{im}(x)| = |\mathrm{im}(x^2)|$. However, since x is a finite degree transformation and $\mathrm{im}(x^2) \subseteq \mathrm{im}(x)$, it follows that $\mathrm{im}(x) = \mathrm{im}(x^2)$, and so x acts as a bijection on $\mathrm{im}(x)$. Hence $\mathrm{rank}(x^3) = \mathrm{rank}(x)$, contradicting the hypothesis of the lemma. Therefore $x^2 \notin D_x$, and since $S^1x^2S^1 \subseteq S^1xS^1$, it follows that $x^2 < x$ under the preorder on S. Similarly, for every $y_1, y_2 \in X$ such that $y_1, y_2 \geq x$, it follows from Lemma 2.2 that $y_1xy_2 < x$, since $\mathrm{rank}(y_1xy_2) < \mathrm{rank}(x)$ and $S^1y_1xy_2S^1 \subseteq S^1xS^1$.

Let $z_1, \ldots, z_m \in X$ be such that $z_1 \cdots z_m \in D_x$. Then $x \leq z_i$ for all i by Lemma 3.1. Hence there are $k \in \mathbb{N}$, $n_1, \ldots, n_k, m_2, \ldots, m_k \geq 1$, and $m_1, m_{k+1} \geq 0$ such that

$$z_1 \cdots z_m = y_{1,1} \cdots y_{1,m_1} x^{n_1} y_{2,1} \cdots y_{k,m_k} x^{n_k} y_{k+1,1} \cdots y_{k+1,m_{k+1}}$$

where $y_{i,j} \in X \backslash \{x\}$ and $y_{i,j} \geq x$ for all i and j. Here we are assuming that

$$m = \sum_{i=1}^{k} n_i + \sum_{i=1}^{k+1} m_i,$$

$z_1 = y_{1,1}$, $z_2 = y_{1,2}$, and so on. Again by Lemma 3.1, if $z = z_i \cdots z_j$ is a subproduct of $z_1 \cdots z_m$, then $x \leq z$. However $x^2 < x$, and thus x^2 is not a subproduct of $z_1 \cdots z_m$. That is, $n_i = 1$ for all $i \in \{1, \ldots, k\}$. Hence

$$z_1 \cdots z_m = y_{1,1} \cdots y_{1,m_1} x y_{2,1} \cdots x y_{k+1,1} \cdots y_{k+1,m_{k+1}}.$$

In a similar fashion, if $y_{i,m_i}, y_{i+1,1} \in X \backslash \{x\}$ for $y_{i,m_i}, y_{i+1,1} \geq x$, then as observed above $y_{i,m_i} x y_{i+1,1} < x$, and so $y_{i,m_i} x y_{i+1,1}$ is not a subproduct of $z_1 \cdots z_m$. Hence $z_1 \cdots z_m$ is one of the following products

$$x, \quad y_1 \cdots y_l, \quad x y_1 \cdots y_l, \quad y_1 \cdots y_l x, \quad \text{or} \quad x y_1 \cdots y_l x$$

where $m \geq 1$, $y_1, \ldots, y_l \in X \backslash \{x\}$ and $y_1, \ldots, y_l \geq x$. Hence it remains to show that $xy_1 \cdots y_l x \notin D_x$.

Suppose that $xy_1 \cdots y_l x \in D_x$ for some $l \geq 1$, $y_1, \ldots, y_l \in X \backslash \{x\}$ such that $y_1, \ldots, y_l \geq x$. Then there are $a, b \in S^1$ such that $axy_1 \cdots y_l xb = x$. Note that unless $a = b = 1$, the product $axy_1 \cdots y_l xb$ is not in one of the above forms, and so cannot be in D_x. Hence $a = b = 1$, and thus $xy_1 \cdots y_m l = x$. Which is only possible if $y_1 \ldots y_l$ acts bijectively on $\mathrm{im}(x)$. Thus y_1 acts bijectively on $\mathrm{im}(x)$. If $\mathrm{im}(y_l x) = \mathrm{im}(x)$, then it follows that

$$\mathrm{rank}(y_l x y_1) = |\mathrm{im}(y_l x y_1)| = |\mathrm{im}(y_l x)| = |\mathrm{im}(x)| = \mathrm{rank}(x),$$

which contradicts the hypothesis of the lemma. Hence $\mathrm{im}(y_l x) \subsetneq \mathrm{im}(x)$. However, it then follows that

$$\mathrm{rank}(x y_1 \cdots y_l x) \leq |\,\mathrm{im}(y_l x)| < \mathrm{rank}(x),$$

contradicting $xy_1 \cdots y_l x = x$. Therefore $xy_1 \cdots y_l x \notin D_x$ for all $l \geq 1$ and all $y_1, \ldots, y_l \in X \backslash \{x\}$ such that $y_1, \ldots y_l \geq x$, as required. □

Corollary 3.3 *Let $S \leq \mathcal{T}_n$, let X be a generating set for S, and let $x \in X$ be such that $\operatorname{rank}(z_1 x z_2) < \operatorname{rank}(x)$ for all $z_1, z_2 \in X$ where $z_1, z_2 \geq x$. Then $pxuys \notin D_x$ for all $p, u, s \in S^1$ and any $y \in X$ such that $x \mathcal{D} y$.*

Proof If $x, y \in X$, $x \mathcal{D} y$, and $pxuys \in D_x = D_y$ for some $p, u, s \in S^1$, then there are $a, b, c, d \in S^1$ such that

$$axsyb = x \quad \text{and} \quad cxsyd = y.$$

Hence $axscxsydb = x \in D_x$, but x occurs twice in the product, which is a contradiction according to Lemma 3.2. □

Finally, we prove Theorem 1.1. Observe that if a transformation semigroup $S \leq \mathcal{T}_n$ is such that all irredundant generating sets have the same cardinality, then every irredundant generating set is a smallest generating set.

Theorem 1.1 *Let $S \leq \mathcal{T}_n$ and suppose that X is a generating set for S such that $\operatorname{rank}(xyz) < \operatorname{rank}(y)$ for all $x, y, z \in X$. Then S is ubiquitous.*

Proof Let $X' \subseteq X$ be irredundant. Then $\operatorname{rank}(xyz) < \operatorname{rank}(y)$ for all $x, y, z \in X'$. It is sufficient to show that every irredundant generating set is of the same cardinality. Moreover, without loss of generality we may assume that X is irredundant and show that every irredundant generating set is of size $|X|$.

Let Y be an irredundant generating set for S. Let $\leq_d$ be a total order defined on $\mathcal{D}$-classes of S such that if D and D' are $\mathcal{D}$-classes of S and $D \leq D'$ under the partial order of $\mathcal{D}$-classes, then $D \leq_d D'$. Let $\{D_1, \ldots, D_d\}$ be the set of all $\mathcal{D}$-classes of S, indexed so that $D_d <_d \ldots <_d D_1$. For $k \in \{1, \ldots, d\}$, define

$$X_k = X \cap \left(\bigcup_{i=1}^{k} D_i \right) \quad \text{and} \quad Y_k = Y \cap \left(\bigcup_{i=1}^{k} D_i \right).$$

Let $k \geq 1$ and let $z \in D_k$. By Lemma 3.1 if $x_1 \cdots x_m \in D_k$ where $x_1, \ldots, x_m \in X$, then $z \leq x_i$, and so there is $j \leq k$ so that $x_i \in D_j$ for all $i \in \{1, \ldots, m\}$. In other words,

$$x_1 \cdots x_m \in D_k \quad \text{where} \quad x_1, \ldots, x_m \in X \implies x_i \in X_k \quad \text{for all} \quad i \in \{1, \ldots, m\}. \tag{1}$$

The same argument applies to Y, and so

$$D_k \subseteq \langle X_k \rangle \quad \text{and} \quad D_k \subseteq \langle Y_k \rangle \tag{2}$$

for all $k \geq 1$.

By the definition of the total order $\leq_d$, the $\mathcal{D}$-class D_1 is maximal, and so both X and Y intersect D_1 non-trivially. For any $i \geq 2$ and $x_1, \ldots, x_i \in X_1$, it follows from

Corollary 3.3 that $x_1 \cdots x_i \notin D_{x_1} = D_1$, and so $D_1 = X_1$. The same argument shows that $D_1 = Y_1$, and so $X_1 = Y_1 = D_1$.

For $k \geq 1$, suppose that $|X_k| = |Y_k|$ and $\langle X_k \rangle = \langle Y_k \rangle$. If $X \cap D_{k+1} = \varnothing$, then $D_{k+1} \subseteq \langle X_k \rangle = \langle Y_k \rangle$. Hence $X_{k+1} = X_k$ and $Y_{k+1} = Y_k$, and thus $|X_{k+1}| = |Y_{k+1}|$ and $\langle X_{k+1} \rangle = \langle Y_{k+1} \rangle$.

Suppose that $X \cap D_{k+1} \neq \varnothing$. Then $D_{k+1} \nsubseteq \langle X_k \rangle = \langle Y_k \rangle$, and so $Y \cap D_{k+1} \neq \varnothing$. Suppose that $t \geq 0$ is largest integer such that there is $X' \subseteq X \cap D_{k+1}$ and $Y' \subseteq Y \cap D_{k+1}$ with $|X'| = |Y'| = t$ and $\langle X_k, X' \rangle = \langle Y_k, Y' \rangle$. If $t = |Y \cap D_{k+1}|$ and $x \in X \cap D_{k+1} \backslash X'$, then

$$x \in D_{k+1} \subseteq \langle Y_{k+1} \rangle = \langle Y_k, Y' \rangle = \langle X_k, X' \rangle,$$

by (2). However, this is impossible, since X is irredundant and $x \notin X_k \cup X' \subseteq X$. Hence if $t = |Y \cap D_{k+1}|$ then $X' = X \cap D_{k+1}$, or in other words $X_{k+1} = X_k \cup X'$ and $Y_{k+1} = Y_k \cup Y'$. Therefore, $|X_{k+1}| = |X_k| + t = |Y_k| + t = |Y_{k+1}|$ and $\langle X_{k+1} \rangle = \langle Y_{k+1} \rangle$. We will now show that $t = |Y \cap D_{k+1}|$.

Suppose that $t < |Y \cap D_{k+1}|$. Then there is $y \in Y \cap D_{k+1} \backslash Y'$ and y is equal to a product of elements of X_{k+1} by (2). It follows from Corollary 3.3 that if $x_1 \cdots x_m \in D_{k+1}$ where $x_1, \ldots, x_m \in X$ then there is at most one $i \in \{1, \ldots, m\}$ such that $x_i \in X \cap D_{k+1}$, otherwise some subword of $x_1 \cdots x_m$ would not be an element of D_{k+1}. Since $y \notin Y_k \cup Y'$, the irredundancy of Y implies that $y \notin \langle Y_k, Y' \rangle = \langle X_k, X' \rangle$. It follows that $y = p_1 \cdots p_m x s_1 \cdots s_l$ for some $x \in X \cap D_{k+1} \backslash X'$, $m, l \geq 0$ and $s_i, p_i \in X_k$. Hence $y \in \langle x, X_k \rangle$. Since $x, y \in D_{k+1}$, it follows that there are $a, b \in S^1$ such that

$$a p_1 \cdots p_m x s_1 \cdots s_l b = ayb = x.$$

It follows from (1) and the discussion above that $a, b \in \langle X_k \rangle^1$, and so $x \in \langle y, X_k \rangle$. Moreover

$$x \in \langle y, X_k, X' \rangle \quad \text{and} \quad y \in \langle x, X_k, X' \rangle.$$

Therefore $\langle x, X_k, X' \rangle = \langle y, X_k, X' \rangle = \langle y, Y_k, Y' \rangle$, since $\langle X_k, X' \rangle = \langle Y_k, Y' \rangle$. However $|X' \cup \{x\}| = |Y' \cup \{y\}| = t+1$, which contradicts the maximality of t. Therefore $t = |Y \cap D_{k+1}|$ and by the previous paragraph $\langle X_{k+1} \rangle = \langle Y_{k+1} \rangle$ and $|X_{k+1}| = |Y_{k+1}|$.

By induction it follows that $\langle X_k \rangle = \langle Y_k \rangle$ and $|X_k| = |Y_k|$ for all $k \in \{1, \ldots, d\}$. In particular, $X_d = X$ and $Y_d = Y$, and thus $|X| = |Y|$, as required. □

4 SmallGeneratingSet

In this section we return to the motivating question about the algorithm SmallGeneratingSet. First, we note that SmallGeneratingSet might return a generating set which is not irredundant. For example, if the semigroup under investigation is a group of size at least 2, the algorithm can first pick an identity and so return a generating set which

includes an identity. However, we show that under the assumptions of Theorem 1.1 the generating set returned by SmallGeneratingSet is irredundant.

Lemma 4.1 *Let $S \leq \mathcal{T}_n$ and suppose that X is a generating set for S such that $\operatorname{rank}(xyz) < \operatorname{rank}(y)$ for all $x, y, z \in X$. Then SmallGeneratingSet returns an irredundant generating set.*

Proof Let $X = \{x_1, \ldots, x_m\}$ be the output of the algorithm, and assume that the elements were selected in the order they are listed. Suppose that $I \subseteq X$ is irredundant and let $x_i \in X \backslash I$. Since x_i was selected by the algorithm, it means that

$$x_i \notin \langle x_1, \ldots, x_{i-1} \rangle,$$

and so there exists $x_j \in I$ such that $x_i \mathcal{D} x_j$ and $j > i$, otherwise $x_i \notin \langle I \rangle$. Without loss of generality, we can assume that i is the largest integer such that $x_i \in X \backslash I$ and $x_i \mathcal{D} x_j$. Then there are $a_1, \ldots, a_{k_a}, b_1, \ldots, b_{k_b} \in I$ such that

$$a_1 \cdots a_{k_a} x_i b_1 \cdots b_{k_b} = x_j.$$

Since $x_j \notin \langle x_1, \ldots, x_i \rangle$, it follows that at least one of the $a_1, \ldots, a_{k_a}, b_1, \ldots, b_{k_b}$ is x_k for some $k > i$. It follows from Lemma 3.1 that $x_k \geq x_j$, and since $k > i$ implies that $x_k \not> x_i$, we have that $x_k \mathcal{D} x_i$.

Finally, there are $c_1, \ldots, c_{k_c}, d_1, \ldots, d_{k_d} \in I$ such that

$$c_1 \cdots c_{k_c} x_j d_1 \cdots d_{k_d} = x_i,$$

and so

$$a_1 \cdots a_{k_a} c_1 \cdots c_{k_c} x_j d_1 \cdots d_{k_d} b_1 \cdots b_{k_b} = x_j.$$

Which contradicts Corollary 3.3 as at least one of $a_1, \ldots, a_{k_a}, b_1, \ldots, b_{k_b}$ is x_k. Therefore, $X = I$. □

The following result is then immediate from Theorem 1.1.

Corollary 4.2 *Let $S \leq \mathcal{T}_n$ and suppose that X is a generating set for S such that $\operatorname{rank}(xyz) < \operatorname{rank}(y)$ for all $x, y, z \in X$. Then SmallGeneratingSet returns a smallest generating set.*

5 Asymptotics

The main aim of this section is to show that if for some fixed $k \geq 1$ we choose $x_1, \ldots, x_k \in \mathcal{T}_n$ with uniform probability, then the probability $\mathbb{P}_k(n)$ that $\langle x_1, \ldots, x_k \rangle$ is ubiquitous and the probability $\mathbb{W}_k(n)$ that SmallGeneratingSet returns a smallest generating set for $\langle x_1, \ldots, x_k \rangle$ both tend to 1 as n increases.

Lemma 5.1 *Let $X \subseteq \mathcal{T}_n$ be such that* $\operatorname{rank}(xyz) = \operatorname{rank}(y)$ *for some* $x, y, z \in X$. *Then one of the following holds:*

(i) *there is $x \in X$ such that $\langle x \rangle$ is a group;*
(ii) *there are distinct $x, y \in X$ such that* $\operatorname{rank}(xyx) = \operatorname{rank}(y)$*;*
(iii) *there are mutually distinct $x, y, z \in X$ such that* $\operatorname{rank}(xyz) = \operatorname{rank}(y)$.

Proof Suppose that $\operatorname{rank}(xyz) = \operatorname{rank}(y)$ for some $x, y, z \in X$ and suppose that not all x, y, and z are distinct. If $x = y = z$, then $\operatorname{rank}(x^3) = \operatorname{rank}(x)$, which is only possible if x acts bijectively on $\operatorname{im}(x)$. However, in that case $\langle x \rangle$ is a group. Hence we only need to consider the case that where exactly two of x, y, and z are equal.

Suppose that $x = y$. Then $\operatorname{rank}(y^2 z) = \operatorname{rank}(y)$, and since

$$\operatorname{rank}(y) \leq \operatorname{rank}(y^2) \leq \operatorname{rank}(y^2 z) = \operatorname{rank}(y),$$

it follows that $\operatorname{rank}(y^2) = \operatorname{rank}(y)$. Hence by an argument similar to above $\langle y \rangle$ is a group. The case $y = z$ can be dealt with in an almost identical fashion. Therefore, there are distinct $x, y \in X$ such that $\operatorname{rank}(xyx) = \operatorname{rank}(y)$. □

In order to show that $\mathbb{P}_k(n) \to 1$ as $n \to \infty$, for every $n \in \mathbb{N}$, we define three probabilities:

$\mathbb{G}_n$ is the probability that $\langle x \rangle$ is a group where $x \in \mathcal{T}_n$ is chosen randomly with uniform probability

$\mathbb{T}_n$ is the probability that $\operatorname{rank}(xyx) = \operatorname{rank}(y)$ where $x, y \in \mathcal{T}_n$ are chosen randomly with uniform probability

$\mathbb{V}_n$ is the probability that $\operatorname{rank}(xyz) = \operatorname{rank}(z)$ where $x, y, z \in \mathcal{T}_n$ are chosen randomly with uniform probability.

For a fixed $k \geq 1$, if $x_1, \ldots, x_k \in \mathcal{T}_n$ are chosen randomly with uniform probability, it follows from Lemma 5.1 that the probability that there are $x, y, z \in \{x_1, \ldots, x_k\}$ such that $\operatorname{rank}(xyz) = \operatorname{rank}(y)$ is bounded from above by

$$k\mathbb{G}_n + k(k-1)\mathbb{T}_n + k(k-1)(k-2)\mathbb{V}_n.$$

Hence by Theorem 1.1

$$\mathbb{P}_k(n) \geq 1 - k\mathbb{G}_n - k(k-1)\mathbb{T}_n - k(k-1)(k-2)\mathbb{V}_n,$$

and the same lower bound hold for $\mathbb{W}_k(n)$ by Corollary 4.2. Hence in order to prove Theorems 1.2 and 1.3it suffices to show that $\mathbb{G}_n \to 0$, $\mathbb{T}_n \to 0$, and $\mathbb{V}_n \to 0$ as $n \to \infty$. We will do so in the remaining three subsections of the paper.

5.1 Preliminary counting results

For $n, r \in \mathbb{N}$ such that $r \leq n$, define $\mathcal{A}(n, r)$ to be the set of partitions of $\{1, \ldots, n\}$ into r non-empty components.

Lemma 5.2 *Let $n, r \in \mathbb{N}$ such that $r \leq n$. Then*

$$\sum_{\{A_1, \ldots, A_r\} \in \mathcal{A}(n,r)} \prod_{i=1}^{r} |A_i| = \binom{n}{r} r^{n-r}.$$

Proof A function $f \in \mathcal{T}_n$ is called idempotent if $f^2 = f$. We prove the lemma by finding the number of idempotent transformation of $\mathcal{T}_n$ of rank r in two ways. Denote this number by N. It can be shown that f is an idempotent if and only if $(x)f = x$ for all $x \in \operatorname{im}(f)$.

If $f \in \mathcal{T}_n$ is an idempotent of rank r, then there are $\binom{n}{r}$ choices for the $\operatorname{im}(f)$ and for every point in $\{1, \ldots, n\} \setminus \operatorname{im}(f)$ there are r choices in $\operatorname{im}(f)$ to map to. Hence

$$N = \binom{n}{r} r^{n-r}.$$

On the other hand, the sets $A_1, \ldots, A_r$ are the kernel classes of $f \in \mathcal{T}_n$ if and only if $\{A_1, \ldots, A_r\} \in \mathcal{A}(n, r)$. If f is an idempotent and $A_1, \ldots, A_r$ are kernel classes of f then $(A_i)f \in A_i$ for all $i \in \{1, \ldots, r\}$, and so there are $\prod_{i=1}^{r} |A_i|$ choices for the $\operatorname{im}(f)$. Hence

$$N = \sum_{\{A_1, \ldots, A_r\} \in \mathcal{A}(n,r)} \prod_{i=1}^{r} |A_i|,$$

as required. □

Since $|\mathcal{A}(n, r)| = \left\{ {n \atop r} \right\}$, the following easy upper bound for the Stirling numbers is an immediate consequence of Lemma 5.2.

Corollary 5.3 *Let $n, r \in \mathbb{N}$ be such that $r \leq n$. Then*

$$\left\{ {n \atop r} \right\} \leq \binom{n}{r} r^{n-r}.$$

We will make use of Stirling's approximation formula

$$\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} \leq n! \leq \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n+\frac{1}{2n}}.$$

If $F : \mathbb{R} \to \mathbb{R}$, then we say that $G \in O(F)$ if there are $c > 0$ and $x_0 \in \mathbb{R}$ such that $|G(x)| \leq c|F(x)|$ for all $x \geq x_0$. Then Stirling's formula can be written as follows

$$\log n! = n \log n - n + O(\log(n)).$$

Let $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$. The final notion required in this paper is the function $W : \mathbb{R}^+ \to \mathbb{R}^+$ defined so that

$$x = W(x)e^{W(x)}$$

for all $x \in \mathbb{R}^+$. Since the function $x \mapsto xe^x$ is strictly increasing on $\mathbb{R}^+$, it follows that $W(x)$ is a well-defined function on $\mathbb{R}^+$. In the literature $W(x)$ is known as *Lambert W function* or *product logarithm*, see e.g. [3]. The value $\Omega = W(1)$ is known as the *omega constant* and it satisfies $\Omega e^{\Omega} = 1$, with the numerical value $\Omega = 0.5671439\ldots$.

5.2 $\mathbb{G}_n$ tends to zero

We begin by obtain an expression for $\mathbb{G}_n$ in terms of n.

Lemma 5.4 *Let $n \in \mathbb{N}$. Then*

$$\mathbb{G}_n = \frac{n!}{n^n} \sum_{k=0}^{n-1} \frac{(n-k)^k}{k!}.$$

Proof First observe that for any $x \in \mathcal{T}_n$, the semigroup $\langle x \rangle$ is a group if and only if x acts as a bijection on $\operatorname{im}(x)$. There are

$$\sum_{r=1}^{n} \binom{n}{r} r^{n-r} r!$$

transformations x such that x acts bijectively on $\operatorname{im}(x)$. That is, if $|\operatorname{im}(x)| = r$, then there are $\binom{n}{r}$ choices for $\operatorname{im}(x)$, $r!$ ways of bijectively mapping $\operatorname{im}(x)$ to itself, and r^{n-r} ways to map every point from $\{1, \ldots, n\} \setminus \operatorname{im}(x)$ to $\operatorname{im}(x)$. Since $|\mathcal{T}_n| = n^n$, the probability of randomly choosing $x \in \mathcal{T}_n$ such that $\langle x \rangle$ is a group is

$$\mathbb{G}_n = \frac{1}{n^n} \sum_{r=1}^{n} \binom{n}{r} r^{n-r} r! = \frac{n!}{n^n} \sum_{r=1}^{n} \frac{r^{n-r}}{(n-r)!}.$$

Finally, rewriting the equation using $k = n - r$ we obtain

$$\frac{n!}{n^n} \sum_{r=1}^{n} \frac{r^{n-r}}{(n-r)!} = \frac{n!}{n^n} \sum_{k=0}^{n-1} \frac{(n-k)^k}{k!},$$

as required. □

In order to prove that $\mathbb{G}_n \to 0$ as $n \to \infty$ we use an auxiliary function for which we prove some analytical properties. Also recall that $\Omega \in \mathbb{R}$ is a unique constant which satisfies $\Omega e^{\Omega} = 1$.

Lemma 5.5 *Let $F : (0, 1) \to \mathbb{R}$ be given by $F(x) = x \log(x^{-1} - 1) + x$. Then F has a unique maximum at $\alpha = \frac{\Omega}{1+\Omega} \in (0, 1)$ and $F(\alpha) = \Omega < 1$.*

Proof First observe that $F(x)$ is continuous on $(0, 1)$, and $F(x) \to 0$ as $x \to 0$ and $F(x) \to -\infty$ as $x \to 1$. The first and second derivative are continuous and given by

$$\frac{dF(x)}{dx} = 1 - \frac{1}{1-x} + \log(x^{-1} - 1) \quad \text{and} \quad \frac{d^2F(x)}{dx^2} = -\frac{1}{(x-1)^2 x}.$$

Clearly, $\frac{d^2F(x)}{dx^2} < 0$ for all $x \in (0, 1)$, but $\frac{dF(x)}{dx} \to \infty$ as $x \to 0$ and so the derivative is positive in a neighbourhood of 0. But $F(x) \to -\infty$ as $x \to 1$ and thus F has a unique maximum at α implicitly given by

$$1 - \frac{1}{1-\alpha} + \log\left(\frac{1-\alpha}{\alpha}\right) = 0,$$

or in other words

$$\frac{\alpha}{1-\alpha} = \log\left(\frac{1-\alpha}{\alpha}\right).$$

It then follows that $\frac{\alpha}{1-\alpha} = \Omega$, by the definition of Ω. Hence $\alpha = \frac{\Omega}{1+\Omega}$ and

$$F(\alpha) = \alpha \log\left(\frac{1-\alpha}{\alpha}\right) + \alpha = \alpha\left(1 + \frac{\alpha}{1-\alpha}\right) = \frac{\alpha}{1-\alpha} = \Omega.$$

□

Finally, we conclude this section by describing the asymptotic behaviour of $\mathbb{G}_n$.

Proposition 5.6 *The probability $\mathbb{G}_n$, that $\langle x \rangle$ is a group where $x \in \mathcal{T}_n$ is chosen with uniform distribution, tends to 0 exponentially at the rate less than $1 - \Omega$.*

Proof By Lemma 5.4

$$\mathbb{G}_n = \frac{n!}{n^n} \sum_{k=0}^{n-1} \frac{(n-k)^k}{k!}.$$

We use the Stirling approximation $\log n! = n \log n - n + O(\log(n))$. Then

$$\frac{\log \mathbb{G}_n}{n} = n^{-1} O(\log(n)) - 1 + n^{-1} \log \sum_{k=0}^{n-1} \frac{(n-k)^k}{k!}.$$

Note that the last term can be bounded from above and below in the following way

$$\log\left(\max_{k\in\{0,\ldots,n-1\}} \frac{(n-k)^k}{k!}\right) \leq \log \sum_{k=0}^{n-1} \frac{(n-k)^k}{k!} \leq \log\left(n \max_{k\in\{0,\ldots,n-1\}} \frac{(n-k)^k}{k!}\right).$$

Hence

$$\log \sum_{k=0}^{n-1} \frac{(n-k)^k}{k!} = \log \left(\max_{k\in\{0,\dots,n-1\}} \frac{(n-k)^k}{k!} \right) + O(\log n),$$

and so

$$\lim_{n\to\infty} \frac{\log \mathbb{G}_n}{n} = -1 + \lim_{n\to\infty} n^{-1} \log \left(\max_{k\in\{0,\dots,n-1\}} \frac{(n-k)^k}{k!} \right).$$

Considering the second term in the above equation, noting that for $n \geq 3$ the maximum does not occur at $k = 0$, it follows that

$$\begin{aligned} n^{-1} \log \left(\max_{k\in\{0,\dots,n-1\}} \frac{(n-k)^k}{k!} \right) &= \max_{k\in\{1,\dots,n-1\}} n^{-1} \log \left(\frac{(n-k)^k}{k!} \right) \\ &= \max_{k\in\{1,\dots,n-1\}} n^{-1} \big(k \log(n-k) - k \log k + k \\ &\quad - O(\log k) \big) \\ &= \max_{x\in M_n} \left(x \log(x^{-1} - 1) + x \right) - n^{-1} O(\log n), \end{aligned}$$

where $M_n = \{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\}$. Since F is continuous on $(0, 1)$ we conclude that $\max_{x\in M_n} F(x) \to \max_{x\in(0,1)} F(x)$ as $n \to \infty$. Therefore

$$\lim_{n\to\infty} \frac{\log \mathbb{G}_n}{n} = -1 + \lim_{n\to\infty} n^{-1} \log \left(\max_{k\in\{0,\dots,n-1\}} \frac{(n-k)^k}{k!} \right) = F(\alpha) - 1 = \Omega - 1 < 0,$$

by Lemma 5.5 as required. □

5.3 $\mathbb{T}_n$ tends to zero

Recall that for $n, r \in \mathbb{N}$ such that $r \leq n$, $\mathcal{A}(n, r)$ denotes the set of partitions of $\{1, \dots, n\}$ into r non-empty components. Similarly, define $\mathcal{B}(n, r)$ to be the set of subsets of $\{1, \dots, n\}$ of cardinality r. Then $|\mathcal{B}(n, r)| = \binom{n}{r}$.

Lemma 5.7 *Let $n \in \mathbb{N}$. Then the probability that* $\operatorname{rank}(xyx) = \operatorname{rank}(y)$*, where $x, y \in \mathcal{T}_n$ are chosen with uniform probability, is*

$$\mathbb{T}_n = \frac{1}{n^{2n}} \sum_{r=1}^{n} \binom{n}{r} r! \sum_{k=1}^{r} \left\{ {r \atop k} \right\} k! k^{n-r} \sum_{\{A_1,\dots,A_r\}\in\mathcal{A}(n,r)} \sum_{B\in\mathcal{B}(r,k)} \prod_{i\in B} |A_i|.$$

Proof Let $x, y \in \mathcal{T}_n$ be such that $\operatorname{rank}(xyx) = \operatorname{rank}(y)$. We first show that $\operatorname{im}(xy)$ is contained in a transversal of x. Let $\mathfrak{T}$ be a transversal of xyx. Then xyx is injective on $\mathfrak{T}$ by definition, and so x is injective on $(\mathfrak{T})xy$. Hence $\operatorname{im}(xy) = (\mathfrak{T})xy$ is contained in a transversal of x.

Suppose that $\operatorname{rank}(x) = r$, $\operatorname{rank}(y) = k$, and $\{A_1, \ldots, A_r\} \in \mathcal{A}(n, r)$ are the kernel classes of x. Then there are $\binom{n}{r} r!$ choices for x. Since

$$\operatorname{rank}(y) \geq \operatorname{rank}(xy) \geq \operatorname{rank}(xyx) = \operatorname{rank}(y),$$

it follows that $\operatorname{rank}(xy) = \operatorname{rank}(y) = k$, and also $\operatorname{im}(y) = \operatorname{im}(xy)$. Since x is injective on $\operatorname{im}(xy)$, there are

$$\sum_{B \in \mathcal{B}(r,k)} \prod_{i \in B} |A_i|$$

choices for $\operatorname{im}(y) = \operatorname{im}(xy)$. That is, $\operatorname{im}(xy)$ contains at most one point from any kernel class of x. Since $(\operatorname{im}(x))y = \operatorname{im}(xy) = \operatorname{im}(y)$, there are $\left\{ {r \atop k} \right\} k!$ ways for y to map $\operatorname{im}(x)$ to $\operatorname{im}(y)$. Finally, $(\{1, \ldots, n\}) \setminus \operatorname{im}(x))y \subseteq \operatorname{im}(y)$, and so there k^{n-r} for y to map $(\{1, \ldots, n\}) \setminus \operatorname{im}(x))$ to $\operatorname{im}(y)$. Hence there are in total

$$\left\{ {r \atop k} \right\} k! k^{n-r} \sum_{B \in \mathcal{B}(r,k)} \prod_{i \in B} |A_i|$$

choices for y. Therefore

$$\mathbb{T}_n = \frac{1}{n^{2n}} \sum_{r=1}^{n} \sum_{k=1}^{r} \sum_{\{A_1, \ldots, A_r\} \in \mathcal{A}(n,r)} \binom{n}{r} r! \left\{ {r \atop k} \right\} k! k^{n-r} \sum_{B \in \mathcal{B}(r,k)} \prod_{i \in B} |A_i|,$$

since $|\mathcal{T}_n| = n^n$. □

Next, we simplify the expression for $\mathbb{T}_n$.

Lemma 5.8 *Let $n, r, k \in \mathbb{N}$ such that $k \leq r \leq n$. Then*

$$\sum_{\{A_1, \ldots, A_r\} \in \mathcal{A}(n,r)} \sum_{B \in \mathcal{B}(r,k)} \prod_{i \in B} |A_i| = \sum_{s=k}^{n+k-r} \binom{n}{s} \left\{ {n-s \atop r-k} \right\} \binom{s}{k} k^{s-k}.$$

Proof Let $B \in \mathcal{B}(r, k)$ and $\{A_1, \ldots, A_r\} \in \mathcal{A}(n, r)$ be fixed and denote the number $|\bigcup \{A_b : b \in B\}|$ by s. Note that every A_i is non-empty, so $k \leq s \leq n - (r - k)$. Now suppose that only B is fixed, then for every value of $s \in \{k, \ldots, n + k - r\}$, there are $\binom{n}{s}$ choices for $\bigcup \{A_b : b \in B\}$, and there are $\left\{ {n-s \atop r-k} \right\}$ many choices to choose $\{A_b : b \notin B\}$. Hence we can write

$$\sum_{\{A_1, \ldots, A_r\} \in \mathcal{A}(n,r)} \sum_{B \in \mathcal{B}(r,k)} \prod_{i \in B} |A_i| = \sum_{s=k}^{n+k-r} \binom{n}{s} \left\{ {n-s \atop r-k} \right\} \sum_{\{A_1, \ldots, A_k\} \in \mathcal{A}(s,k)} \prod_{i=1}^{k} |A_i|.$$

The result follows by Lemma 5.2. □

Finally, we prove the main lemma of this section.

Lemma 5.9 *There exist $r \in (0,1)$ and $c > 0$ such that $\mathbb{T}_n \leq cn^{7/2}r^n$.*

Proof Note that by Stirling's approximation there are constants $a, b > 0$ such that $an^n e^{-n} \leq n! \leq bn^{n+\frac{1}{2}}e^{-n}$ for all $n \in \mathbb{N}$.

It follows from Lemmas 5.3, 5.7, and 5.8 that

$$\begin{aligned} n^{2n}\mathbb{T}_n &= \sum_{r=1}^{n} \binom{n}{r} r! \sum_{k=1}^{r} \left\{ {r \atop k} \right\} k! k^{n-r} \sum_{s=k}^{n+k-r} \binom{n}{s} \left\{ {n-s \atop r-k} \right\} \binom{s}{k} k^{s-k} \\ &\leq \sum_{r=1}^{n} \binom{n}{r} r! \sum_{k=1}^{r} \binom{r}{k} k^{n-k} k! \sum_{s=k}^{n+k-r} \binom{n}{s} \binom{s}{k} \binom{n-s}{r-k} (r-k)^{n-s-r+k} k^{s-k}. \end{aligned}$$

Observe that

$$\binom{n}{s}\binom{s}{k} = \binom{n}{k}\binom{n-k}{n-s} \quad \text{and} \quad \binom{n-k}{n-s}\binom{n-s}{r-k} = \binom{n-k}{r-k}\binom{n-r}{s-k}. \tag{3}$$

Hence

$$\begin{aligned} n^{2n}\mathbb{T}_n &\leq \sum_{r=1}^{n} \binom{n}{r} r! \sum_{k=1}^{r} \binom{r}{k} k^{n-k} k! \sum_{s=k}^{n+k-r} \binom{n}{k} \binom{n-k}{r-k} \binom{n-r}{s-k} (r-k)^{n-s-r+k} k^{s-k} \\ &= \sum_{r=1}^{n} \binom{n}{r} r! \sum_{k=1}^{r} \binom{r}{k} k^{n-k} k! \binom{n}{k} \binom{n-k}{r-k} \sum_{i=0}^{n-r} \binom{n-r}{i} (r-k)^{n-r-i} k^i \\ &= \sum_{r=1}^{n} \binom{n}{r} r! \sum_{k=1}^{r} \binom{r}{k} k^{n-k} k! \binom{n}{k} \binom{n-k}{r-k} r^{n-r}. \end{aligned}$$

It can also be show that

$$\binom{n}{k}\binom{n-k}{r-k} = \binom{n}{r}\binom{r}{k}, \tag{4}$$

and so

$$n^{2n}\mathbb{T}_n \leq \sum_{r=1}^{n} \binom{n}{r}^2 r! \sum_{k=1}^{r} \binom{r}{k}^2 k! k^{n-k} r^{n-r} = \sum_{r=1}^{n} \frac{n!^2}{(n-r)!^2} \sum_{k=1}^{r} \frac{r!}{k!(r-k)!^2} k^{n-k} r^{n-r}.$$

Hence using Stirling's formula there is a constant $c > 0$ such that

$$n^{2n}\mathbb{T}_n \leq c \sum_{r=1}^{n} \frac{n^{2n+1}e^{-2n}}{(n-r)^{2(n-r)}e^{-2(n-r)}} \sum_{k=1}^{r} \frac{r^{r+\frac{1}{2}}e^{-r}}{e^{-k}k^k e^{-2(r-k)}(r-k)^{2(r-k)}} k^{n-k} r^{n-r},$$

which can be simplified to

$$\begin{aligned}\mathbb{T}_n &\leq c\sum_{r=1}^{n}\sum_{k=1}^{r}\frac{nr^{n+\frac{1}{2}}k^{n-2k}}{e^{r+k}(n-r)^{2(n-r)}(r-k)^{2(r-k)}}\\ &\leq cn^2\max_{\substack{1\leq r\leq n\\ 1\leq k\leq r}}\left\{\frac{nr^{n+\frac{1}{2}}k^{n-2k}}{e^{r+k}(n-r)^{2(n-r)}(r-k)^{2(r-k)}}\right\}.\end{aligned}$$

Let $x, y \in [0, 1]$ be such that $r = xn$ and $k = yr = xyn$. Then

$$\begin{aligned}\mathbb{T}_n &\leq cn^2\max_{\substack{1\leq r\leq n\\ 1\leq k\leq r}}\left\{n^{\frac{3}{2}}\frac{x^{n+\frac{1}{2}}(xy)^{n-2k}}{e^{r+k}(1-x)^{2(n-r)}(x-xy)^{2(r-k)}}\right\}\\ &\leq cn^{\frac{7}{2}}\sup_{(x,y)\in[0,1]^2}\left\{\frac{x^{2(n-xn)+\frac{1}{2}}y^{n-2xyn}}{e^{xn+xyn}(1-x)^{2(n-xn)}(1-y)^{2(xn-xyn)}}\right\}\\ &\leq cn^{\frac{7}{2}}\left(\sup_{(x,y)\in[0,1]^2}\left\{\frac{x^{2(1-x)}y^{1-2xy}}{e^{x(1+y)}(1-x)^{2(1-x)}(1-y)^{2x(1-y)}}\right\}\right)^n.\end{aligned}$$

It only remains to show that the supremum in the above equation is less than 1. In order to do so, define $F : [0, 1]^2 \to \mathbb{R}$ by

$$F(x, y) = \frac{x^{2(1-x)}y^{1-2xy}}{e^{x(1+y)}(1-x)^{2(1-x)}(1-y)^{2x(1-y)}}.$$

Note that F is continuous on a compact set $[0, 1]^2$, and so has a maximum. Hence we only need to consider the boundary of the domain and stationary points of F, that is points in $[0, 1]^2$ where $\partial F/\partial x = 0 = \partial F/\partial y$. However, while it can be immediately be deduced from plots, using any mathematical software, that the maximum of F is strictly less than 1, we show it here analytically. To this end, define the functions $F_1, F_3 : [0, 1] \to \mathbb{R}$ and $F_2 : [0, 1]^2 \to \mathbb{R}$ by

$$F_1(x) = \frac{x^{2(1-x)}}{(1-x)^{2(1-x)}}, \quad F_2(x, y) = \frac{y^{1-2xy}}{e^{x(1+y)}(1-y)^{2x(1-y)}},$$

and

$$F_3(y) = -1 - y - 2(1-y)\log(1-y) - 2y\log y.$$

Then $F(x, y) = F_1(x)F_2(x, y)$, and it can be shown that $\partial F_2(x, y)/\partial x = F_2(x, y)F_3(y)$. Also note that that F_1, F_2, and F_3 are all continuous.

Since $F_1(x)$ is continuous on a compact set, we can perform standard analysis of stationary points. Then

$$\frac{dF_1(x)}{dx} = F_1(x)(1 + x\log(1-x) - x\log x).$$

and $F_1(x) > 0$ for all $x \in (0, 1]$. Thus the stationary points of F_1 are either 0, 1, or x_0, which is given by the equation

$$(1 + x_0\log(1-x_0) - x_0\log x_0) = 0,$$

or in other words, $x_0 = 1/(1+W(e^{-1}))$ where W is the Lambert-W function. It follows that F_1 is bounded from above by $\max\{F_1(0), F_1(1), F(x_0)\}$. A simple algebraic manipulation gives

$$F_1(x_0) = W(e^{-1})^{-\frac{2}{1+W(e^{-1})^{-1}}} < 1.75.$$

Since $F_1(0) = 0$ and $F_1(1) = 1$, it follows that $F_1(x) \leq 1.75$ for all $x \in [0, 1]$. We also note here, that $dF_1(x)/dx$ is positive for all $x \in [0, x_0)$.

Next we show that $\partial F_2(x, y)/\partial x \leq 0$ for all $xy \in [0, 1]$. First, observe that $F_2(x, y) \geq 0$ over $[0, 1]^2$. Since $\partial F_2(x, y)/\partial x = F_2(x, y)F_3(y)$, we are left to show that $F_3(y) \leq 0$ for $y \in [0, 1]$. Note that $F_3(0) = -1$, $F_3(1) = -2$, and

$$\frac{dF_3(y)}{dy} = -1 + 2\log(1-y) - 2\log y \quad \text{and} \quad \frac{d^2F_3(y)}{dy^2} = \frac{2}{(y-1)y}.$$

Since $d^2F_3(y)/dy^2 < 0$ for all $y \in (0, 1)$, F_3 has a unique maximum at $(1+\sqrt{e})^{-1}$, and

$$F_3\left(\frac{1}{1+\sqrt{e}}\right) = 2\log(1+\sqrt{e}) - 2 < 0.$$

Hence $\partial F_2(x, y)/\partial x \leq 0$ for all $x, y \in [0, 1]$, and so $F_2(x, y) \leq F_2(0, y) = y$ and in particular $F_2(x, y) \leq 1$.

For the last step of the proof consider

$$F_2\left(\frac{1}{2}, y\right) = e^{-\frac{y+1}{2}}\left(\frac{y}{1-y}\right)^{1-y}.$$

Then

$$\frac{dF_2\left(\frac{1}{2}, y\right)}{dy} = F_2\left(\frac{1}{2}, y\right)\left(y^{-1} - \frac{1}{2} + \log\left(y^{-1} - 1\right)\right)$$

and the derivative has a single root at $y_0 = 1/(1 + W(e^{-1/2}))$. Hence if $x \in [1/2, 1]$ and $y \in [0, 1]$, then

$$F_2(x, y) \leq F_2(1/2, y) \leq \max\{F_2(1/2, 0), F_2(1/2, 1), F_2(1/2, y_0)\} < 0.56,$$

and so $F(x, y) \leq 1.75 \cdot 0.56 = 0.98$. Since $F(x, y)$ continuous on $[0, 1]^2$ there is $\varepsilon > 0$ and $\beta < 1$ such that $F(x, y) \leq \beta$ for all $x \in [1/2 - \varepsilon, 1]$ and all $y \in [0, 1]$.

Finally, recall that $x_0 = 1/(1+W(e^{-1})) > 0.78$ and $F_1(x)$ is increasing on $[0, x_0)$. We observe that if $x \in [0, 1/2 - \varepsilon] \subseteq [0, x_0)$ then $0 = F_1(0) \leq F_1(x) \leq F_1(1/2 - \varepsilon) < F_1(1/2) = 1$. Since $F_2(x, y) \leq 1$, it follows that $F(x, y) \leq F_1(1/2-\varepsilon) < 1$ for all $x \in [0, 1/2-\varepsilon]$ and all $y \in [0, 1]$. Therefore $F(x, y) \leq \max\{\beta, F_1(1/2-\varepsilon)\} < 1$ for all $x, y \in [0, 1]$, as required. □

The following is an immediate corollary of Lemmas 5.7 and 5.9.

Corollary 5.10 *The probability $\mathbb{T}_n$, that $\operatorname{rank}(xyx) = \operatorname{rank}(y)$ where $x, y \in \mathcal{T}_n$ are chosen with uniform distribution, tends to 0 as $n \to \infty$ exponentially fast.*

5.4 $\mathbb{V}_n$ tends to zero

We start by finding an expression for $\mathbb{V}_n$ in terms of n. The argument is similar to the proof of Lemma 5.7.

Lemma 5.11 *Let $n \in \mathbb{N}$. Then the probability that $\operatorname{rank}(xyz) = \operatorname{rank}(y)$, where $x, y, z \in \mathcal{T}_n$ are chosen with uniform probability, is*

$$\mathbb{V}_n = \frac{1}{n^{3n}} \sum_{r=1}^{n} \sum_{k=1}^{r} \sum_{t=1}^{\min(r,k)} \left\{ {n \atop r} \right\} \binom{n}{r} r! \binom{n}{k} k! \left\{ {r \atop t} \right\} t! t^{n-r} \sum_{s=t}^{n+t-k} \binom{n}{s} \binom{n-s}{r-k} \binom{s}{t} t^{s-t}.$$

Proof If $x, y, z \in \mathcal{T}_n$ are such that $\operatorname{rank}(xyz) = \operatorname{rank}(y)$. We first show that $\operatorname{im}(xy)$ is contained in a transversal of z. Let $\mathfrak{T}$ be a transversal of xyz. Then xyz is injective on $\mathfrak{T}$ by definition, and so z is injective on $(\mathfrak{T})xy$. Hence $\operatorname{im}(xy) = (\mathfrak{T})xy$ is contained in a transversal of z.

Suppose that $\operatorname{rank}(x) = r$, $\operatorname{rank}(z) = k$, $\operatorname{rank}(y) = t$, and $\{A_1, \ldots, A_k\} \in \mathcal{A}(n, k)$ are the kernel classes of z. Note that $t \leq r$ and $t \leq k$. Then there are $\binom{n}{r} \left\{ {n \atop r} \right\} r!$ choices for x and $\binom{n}{k} k!$ choices for z. Since

$$\operatorname{rank}(y) \geq \operatorname{rank}(xy) \geq \operatorname{rank}(xyz) = \operatorname{rank}(y),$$

it follows that $\operatorname{rank}(xy) = \operatorname{rank}(y) = t$, and also $\operatorname{im}(y) = \operatorname{im}(xy)$. Since z is injective on $\operatorname{im}(xy)$, there are

$$\sum_{B \in \mathcal{B}(r,t)} \prod_{i \in B} |A_i|$$

choices for $\operatorname{im}(y) = \operatorname{im}(xy)$. That is, $\operatorname{im}(xy)$ contains at most one point from any kernel class of z. Since $(\operatorname{im}(x))y = \operatorname{im}(xy) = \operatorname{im}(y)$, there are $\left\{ {r \atop t} \right\} t!$ ways for y to map $\operatorname{im}(x)$ to $\operatorname{im}(y)$. Finally, $(\{1, \ldots, n\}) \setminus \operatorname{im}(x))y \subseteq \operatorname{im}(y)$, and so there t^{n-r} ways for y to map $(\{1, \ldots, n\}) \setminus \operatorname{im}(x))$ to $\operatorname{im}(y)$. Hence there are in total

$$\left\{ {r \atop t} \right\} t! t^{n-r} \sum_{B \in \mathcal{B}(r,t)} \prod_{i \in B} |A_i|$$

choices for y. Therefore

$$\mathbb{V}_n = \frac{1}{n^{3n}} \sum_{r=1}^{n} \sum_{k=1}^{r} \sum_{t=1}^{\min(r,k)} \sum_{\{A_1,\ldots,A_k\} \in \mathcal{A}(n,k)} \binom{n}{r} r! \binom{n}{k} k! \left\{ {n \atop r} \right\} \left\{ {r \atop t} \right\} t! t^{n-r} \sum_{B \in \mathcal{B}(r,t)} \prod_{i \in B} |A_i|,$$

since $|\mathcal{T}_n| = n^n$. It follows from Lemma 5.8 that

$$\mathbb{V}_n = \frac{1}{n^{3n}} \sum_{r=1}^{n} \sum_{k=1}^{r} \sum_{t=1}^{\min(r,k)} \left\{ {n \atop r} \right\} \binom{n}{r} r! \binom{n}{k} k! \left\{ {r \atop t} \right\} t! t^{n-r} \sum_{s=t}^{n+t-k} \binom{n}{s} \binom{n-s}{r-k} \binom{s}{t} t^{s-t},$$

as required. □

Finally, we prove the main two lemmas of this section. This is an analogue of Lemma 5.9.

Lemma 5.12 *There exist $c > 0$ such that*

$$\mathbb{V}_n \leq c n^5 \left(\max_{\substack{x,y,z \in (0,1] \\ z \leq \min(x,y)}} G(x,y,z) \right)^n,$$

where

$$G(x,y,z) = \frac{x^{1-x} y^{1-y} z^{1-2z}}{e^{x+y+z}(1-x)^{2(1-x)}(1-y)^{2(1-y)}(x-z)^{x-z}(y-z)^{y-z}}.$$

Proof We begin by applying the same strategy as in Lemma 5.9. That is we use Lemma 5.3 to give an upper bound without Stirling numbers of the second kind and then use Eqs. (3) and (4). It follows that

$$\begin{aligned} n^{3n} \mathbb{V}_n &= \sum_{r=1}^{n} \sum_{k=1}^{r} \sum_{t=1}^{\min(r,k)} \left\{ {n \atop r} \right\} \binom{n}{r} r! \binom{n}{k} k! \left\{ {r \atop t} \right\} t! t^{n-r} \sum_{s=t}^{n+t-k} \binom{n}{s} \binom{n-s}{r-k} \binom{s}{t} t^{s-t} \\ &\leq \sum_{r=1}^{n} \sum_{k=1}^{n} \sum_{t=1}^{\min(r,k)} \left(\frac{n!}{(n-r)!} \right)^2 \left(\frac{n!}{(n-k)!} \right)^2 \frac{t^{n-t} r^{n-r} k^{n-k}}{(r-t)!(k-t)! t!} \end{aligned}$$

Replacing the sums with n times their maximal value we obtain after some algebraic manipulation

$$n^{3n} \mathbb{V}_n \leq n^3 \max_{\substack{1 \leq r \leq n \\ 1 \leq k \leq n \\ 1 \leq t \leq \min(r,k)}} \left\{ \frac{(n!)^4 k^{n-k} r^{n-r} t^{n-t}}{((n-r)!(n-k)!)^2 (r-t)!(k-t)! t!} \right\}$$

Using Stirling's approximation $\mathbb{V}_n$ can be bounded by

$$n^{3n}\mathbb{V}_n \leq cn^3 \max_{\substack{1\leq r\leq n\\ 1\leq k\leq n\\ 1\leq t\leq \min(r,k)}} \left\{ \frac{n^{4n+2}k^{n-k}r^{n-r}t^{n-2t}}{(n-r)^{2(n-r)}(n-k)^{2(n-k)}(r-t)^{r-t}(k-t)^{k-t}e^{r+k+t}} \right\}$$

for some $c > 0$. Let $x = r/n$, $y = k/n$, and $z = t/n$. The above equation can be rearranged to obtain

$$\mathbb{V}_n \leq cn^5 \left(\max_{\substack{x,y,z\in(0,1]\\ z\leq \min(x,y)}} \left\{ \frac{x^{1-x}y^{1-y}z^{1-2z}}{e^{x+y+z}(1-x)^{2(1-x)}(1-y)^{2(1-y)}(x-z)^{x-z}(y-z)^{y-z}} \right\} \right)^n.$$

Hence

$$\mathbb{V}_n \leq cn^5 \left(\max_{\substack{x,y,z\in(0,1]\\ z\leq \min(x,y)}} G(x,y,z) \right)^n,$$

as required. □

By inspection we see that G is continuous and bounded on

$$X = \left\{ (x,y,z) \in \mathbb{R}^3 \mid 0 < x, y < 1 \text{ and } 0 < z < \min(x,y) \right\}.$$

We can further extend the definition of G to the closure $\overline{X}$. It remains to find the maximum of G, which we do in the last lemma of this section.

Lemma 5.13 *There exists $r \in (0,1)$ such that $G(x,y,z) \leq r$ for all $x, y, z \in [0,1]$ such that $z \leq \min(x,y)$.*

Proof First we establish the value of G on the boundary $\overline{X}\setminus X$. Clearly for either $x = 0$, $y = 0$ or $z = 0$ we have $G(x,y,z) = 0$. If $x = 1$

$$G(1,y,z) = \frac{y^{1-y}z^{1-2z}}{e^{1+y+z}(1-y)^{2(1-y)}(1-z)^{1-z}(y-z)^{y-z}}.$$

By considering the derivative of x^{-x}, we can show that $x^{-x} \leq e^{e^{-1}}$ for all $x \in [0,1]$. Hence

$$G(1,y,z) = e^{7e^{-1}-1} \cdot \frac{yz}{e^{y+z}}.$$

Also note that $x \to xe^{-x}$ is increasing on $[0,1]$, and so $xe^{-x} \leq e^{-1}$. Therefore

$$G(1,y,z) \leq e^{7e^{-1}-3} \leq 0.7.$$

By symmetry this also holds for $y = 1$.

Let $x = z$. Then

$$G(x, y, x) = \frac{x^{2-3x} y^{1-y}}{e^{2x+y}(1-x)^{2(1-x)}(1-y)^{2(1-y)}(y-x)^{y-x}}.$$

Similarly,

$$G(x, y, x) \leq e^{3e^{-1}} \left(\frac{x^{1-x}}{e^x (1-x)^{1-x}} \right)^2 \cdot \frac{y^{1-y}}{e^y (1-y)^{1-y}}.$$

By considering the derivatives of $x \to \frac{x^{1-x}}{e^x (1-x)^{1-x}}$, we can show that the function has a unique maximum at $x_0 = \frac{1}{1+\Omega}$. Hence after some algebraic manipulations we get

$$\frac{x_0^{1-x_0}}{e^{x_0}(1-x_0)^{1-x_0}} = e^{-\frac{1}{1+\Omega}} \cdot \Omega^{-\frac{\Omega}{1+\Omega}} = e^{\Omega - 1},$$

and so

$$G(x, y, x) \leq e^{3(e^{-1}+\Omega-1)} < 1.$$

By symmetry the same holds for $y = z$.

The partial derivatives of G are as follows

$$\begin{aligned}
\frac{\partial G(x, y, z)}{\partial x} &= \left(\log \left(\frac{(1-x)^2}{x^2 - xz} \right) + \frac{1-x}{x} \right) G(x, y, z), \\
\frac{\partial G(x, y, z)}{\partial y} &= \left(\log \left(\frac{(1-y)^2}{y^2 - yz} \right) + \frac{1-y}{y} \right) G(x, y, z), \\
\frac{\partial G(x, y, z)}{\partial z} &= \left(\log \left(\frac{(x-z)(y-z)}{z^2} \right) + \frac{1-z}{z} \right) G(x, y, z).
\end{aligned}$$

Suppose that $(\alpha, \beta, \gamma) \in (0, 1]^3$ is a stationary point of $G(x, y, z)$. Note that $G(x, y, z) > 0$ if $x, y, z \neq 0$, and so

$$\log \left(\frac{(1-\alpha)^2}{\alpha^2 - \alpha\gamma} \right) + \frac{1-\alpha}{\alpha} = 0 \quad \text{and} \quad \log \left(\frac{(1-\beta)^2}{\beta^2 - \beta\gamma} \right) + \frac{1-\beta}{\beta} = 0$$

Hence

$$\gamma = \alpha - \frac{e^{\alpha^{-1}}(1-\alpha)^2}{e\alpha} = \beta - \frac{e^{\beta^{-1}}(1-\beta)^2}{e\beta} \tag{5}$$

However, the function $x \to x - \frac{e^{x^{-1}}(x-1)^2}{ex}$ is increasing and thus injective, implying that $\alpha = \beta$. Hence all stationary points of $G(x, y, z)$ are of the form (α, α, γ).

Substituting $\alpha = \beta$ into $\partial G(x, y, z)/\partial z = 0$ and rearranging we obtain that

$$\frac{1}{\gamma} + 2\log\left(\frac{\alpha - \gamma}{\gamma}\right) = 1.$$

Combining the above equation with (5) we get that α satisfies

$$\frac{e\alpha}{e\alpha^2 - e^{\alpha^{-1}}(1-\alpha)^2} + 2\log\left(\frac{e^{\alpha^{-1}}(1-\alpha)^2}{e\alpha^2 - e^{\alpha^{-1}}(1-\alpha)^2}\right) = 1. \tag{6}$$

It can be shown that the derivative of the function given by the left hand side of the above equation is

$$D(x) = -\frac{e\left(ex^3(x+3) - e^{x^{-1}}(1-x)^2(x^2+2x-1)\right)}{(1-x)\left(ex^2 - e^{x^{-1}}(1-x)^2\right)^2}.$$

Note that since the function in (5) is increasing, and so if $\alpha \leq 0.587$, then

$$\gamma \leq 0.587 - \frac{e^{0.587^{-1}}(1-0.587)^2}{e \cdot 0.587} < 0,$$

which contradicts $\gamma \in (0, 1]$. Hence $\alpha > 0.587$. It is easy to see that $x \to e^{x^{-1}}(1-x)^2$ is decreasing for $x \in (0, 1)$ and that $x \to x^2 + 2x - 1$ is increasing for $x \geq -1$. Thus

$$e^{x^{-1}}(1-x)^2(x^2+2x-1) \leq 2e^{0.587^{-1}}(1-0.587)^2 < 1.88$$

for $x \in [0.587, 1]$. On the other hand, $x \to ex^3(x+3)$ is increasing, and so for $x \geq 0.587$

$$ex^3(x+3) \geq e \cdot 0.587^3(0.587+3) > 1.97.$$

Therefore

$$ex^3(x+3) - e^{x^{-1}}(1-x)^2(x^2+2x-1) > 0$$

for $x > 0.587$, and thus $D(x) < 0$, implying that the left hand side of (6) is strictly decreasing. Hence there is a unique value α satisfying the (6). Moreover, we can see by inspection that $0.68152 < \alpha < 0.68153$. Since (5) is strictly increasing, it also follows that $0.44403 < \gamma < 0.44407$. Finally

$$\begin{aligned} G(\alpha, \alpha, \gamma) &= \frac{\alpha^{2(1-\alpha)}\gamma^{1-2\gamma}}{e^{2\alpha+\gamma}(1-\alpha)^{4(1-\alpha)}(\alpha-\gamma)^{2(\alpha-\gamma)}} \\ &\leq \frac{0.68153^{2(1-0.68153)}0.44407^{1-2\cdot 0.44407}}{(1-0.68153)^{4(1-0.68152)}(0.68152-0.44407)^{2(0.68153-0.44403)}e^{2\cdot 0.68152+0.44403}} \\ &< 0.999. \end{aligned}$$

Therefor there exists $r \in (0, 1)$ such that $G(x, y, z) \leq r$ for all $x, y, z \in [0, 1]$ such that $z \leq \min(x, y)$. □

The following is an immediate corollary of Lemmas 5.12 and 5.13, which concludes the proof of Theorems 1.2 and 1.3.

Corollary 5.14 *The probability* $\mathbb{V}_n$*, that* $\operatorname{rank}(xyz) = \operatorname{rank}(y)$ *where* $x, y, z \in \mathcal{T}_n$ *are chosen with uniform distribution, tends to 0 as* $n \to \infty$ *exponentially fast.*

Acknowledgements Open access funding provided by TU Wien (TUW).

References

1. Arvind, V., Torán, J.: The complexity of quasigroup isomorphism and the minimum generating set problem. In: Algorithms and Computation, Volume 4288 of Lecture Notes in Computer Science, pp. 233–242. Springer, Berlin (2006)
2. Cameron, P.J.: Dixon's theorem and random synchronization. Discrete Math. **313**(11), 1233–1236 (2013)
3. Corless, R.M., Gonnet, G.H., Hare, D.E.G., Jeffrey, D.J., Knuth, D.E.: On the Lambert *W* function. Adv. Comput. Math. **5**(4), 329–359 (1996)
4. Distler, A., Mitchell, J.D.: Smallsemi—GAP Package, Version 0.6.10 (2016)
5. East, J., Egri-Nagy, A., Mitchell, J.D., Péresse, Y.: Computing finite semigroups. J. Symb. Comput. **92**, 110–155 (2019)
6. Froidure, V., Pin, J.E.: Algorithms for computing finite semigroups. In: Cucker, F., Shub, M. (eds.) Foundations of Computational Mathematics. Springer, Berlin, Heidelberg (1997)
7. Gomes, G.M.S., Howie, J.M.: On the ranks of certain finite semigroups of transformations. Math. Proc. Camb. Philos. Soc. **101**(3), 395–403 (1987)
8. Gomes, G.M.S., Howie, J.M.: On the ranks of certain semigroups of order-preserving transformations. Semigroup Forum **45**(3), 272–282 (1992)
9. Gray, R.D.: The minimal number of generators of a finite semigroup. Semigroup Forum **89**(1), 135–154 (2014)
10. Howie, J.M.: Fundamentals of semigroup theory. In: LMS Monographs. Clarendon Press (1995)
11. Jonušas J, Mitchell, J.D., Pfeiffer, M.: Parallel algorithms for computing finite semigroups. Preprint (2017)
12. Mitchell, J.D. et al.: Semigroups—GAP Package, Version 2.8.0, (May 2016)
13. Papadimitriou, C.H., Yannakakis, M.: On limited nondeterminism and the complexity of the VC dimension. J. Comput. Syst. Sci. **53**(21), 161–170 (1996). **(Eighth Annual Structure in Complexity Theory Conference (San Diego, CA, 1993))**
14. The GAP Group: GAP—Groups, Algorithms, and Programming, Version 4.8.7 (2017)