comput. complex. **20** (2011), 21 – 50

1016-3328/11/010021-30
published online April 19, 2011
DOI 10.1007/s00037-011-0005-5

computational complexity

GENERIC COMPLEXITY OF FINITELY PRESENTED MONOIDS AND SEMIGROUPS

Mark Kambites

Abstract. We study the generic properties of finitely presented monoids and semigroups, and the generic-case complexity of decision problems for them. We show that for a finite alphabet A of size at least 2 and positive integers k and m, the generic A-generated k-relation monoid and semigroup (defined using any of several stratifications) satisfies the small overlap condition $C(m)$. It follows that the generic finitely presented monoid has a number of interesting semigroup-theoretic properties and, by a recent result of the author, admits a linear time solution to its word problem and a regular language of unique normal forms for its elements. Moreover, the uniform word problem for finitely presented monoids is generically solvable in polynomial time.

Keywords. Generic-case complexity, monoids, semigroups, finite presentations.

Subject classification. 68Q25, 20M05.

1. Introduction

Traditional complexity theory studies the time taken to solve a problem or execute an algorithm in the "worst case", but for many problems the "worst case" arises very infrequently. Probably the best known example is Dantzig's simplex method for linear programming (Dantzig 1951), which has exponential worst-case time complexity but in practice almost invariably terminates in linear time (see for example Klee & Minty 1972). Now over 60 years old, it remains the preferred choice for practical applications, even though there are now alternative algorithms with worst-case polynomial time complexity. Phenomena such as this motivated the development of *average-case complexity* (Gurevich 1991), which measures, roughly speaking, the *mean* difficulty of a problem across instances, with respect to some measure. Average-case complexity has proved extremely helpful for obtaining a theoretic understanding of the practical difficulty of problems, especially within the class NP of problems admitting non-deterministic worst-case polynomial time solution.

Average-case analysis can also be applied outside NP, but here it is less clear whether it serves the intended purpose. For most applications, what matters is not so much the *mean* difficulty of a problem of across instances, but rather the *typical* difficulty of instances encountered in practice. As is well known to statisticians, the mean value of a data set is not necessarily a guide to the typical values, since the former can be heavily skewed in one direction by a very small number of very extreme outliers. Likewise, the average-case complexity of a problem can be skewed upwards by a very small proportion of very difficult instances. Within NP, one at least has an exponential upper bound on worst cases; this imposes a limit on the "extremeness" of outlying instances and hence their ability to distort the mean. Outside NP, however, the distortion can be much more dramatic, with a tiny minority of extremely difficult instances potentially inflating the average-case complexity well beyond the complexity of the typical instance encountered in practice. This culminates in the extreme case of recursively unsolvable problems, whose average-case complexity is not defined at all, even though algorithms may exist to solve such problems efficiently for an overwhelming majority of cases (Gilman *et al.* 2007).

The aim of *generic-case complexity* is directly to analyse the complexity of *typical* problem instances, as distinct from the *average* difficulty of problem instances. Rather than introducing a *measure* on the instance space, the key idea is the *stratification* of an instance space (or indeed any other set) into an infinite sequence of finite subsets. A subset X of the space is called *generic* if the proportion of elements in each finite set which belong to X approaches 1 as one moves along the sequence. The generic complexity is (very roughly speaking) the minimum complexity attainable on a generic set. Compared with the average-case approach, the key feature is that no single instance (indeed no finite set of instances) makes any contribution at all to the generic properties of the space. Generic-case complexity was introduced by group theorists (Kapovich *et al.* 2003), investigating the large stock of hard algorithmic problems which occur in the study of finitely generated infinite groups. It has proved especially useful in view of recent interest in the use of non-commutative algebraic structures as a basis for cryptographic systems (see for example, Shpilrain & Zapata 2006), permitting for example a theoretical understanding of the success of the *length-based attack* (Ruinskiy *et al.* 2007) on the Shpilrain–Ushakov key establishment protocol based on the Thompson group (Shpilrain & Ushakov 2005).

The main aim of this paper is to study generic properties of finitely presented monoids and semigroups and hence to understand the generic-case complexity of uniform decision problems for monoids and semigroups. Our main results show that, with respect to a number of very natural stratifications, the generic[1] finite monoid presentation (over a given alphabet and with a given number of generators) satisfies *small overlap conditions* in the sense introduced by Remmers (1971, 1980) (see also Higgins 1992). Small overlap conditions are natural semigroup-theoretic analogues of the *small cancellation conditions* extensively used by combinatorial group theorists (see Lyndon & Schupp 1977), and so, our main result can be viewed as loosely analogous (although our objectives and hence our formalism are rather different) to the well-known fact, first asserted by Gromov (1987) and proved in detail by Ol′shanskiĭ (1992), that the generic finitely presented group is word hyperbolic.

These results immediately tell us a great deal about the algebraic structure of the generic finitely presented monoid. For example, we learn that it is $\mathscr{J}$*-trivial* (that is, every element

[1]For brevity, we use statements such as "the generic X has property Y" as shorthand for "there is a generic subset of the set of X's, every member of which has property Y". Of course the generic X truly "exists" only in the case that a single isomorphism type forms a generic subset of X's; in this case the isomorphism type has all the ascribed properties, so the terminology is unambiguous!

generates a distinct principal ideal) and hence torsion-free with no non-trivial subgroups. Even more important, by recent results of the author (Kambites 2009a), the uniform word problem for such presentations is solvable in (worst-case RAM) time linear in the word lengths and quadratic in the presentation size. Since it can be checked in (worst-case RAM) polynomial time whether a presentation satisfies a small overlap condition, it follows that the uniform word problem for finitely presented monoids is generically solvable in polynomial time (in the RAM model, linear in the word lengths and quadratic in the presentation size). All of these results apply equally to semigroups without identity elements.

As already remarked, generic-case complexity has been developed by combinatorial and geometric group theorists, and the literature is largely concerned with applications to advanced group theory; as a result, much of it is not readily accessible to non-algebraists. An additional objective of this article is to provide a relatively gentle exposition of generic sets, generic properties and generic-case complexity, in a form fully intelligible to the reader without a specialist algebraic background. Monoid presentations are combinatorially simpler objects than group presentations; the relatively straightforward combinatorial nature of many of our proofs should allow them to double as detailed worked examples to give the reader a feel for generic-case complexity.

In addition to this introduction, this article comprises four sections. Section 2 provides a gentle introduction to generic sets and generic-case complexity. In Section 3, we prove our main results about generic finitely presented monoids and semigroups with respect to certain stratifications. In Section 4, we prove some technical results regarding the relationships between different stratifications; these may be of some independent interest and are applied to show that our results about generic finitely presented monoids apply regardless of which of several natural stratifications are chosen. Finally, Section 5 explores the consequences of our characterisations of generic finitely presented monoids and semigroups, including the fact that the uniform word problems for finitely presented monoids and semigroups are generically solvable in time quadratic in the presentation lengths and linear in the word lengths.

2. Generic Properties and Generic-case Complexity

In this section, we provide a brief introduction to generic sets and generic complexity. A more comprehensive treatment can be found in Gilman *et al.* (2007). Our aim is to make the paper accessible to as wide an audience as possible, and so, we endeavour to keep mathematical prerequisites to a minimum. However, we cannot avoid assuming some elementary familiarity with the theory of sets and sequences.

Let S be a countably infinite set. A *stratification* of S is an infinite sequence $S_1, S_2, \ldots, S_n, \ldots$ of finite, non-empty subsets of S whose union is S. The computationally orientated reader may like to bear in mind the example where S is the instance space for some problem, and S_n is the set of instances of size n for some suitable notion of size; however, we caution that in general the subsets S_n need not be disjoint. We call the stratification *spherical* if the sets S_n are pairwise disjoint ($S_i \cap S_j = \emptyset$ for all $i \neq j$) and at the other extreme *ascending* if they form an ascending sequence under containment ($S_i \subseteq S_j$ for all $i < j$).

Now let X be a subset of S. We say that X is *generic* (with respect to the given stratification) if

$$\lim_{n\to\infty} \frac{|X \cap S_n|}{|S_n|} \tag{2.1}$$

is defined and equal to 1. The subset X is called *negligible* if $S \backslash X$ is generic, or equivalently, if above limit is defined and equal to 0. Intuitively, X is generic if the probability that an instance of size n, chosen uniformly at random, lies in X can be made arbitrarily high by choosing large enough n.

Note that, for any given set X, the limit in (2.1) may not be defined, and indeed for almost any stratification, it is easy to construct a set X for which it is not. The function

$$X \mapsto \lim_{n\to\infty} \frac{|X \cap S_n|}{|S_n|}$$

is a finitely additive probability measure defined on those subsets of X for which the limit converges, but it is typically *not* a measure in the usual sense, since it lacks countable additivity. This fact is no accident: a countably additive probability measure on a countable set clearly cannot assign measure 0 to all the singletons, but as we noted in the introduction, a key feature of the generic approach is that single instances are regarded as negligible. Nonetheless, the intuition that the generic sets are those of "full measure" can be helpful, and they satisfy many of the elementary properties of such sets. In particular, the reader can easily verify that if X is generic and $X \subseteq Y$ then Y is generic, while if X and Y are both generic then $X \cap Y$ is generic. Obvious dual statements hold for negligible sets.

Notice that, in our initial definition of generic sets, we have placed no requirements on the *rate* of convergence of the limit in (2.1). Genericity is an asymptotic property, and if convergence is very slow, then the asymptotic behaviour may not be reflected in "practical sized" instances. We call a set X *superpolynomially* generic/negligible if the appropriate limit converges faster than $1/n^p$ for *every* $p \in \mathbb{N}$, and *exponentially* generic/negligible if it converges faster than p^n for some $p \in (0, 1]$. (In the literature, some authors use the term "strongly generic" for what we have called exponentially generic sets, while some use "strongly generic" to mean superpolynomially generic and "supergeneric" to mean exponentially generic. To avoid confusion, we shall avoid these terms in favour of less concise but more descriptive ones.)

We now turn our attention to the application of generic sets in computational complexity. This requires us to consider explicitly not just abstract algorithmic problems, but also stratifications of instance spaces. We define a *stratified problem* to be an algorithmic decision problem equipped with a stratification on its instance space. (We shall restrict our attention here to decision problems, but analogous definitions can be made for more general computational problems.)

Of course traditional complexity theory is implicitly concerned with stratified problems: to study the asymptotic complexity of a problem, one requires a notion of the *size* of each member of the instance space S. As we have already remarked, this automatically induces a stratification given by setting S_n to be the set of all instances of size n. We call this the *input size stratification* for the problem. However, the dependence on stratification is much tighter in generic complexity theory than it is in traditional complexity theory—many authors discussing traditional complexity of algorithmic problems prefer to avoid detailed discussion of data encoding and hence of exact instance sizes; this is entirely reasonable since traditional complexity classes are largely insensitive to minor encoding issues. But for generic-case complexity, these issues can make a very big difference.

Note also that, while the input size stratification is a natural, canonical one to associate to any algorithmic problem, it is only one of many possible stratifications and may not be the appropriate one for any given application. The ideal is rather to find a stratification that reflects the empirical distribution of problem instances, that is, the frequency with which they arise in practice in a particular application, and there is often no reason to suppose that this is strongly correlated with size.

Now let **C** be any class of decision problems (typically a complexity class of some kind). We say that a stratified problem $\mathscr{P}$ is *generically in* **C** if there exists a generic subset Y of the instance space such that

(i) the membership problem for Y lies in **C**; and

(ii) the problem $\mathscr{P}$ restricted to Y lies in **C**.

Intuitively, a stratified decision problem is generically in **C** if the decision problem admits a partial algorithm (that is, an algorithm which outputs “yes”, “no” or “don’t know”, and which in the former two cases is always correct) in **C**, such that the probability of a “don’t know” is negligible. We write **GenC** for the class of all stratified problems generically in **C**.

Obvious examples are the class **GenP** of *generically polynomial-time stratified problems* and **GenNP** of *generically non-deterministic polynomial-time stratified problems*. Another interesting example is the class **GenBPP**, which consists of stratified problems admitting a randomised polynomial-time algorithm with probability of error uniformly bounded away from 1/2 for every instance in some generic subset whose membership problem also lies in **BPP**.

3. Generic Monoid Presentations

In this section, we study the generic properties of finite monoid presentations. We begin with some basic definitions.

Recall that a *semigroup* is a set S equipped with an associative binary operation (that is, a rule for multiplication satisfying $(xy)z = x(yz)$ for all elements $x, y, z \in S$). A *monoid* is a semigroup containing an *identity element* (that is, an element $1 \in S$ satisfying $1x = x1 = x$ for all elements $x \in S$).

Let A be a finite alphabet (set of symbols). A *word* over A is a finite sequence of zero or more elements from A. The set of all words over A is denoted A^*; under the operation of *concatenation* it forms a monoid, called the *free monoid* on A. The length of a word $w \in A^*$ is denoted $|w|$. The unique *empty word* of length 0 is denoted ϵ; it forms the identity element of the monoid A^*. The set $A^+ \setminus \{\epsilon\}$ of non-empty words forms a subsemigroup of A^*, called the *free semigroup on* A.

A finite monoid presentation $\langle A \mid R \rangle$ consists of a finite alphabet A, together with a finite sequence $R \subseteq A^* \times A^*$ of ordered pairs of words.[2] We say that $u, v \in A^*$ are *one-step equivalent* if $u = axb$ and $v = ayb$ for some possibly empty words $a, b \in A^*$ and relation $(x, y) \in R$ or $(y, x) \in R$. We say that u and v are *equivalent* and write $u \equiv_R v$ or just $u \equiv v$, if there is a finite sequence of words beginning with u and ending with v, each term of which but the last is one-step equivalent to its successor. Equivalence is clearly an equivalence relation; in fact, it is the least equivalence relation containing R and compatible with the multiplication in A^*. The equivalence classes form a monoid with multiplication well defined by $[u]_\equiv [v]_\equiv = [uv]_\equiv$; this is called the *monoid presented* by the presentation.

The *word problem* for a (fixed) monoid presentation $\langle A \mid R \rangle$ is the algorithmic problem of, given as input two words $u, v \in A^*$, deciding whether $u \equiv_R v$. The *uniform word problem for finitely presented monoids* is the algorithmic problem of, given as input a monoid presentation $\langle A \mid R \rangle$ and two words $u, v \in A^*$, deciding

[2] The reader may think it more natural to consider a *set* of *unordered* pairs, but the definition we use simplifies the combinatorics in our analysis, and Theorem 4.10 below will show that it makes no difference to the end results.

whether $u \equiv_R v$. It is well known that there exist finite monoid presentations for which the word problem is undecidable and hence that the uniform word problem for finitely presented monoids is undecidable (Markov 1947; Post 1947). More generally, if $\mathscr{C}$ is some class of finite monoid presentations, then the uniform word problem for $\mathscr{C}$ monoids is the algorithmic problem of, given as input a monoid presentation $\langle A \mid R \rangle$ in $\mathscr{C}$ and two words $u, v \in A^*$, deciding whether $u \equiv_R v$.

Now suppose we have a fixed monoid presentation $\langle A \mid R \rangle$. A *relation word* is a word which appears as one side of a relation in R. A *piece* is a word which appears more than once as a factor in the relations, either as a factor of two different relation words or as a factor of the same relation word in two different (but possibly overlapping) places. Let $m \in \mathbb{N}$ be a positive integer. The presentation is said to *satisfy the small overlap condition* $C(m)$ if no relation word can be written as a product of *strictly fewer than* m pieces. Thus, $C(1)$ says that no relation word is empty; $C(2)$ says that no relation word is a factor of another. Small overlap conditions were introduced by Remmers (1971); an accessible introduction is given in the book of Higgins (1992), while more recent developments can be found in papers of the author (Kambites 2009a,b).

Definitions corresponding to all of those above can also be made for semigroups (without necessarily an identity element), by taking A^+ in place of A^* (in all places except the definition of one-step equivalence, where a and b must still be allowed to be empty).

Now fix an alphabet A. To study generic properties of k-relation presentations over A, we need a stratification on the (countable) set of all such. There are two obvious ways to define the size of a presentation and hence two natural stratifications of the A-generated k-relation presentations. First, one can take the size of the presentation to be the *sum length of the relation words*; this gives rise to the *sum length stratification* of presentations. Alternatively, one can define the size to be the *length of the longest relation word*; this leads to the *maximum length stratification*. Which choice is most appropriate depends on the application. For example, the sum length of a presentation is a good approximation to the space required to encode the presentation in the obvious way and hence for computational applications seems quite natural.

Intuitively, the sum length stratification lends greater weight to uneven distributions of the relation word lengths within a presentation; in particular, it results in a greater frequency of relatively short words, which makes it seem less likely that small overlap conditions will hold. Nevertheless, it transpires that our main results hold for both stratifications, which may be regarded as some evidence of their "robustness".

We emphasise that we are attempting here to stratify only the set of A-generated, k-relation semigroup presentations, where the alphabet A and set of relations k are fixed. There are, of course, also natural stratifications across *all* A-generated semigroup presentations, allowing the number of relations to vary. These typically lead to a high frequency of "short" relation words, which means that small overlap-type conditions do not hold generically. However, it seems likely that, for at least some natural stratifications of this type, the word problem remains generically solvable for other reasons. This interesting issue deserves further study.

We shall need a couple of elementary definitions from combinatorics. Let n and k be non-negative integers. Recall that a *composition of* n *into* k is an ordered k-tuple of positive integers which sum to n, while a *weak composition of* n *into* k is an ordered k-tuple of non-negative integers which sum to n.

Having fixed the alphabet A, a k-relation monoid presentation of sum length n is uniquely determined by its sequence of relation words; this in turn is uniquely determined by the concatenation in order of those words (a word in A^n) and the lengths of those words (a weak composition of n into $2k$, called the *shape* of the presentation). Thus, k-relation monoid presentations of sum relation length n are in a natural bijective correspondence with ordered pairs whose first component is a word of length n and whose second component is a weak composition of n into $2k$. This fact makes the sum length stratification particularly easy to analyse.

We shall need the following simple combinatorial lemma.

LEMMA 3.1. *Let A be a finite alphabet and c and p be positive integers. The number of distinct words of length c which admit factorisations as x_1vy_1 and as x_2vy_2 for some $x_1, x_2, y_1, y_2, v \in A^*$ with $|v| \geq p$ and $x_1 \neq x_2$ is bounded above by $c^2|A|^{c-p}$.*

Proof. Clearly if a word admits such factorisations at all, then it admits such a factorisation with $|v| = p$, so we need count only those words which admit such factorisations with $|v| = p$.

We claim that, once A, c and p are fixed, any such word is uniquely determined by x_1, y_1 and the length of x_2. Clearly, there are fewer than c^2 ways to choose the lengths of x_1 and x_2; doing so also fixes the length of y_1, since we must have

$$|x_1| + |v| + |y_1| = |x_1| + p + |y_1| = c.$$

Now, there are at most

$$|A|^{|x_1|+|y_1|} \;=\; |A|^{c-|v|} \;=\; |A|^{c-p}$$

ways to choose the words x_1 and y_1 with the given lengths, so proving the claim will suffice to prove the lemma.

Since x_1 and x_2 are distinct prefixes of the same word, their lengths cannot be equal. Suppose first that x_1 is longer than x_2 and write $v = v^{(1)} \dots v^{(|v|)}$ and $x_1 = x_1^{(1)} \dots x_1^{(|x_1|)}$ with each $v^{(i)}$ and $x_1^{(i)}$ in A. Then since $x_1 v y_1 = x_2 v y_2$, we have

$$v^{(i)} = \begin{cases} x_1^{(|x_2|+i)} & \text{for } 1 \leq i \leq |x_1| - |x_2| \\ v^{(i-|x_1|+|x_2|)} & \text{for } |x_1| - |x_2| < i \leq |v| \end{cases}$$

from which the claim follows.

If, on the other hand, x_1 is shorter than x_2, then we use the lengths of v and x_2 to deduce the length of y_2, whereupon a symmetric argument suffices to complete the proof. □

Proposition 3.2. *Let A be a finite alphabet, and n and r be positive integers, and fix a weak composition σ of n (into any number). Then, the proportion of presentations of shape σ which have a piece of length r or more is bounded above by $n^2|A|^{-r}$.*

Proof. The set of presentations over A of shape σ is in 1:1 correspondence with the set A^n via the map which takes each presentation to the concatenation, in the obvious order, of its relation words. If the presentation has a piece of length r or more then the corresponding word will feature that piece as a factor in at least two different places. By Lemma 3.1, it follows that the number of presentations with a piece of length r or more is bounded above by $n^2|A|^{n-r}$. The total number of such presentations is $|A|^n$, so the proportion of presentations with the desired property is bounded above by $n^2|A|^{-r}$ as required. □

COROLLARY 3.3. *Let A be a finite alphabet and k, n, m and K be positive integers with $m \geq 2$, and fix a weak composition σ of n into $2k$ such that no block has size less than K. Then, the proportion of presentations with alphabet A and shape σ which do not satisfy $C(m)$ is bounded above by*

$$\frac{n^2}{|A|^{K/(m-1)}}.$$

PROOF. If a presentation fails to satisfy $C(m)$, then some relation word can be written as a product of $m-1$ pieces. By assumption, this relation word must have length at least K, so one of the pieces must have length at least $K/(m-1)$. The result now follows immediately from Proposition 3.2. □

Before proving the first of our main theorems, we will need an elementary combinatorial result concerning weak compositions; this will serve to bound the proportion of presentations which feature a "short" relation word.

LEMMA 3.4. *Let k be a positive integer, and $f : \mathbb{N} \to \mathbb{N}$ be a function such that $f(n)/n$ tends to zero as n tends to infinity. Then, the proportion of weak compositions of n into k which feature a block of size $f(n)$ or less tends to zero as n tends to infinity.*

PROOF. It is well known and easy to prove (see for example Bóna 2002, Theorem 5.2) that the number of weak compositions of n into k is given by

$$C'_k(n) = \frac{(n+k-1)!}{n!(k-1)!}$$

Clearly, every partition of n into k featuring a block of size $f(n)$ or less can be obtained by refining a partition of n into $k-1$, with the extra decomposition in one of $k(f(n)+1)$ places. Thus, the number of such partitions is bounded above by

$$k(f(n)+1)C'_{k-1}(n) = k(f(n)+1)\frac{(n+k-2)!}{n!\,(k-2)!}$$

Hence, the proportion of such partitions amongst all weak compositions of n into k is bounded above by

$$\begin{aligned}\frac{k\ (f(n)+1)\ C'_{k-1}(n)}{C'_k(n)} &= \frac{k\ (f(n)+1)\ (n+k-2)!\ n!\ (k-1)!}{(n+k-1)!\ n!\ (k-2)!} \\ &= \frac{k(k-1)\ (f(n)+1)}{n+k-1} \\ &= k(k-1)\left(\frac{f(n)}{n+k-1} + \frac{1}{n+k-1}\right) \\ &\leq k(k-1)\left(\frac{f(n)}{n} + \frac{1}{n}\right)\end{aligned}$$

which clearly tends to zero as n tends to infinity. □

We are now ready to prove our main theorem for the sum relation length stratification.

THEOREM 3.5. *Let A be an alphabet of size at least 2, and k and m be positive integers. Then, the set of A-generated, k-relation monoid presentations which satisfy the condition $C(m)$ is generic with respect to the sum length stratification.*

PROOF. Since $C(2)$ implies $C(1)$, we may clearly assume without loss of generality that $m \geq 2$. We need to show that the proportion of A-generated k-relation monoid presentations of length n which fail to satisfy $C(m)$ tends to zero as n tends to infinity.

For each n, let P_n be the set of all weak compositions of n into k, let Q_n be the set of weak compositions of n into k featuring a block of size $3(m-1)\log_{|A|} n$ or less and let $R_n = P_n \backslash Q_n$. By an application of Lemma 3.4, with the function $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = 3(m-1)\log_{|A|} n$, we see that the proportion $|Q_n|/|P_n|$ tends to 0 as n tends to infinity.

For each weak composition σ, let x_σ be the proportion of presentations of shape σ which fail to satisfy $C(m)$. Note that by Corollary 3.3 we have

$$x_\sigma \leq \frac{n^2}{|A|^{K_\sigma/(m-1)}}$$

where K_σ denotes the smallest block size in σ. For each fixed n, there are clearly equally many ($|A|^n$ to be precise) presentations of each shape, so the proportion of presentations of length n failing to satisfy $C(m)$ is just the average over shapes σ of x_σ, that is:

$$\begin{aligned}
\frac{1}{|P_n|}\left(\sum_{\sigma\in P_n} x_\sigma\right) &= \frac{1}{|P_n|}\left(\sum_{\sigma\in Q_n} x_\sigma\right) + \frac{1}{|P_n|}\left(\sum_{\sigma\in R_n} x_\sigma\right) \\
&\leq \frac{1}{|P_n|}\left(\sum_{\sigma\in Q_n} 1\right) + \frac{1}{|P_n|}\left(\sum_{\sigma\in R_n} \frac{n^2}{|A|^{K_\sigma/(m-1)}}\right) \\
&= \frac{|Q_n|}{|P_n|} + \frac{1}{|P_n|}\left(\sum_{\sigma\in R_n} \frac{n^2}{|A|^{K_\sigma/(m-1)}}\right).
\end{aligned}$$

We have already observed that $|Q_n|/|P_n|$ tends to zero as n tends to infinity. Moreover, by the definition of R_n we have $K_\sigma > 3(m-1)\log_{|A|} n$ for all $\sigma \in R_n$ so that

$$\begin{aligned}
\frac{1}{|P_n|}\sum_{\sigma\in R_n} \frac{n^2}{|A|^{K_\sigma/(m-1)}} &\leq \frac{1}{|P_n|}\sum_{\sigma\in R_n} \frac{n^2}{|A|^{(3(m-1)\log_{|A|} n)/(m-1)}} \\
&= \frac{|R_n|}{|P_n|}\frac{n^2}{|A|^{(3(m-1)\log_{|A|} n)/(m-1)}} \\
&= \frac{|R_n|}{|P_n|}\frac{n^2}{|A|^{\log_{|A|}(n^3)}} \\
&\leq \frac{n^2}{n^3}
\end{aligned}$$

which tends to zero as required. □

An analysis of the proof shows, approximately speaking, that the proportion of presentations failing to satisfy any given small overlap condition goes to zero like $(\log_{|A|} n)/n$, which for practical purposes may be rather slow. The barrier to showing a faster convergence is the proportion of presentations featuring a "short" relation word ($|Q_n|/|P_n|$ in the notation of the proof); this proportion really does seem to decrease very slowly, suggesting that for the sum length stratification, fast convergence to small overlap conditions is not possible. To obtain statements about the "superpolynomially generic finitely presented monoid" or "exponentially generic finitely presented monoid" with respect to the sum length stratification, one would require arguments that take detailed account of the "short" relation words.

Our next task is to prove that an equivalent result holds for the maximum length stratification. We begin with an analogue of Lemma 3.4, which will show that the frequency of presentations featuring a "small" relation word is again negligible. This time, because the number of presentations of each shape of maximum length k is not fixed, we must reason directly with presentations rather than just shapes. Having taken account of this, the result is easier and, as one might expect given our remarks above on the relative frequency of "short" relation words in this stratification, stronger.

LEMMA 3.6. *Let A be an alphabet of size at least 2, k be a positive integer, and $f : \mathbb{N} \to \mathbb{N}$ be a function such that $n - f(n)$ tends to infinity as n tends to infinity. Then, the proportion of A-generated k-relation presentations of maximum relation word length n which feature a relation word of length $f(n)$ or less tends to zero as n tends to infinity. Moreover, if there exists a constant $p > 0$ such that $n - f(n) > pn$ for sufficiently large n then the given proportion tends to zero exponentially fast.*

PROOF. Let X_n be the set of all presentations over A of maximum relation length n, let Y_n be the presentations in X_n which have a relation word of length $f(n)$ or less and let $Z_n = X_n \backslash Y_n$. The quantity we seek is thus the limit as n tends to infinity of $|Y_n|/|X_n|$. Let $I = \{1, \ldots, 2k\}$ and define a map σ from $I \times X_n$ to the set of all k-relation presentations over A, which takes (i, P) to the presentation obtained from P by removing $n - f(n)$ characters from the end of the ith relation word, or replacing this relation word with the empty word if its length is less than $n - f(n)$.

We claim that under the map σ, every presentation in Y_n has at least $|A|^{n-f(n)}$ pre-images in $I \times X_n$. Indeed, if $Q \in Y_n$, then Q has some relation word (say the jth) of length less than $f(n)$, say length p. Now for each of $|A|^{n-f(n)}$ words $w \in A^{n-f(n)}$, we can obtain from Q a presentation $P_w \in X_n$ by appending w to the end of the jth relation word, and it is easily seen $\sigma(j, P_w) = Q$ for all such w.

Thus, we have $2k|X_n| = |I \times X_n| \geq |A|^{n-f(n)}|Y_n|$, and so

$$\frac{|Y_n|}{|X_n|} \leq \frac{2k}{|A|^{n-f(n)}}.$$

Since $n - f(n)$ tends to infinity with n, this clearly tends to zero. If moreover $p > 0$ is such that $n - f(n) \geq pn$ for n sufficiently large, then we have

$$\frac{|Y_n|}{|X_n|} \leq \frac{2k}{|A|^{pn}}$$

so that the given quantity tends to zero exponentially fast. □

Corollary 3.7. *Let A be an alphabet of size at least 2, k be a positive integer, and c a constant with $0 < c < 1$. Then, the proportion of A-generated, k-relation presentations of maximum relation word length n which feature a relation word of length cn tends to zero exponentially fast as n tends to infinity.*

Proof. Define $f : \mathbb{N} \to \mathbb{N}$ by $f(n) = cn$, and choose p with $0 < p < 1 - c$. Then, $n - f(n) = (1 - c)n > pn$ for all n, so the result follows from Lemma 3.6. □

We are now ready to prove our main result for the maximum length stratification.

Theorem 3.8. *Let A be an alphabet of size at least 2, and let k and m be positive integers. Then, the set of A-generated, k-relation monoid presentations which satisfy $C(m)$ is exponentially generic with respect to the maximum length stratification.*

Proof. The structure of the proof is essentially the same as that for Theorem 3.8, but it is slightly complicated by the fact that the number of presentations of each shape for a given maximum relation word n is not fixed. In addition, we must show that the rate of convergence is exponential. Once again, we assume without loss of generality that $m \geq 2$.

Let C_n be the total number of presentations over A of maximum relation word length n. Let P_n be the set of all weak compositions of any integer into $2k$ with largest block size n. Choose d with $0 < d < 1$ and let Q_n be the set of all shapes in P_n with a word of length dn or less. Let $R_n = P_n \setminus Q_n$. For each weak composition $\sigma \in P_n$, let c_σ be the total number of presentations of shape σ and let x_σ be the proportion of presentations of shape σ which fail to satisfy $C(m)$. For each shape σ, by Corollary 3.3 we have

$$x_\sigma \leq \frac{(n_\sigma)^2}{|A|^{K_\sigma/(m-1)}}$$

where n_σ is the total size of σ (that is, the sum of the block sizes of σ, or the *sum* relation word length of a presentation of shape σ), and K_σ is the smallest block size in σ. But σ has $2k$ blocks, none of which is larger than n, so we must have $n_\sigma \leq 2kn$ and hence

$$x_\sigma \leq \frac{(2kn)^2}{|A|^{K_\sigma/(m-1)}} = \frac{4\ k^2\ n^2}{|A|^{K_\sigma/(m-1)}}.$$

Now the proportion we seek is given by

$$\begin{aligned} \frac{1}{C_n}\left(\sum_{\sigma\in P_n} c_\sigma x_\sigma\right) &= \frac{1}{C_n}\left(\sum_{\sigma\in Q_n} c_\sigma x_\sigma\right) + \frac{1}{C_n}\left(\sum_{\sigma\in R_n} c_\sigma x_\sigma\right) \\ &\leq \frac{1}{C_n}\left(\sum_{\sigma\in Q_n} c_\sigma\right) + \frac{1}{C_n}\left(\sum_{\sigma\in R_n} c_\sigma \frac{4k^2n^2}{|A|^{K_\sigma/(m-1)}}\right). \end{aligned}$$

The first term in the last line is the proportion of presentations of maximum relation word length n which feature a relation word of length dn or less; by Corollary 3.7, this tends to zero exponentially fast. Considering now the second term, by the definition of R_n we have that $K_\sigma > dn$ for all $\sigma \in R_n$ so that

$$\begin{aligned}\frac{1}{C_n}\sum_{\sigma\in R_n} c_\sigma \frac{4k^2n^2}{|A|^{K_\sigma/(m-1)}} &\leq \frac{1}{C_n}\sum_{\sigma\in R_n} c_\sigma \frac{4k^2n^2}{|A|^{dn/(m-1)}} \\ &= \left(\frac{4k^2n^2}{|A|^{dn/(m-1)}}\right)\left(\frac{\sum_{\sigma\in R_n} c_\sigma}{C_n}\right) \\ &\leq \frac{4k^2n^2}{(|A|^{d/(m-1)})^n}\end{aligned}$$

which since $|A| \geq 2$ and $d > 0$ clearly tends to zero exponentially fast. □

4. Equivalence of Stratifications

It often happens that two stratifications (on the same set, or on related sets) are closely related, so that knowledge of the generic sets with respect to one yields corresponding information about the generic sets with respect to the other. In this section, we establish some technical conditions under which this holds and use this to extend many of our earlier results to additional natural stratifications.

First, we consider the relationship between spherical and ascending stratifications. So far, we have seen examples only of spherical stratifications of instance spaces, but to each such stratification is associated an equally natural ascending stratification, the sets in the latter being unions of the sets in the former. The following proposition, which was first observed by Gilman *et al.* (2007) to be an easy consequence of the Stolz–Cesaro Theorem, says that the generic sets are independent of which of these stratifications are used (see Gilman *et al.* 2007 for a more detailed explanation).

PROPOSITION 4.1. (*Gilman et al. 2007, Lemma 3.2*) *Let S_n be a spherical stratification of a set S. Define a new stratification on S by*

$$B_n = \bigcup_{j=1}^{n} S_j.$$

Then, any set $X \subseteq S$ is generic with respect to the stratification S_n if and only if it is generic with respect to the stratification B_n.

We shall need the following elementary proposition, which essentially says that the restriction of a stratification to a generic set preserves generic sets.

LEMMA 4.2. *Let X be a stratified set, and X' a generic subset of X. Then for any $P \subseteq X$, we have*

$$\lim_{n\to\infty} \frac{|P \cap X_n|}{|X_n|} = \lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n \cap X'|}$$

and one limit is defined if and only if the other is.

PROOF. First notice that, since X' is generic, we have

$$\lim_{n\to\infty} \frac{|P \cap X_n \cap (X \backslash X')|}{|X_n|} = \lim_{n\to\infty} \frac{|(X \backslash X') \cap X_n|}{|X_n|} = 0 \tag{4.3}$$

Now

$$\begin{aligned}\lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n \cap X'|} &= \lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n|} \frac{|X_n|}{|X_n \cap X'|} \\ &= \left(\lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n|}\right)\left(\lim_{n\to\infty} \frac{|X_n \cap X'|}{|X_n|}\right)^{-1}\end{aligned}$$

Now since X' is generic, the right-hand factor in the above expression is equal to 1. Using also Eq. (4.3) to tell us that the right-hand term in the next but one line is 0, we obtain

$$\begin{aligned}&\lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n \cap X'|} \\ &\quad = \left(\lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n|}\right) + \left(\lim_{n\to\infty} \frac{|P \cap X_n \cap (X \setminus X')|}{|X_n|}\right) \\ &\quad = \lim_{n\to\infty} \frac{|P \cap X_n \cap X'|}{|X_n|} + \frac{|P \cap X_n \cap (X \backslash X')}{|X_n|} \\ &\quad = \lim_{n\to\infty} \frac{|P \cap X_n|}{|X_n|}\end{aligned}$$

as required. □

Next, we introduce a very useful sufficient condition for a map between stratified sets to preserve generic sets. To do so, we need some terminology. Let X and Y be stratified sets, $X' \subseteq X$ and $Y' \subseteq Y$, and $f : X' \to Y'$ a map. Then, f is called *stratification-preserving* if for every $x \in X'$ and $n \in \mathbb{N}$ we have $x \in X_n$ if and only if $f(x) \in Y_n$. If $P \subseteq X$, then f is said to *respect* P if $f(P \cap X')$ and $f((X \setminus P) \cap X')$ are disjoint, that is, if whenever $x_1, x_2 \in X'$ are such that $f(x_1) = f(x_2)$ we have either $x_1, x_2 \in P$ or $x_1, x_2 \notin P$. Recall that the *fibre size* of f at a point $y \in Y'$ is the cardinality of the set of elements $x \in X'$ such that $f(x) = y$. The map f is called *bounded-to-one* if there is a finite upper bound on its fibre sizes.

LEMMA 4.4. *Let X and Y be stratified sets, $X' \subseteq X$ and $Y' \subseteq Y$ be generic subsets of X and Y, respectively, $d \in \mathbb{N}$ and $f : X' \to Y'$ a surjective, stratification-preserving map, such that for every positive integer n there exists a positive integer k_n such that the fibre sizes of f at points in $Y_n \cap Y'$ all lie between k_n and dk_n. Then for any set $P \subseteq X$, we have*

(i)

$$\frac{1}{d} \lim_{n\to\infty} \frac{|f(P \cap X') \cap Y_n|}{|Y_n|} \leq \lim_{n\to\infty} \frac{|P \cap X_n|}{|X_n|} \leq d \lim_{n\to\infty} \frac{|f(P \cap X') \cap Y_n|}{|Y_n|}$$

wherever both limits are defined;

(ii)

$$\frac{1}{d} \lim_{n\to\infty} \frac{|P \cap X_n|}{|X_n|} \leq \lim_{n\to\infty} \frac{|f(P \cap X') \cap Y_n|}{|Y_n|} \leq d \lim_{n\to\infty} \frac{|P \cap X_n|}{|X_n|}$$

wherever both limits are defined;

(iii) *P is negligible in X if and only if $f(P \cap X')$ is negligible in Y;*

(iv) *If P is generic in X then $f(P \cap X')$ is generic in Y;*

(v) *If* $d = 1$ *and* $f(P \cap X')$ *is generic in* Y *then* P *is generic in* X*; and*

(vi) *If* f *respects* P *and* $f(P \cap X')$ *is generic in* Y *then* P *is generic in* X*.*

Before proving Lemma 4.4, we emphasise that parts (i) and (ii) do *not* guarantee that one of the limits involved is defined exactly if the other is defined. If one of the sequences converges to some value c, then only in the case $c = 0$ can we be certain that the other will converge. If $c \neq 0$, then the other may fail to converge, although it will eventually be constrained to vary within the range $[d^{-1}c, dc]$. We now turn to proving Lemma 4.4.

PROOF. By the bounds on the fibre sizes of f, we clearly have

$$|f(P \cap X' \cap X_n)| \leq |P \cap X' \cap X_n| \leq d|f(P \cap X' \cap X_n)| \quad \text{and}$$
$$|f(X' \cap X_n)| \leq |X' \cap X_n| \leq d|f(X' \cap X_n)|$$

for all $n \in \mathbb{N}$. It follows from the fact that f is surjective and stratification-preserving that $f(X' \cap X_n) = Y' \cap Y_n$ and $f(P \cap X' \cap X_n) = f(P \cap X') \cap Y_n$, so the above inequalities become

$$|f(P \cap X') \cap Y_n| \leq |P \cap X' \cap X_n| \leq d|f(P \cap X') \cap Y_n| \quad \text{and}$$
$$|Y' \cap Y_n| \leq |X' \cap X_n| \leq d|Y' \cap Y_n|$$

respectively. Now combining these yields

$$\begin{aligned} \frac{1}{d}\frac{|f(P \cap X') \cap Y_n|}{|Y' \cap Y_n|} &\leq \frac{|P \cap X_n \cap X'|}{|X_n \cap X'|} \\ &\leq d\frac{|f(P \cap X') \cap Y_n|}{|Y' \cap Y_n|}. \end{aligned} \tag{4.5}$$

It follows also that

$$\begin{aligned} \frac{1}{d}\frac{|P \cap X_n \cap X'|}{|X_n \cap X'|} &\leq \frac{|f(P \cap X') \cap Y_n|}{|Y_n \cap Y'|} \\ &\leq d\frac{|P \cap X_n \cap X'|}{|X_n \cap X'|} \end{aligned} \tag{4.6}$$

where the left-hand [respectively, right-hand] inequality is obtained by dividing [multiplying] both sides of the right-hand [left-hand] inequality in (4.5) by d.

Now since X' and Y' are generic in X and Y, respectively, Lemma 4.2 gives

$$\lim_{n\to\infty} \frac{|P \cap X_n|}{|X_n|} = \lim_{n\to\infty} \frac{|P \cap X' \cap X_n|}{|X_n \cap X'|}$$

and

$$\begin{aligned} \lim_{n\to\infty} \frac{|f(P \cap X') \cap Y_n|}{|Y_n|} &= \lim_{n\to\infty} \frac{|f(P \cap X') \cap Y_n \cap Y'|}{|Y_n \cap Y'|} \\ &= \lim_{n\to\infty} \frac{|f(P \cap X') \cap Y_n|}{|Y_n \cap Y'|} \end{aligned}$$

where the second equality on the second line holds because $f(P \cap X') \subseteq Y'$. It is now clear that parts (i) and (ii) follow from (4.5) and (4.6), respectively.

If $f(P \cap X')$ is negligible in Y, then the left- and right-hand sides of (i) converge to 0, from which it follows that the middle expression converges to 0, and so P is negligible. Conversely, if P is negligible, then exactly the same argument applies with (ii) in place of (i) to show that $f(P \cap X')$ is negligible. This proves part (iii).

If P is generic in X, then $X \backslash P$ is negligible in X, so by part (iii), $f((X \backslash P) \cap X')$ is negligible in Y. But by surjectivity, we must have

$$Y' \backslash f(P \cap X') \subseteq f((X \backslash P) \cap X')$$

so that $Y' \backslash f(P \cap X')$ is negligible in Y. Since $Y \setminus Y'$ is negligible in Y and negligible sets are closed under union, it follows that

$$Y \backslash f(P \cap X') = (Y' \backslash f(P \cap X')) \cup (Y \setminus Y')$$

is negligible in Y, so that $f(P \cap X')$ is generic in Y as required to prove part (iv).

If $d = 1$ and $f(P \cap X')$ is generic in Y then it is immediate from part (i) that P is generic in X, so that part (v) holds.

Finally, suppose that f respects P and that $f(P \cap X')$ is generic in Y. Since f is surjective, we have

$$Y' = f(X') = f((X \backslash P) \cap X') \cup f(P \cap X').$$

Now since f respects P, we know that $f((X\backslash P)\cap X')$ and $f(P\cap X')$ are disjoint, and since Y' is generic in Y is follows that

$$f((X\backslash P) \cap X') = Y' \backslash f(P \cap X')$$

is negligible in Y. But now by part (iii), we deduce that $X\backslash P$ is negligible in X, and hence that P is generic in X, as required to prove part (vi). □

A particularly useful special case is the following immediate corollary.

Corollary 4.7. *Let X and Y be stratified sets, $X' \subseteq X$ and $Y' \subseteq Y$ generic subsets of X and Y, respectively, and $f : X' \to Y'$ a surjective, stratification-preserving, bounded-to-one map. Then for any $P \subseteq X$ such that f respects P, we have that P is generic [respectively, negligible] in X if and only if $f(P \cap X')$ is generic [respectively, negligible] in Y.*

Next, we apply Lemma 4.4 to show that generic properties of finitely presented semigroups are essentially governed by those of finitely presented monoids. Recall that if S is a semigroup then S^1 denotes the monoid with set of elements $S \cup \{1\}$ where 1 is a new symbol not in S, and multiplication defined by

$$st = \begin{cases} \text{the } S - \text{product } st & \text{if } s, t \in S; \\ s & \text{if } t = 1; \\ t & \text{if } s = 1. \end{cases}$$

Theorem 4.8. *Let $\mathscr{C}$ be a class of monoids, A a finite alphabet and k a positive integer. Then, the generic A-generated k-relation monoid (with respect to either the sum length stratification or the maximum length stratification) belongs to $\mathscr{C}$ if and only if the generic A-generated k-relation semigroup S (with respect to the corresponding stratification) is such that S^1 belongs to $\mathscr{C}$.*

Proof. Let X and Y be the sets of k-relation monoid and semigroup presentations, respectively, over A. Suppose X and Y are equipped with either the sum length or the maximum length stratification. Let P be the set of presentations in X such that the monoid presented lies in $\mathscr{C}$, and let Q be the set of presentations in Y such that the semigroup S presented is such that S^1 lies in $\mathscr{C}$.

Let $Y' = Y$ and let $X' = Y \subseteq X$ be the set of semigroup presentations viewed as a subset of the set of monoid presentations, that is, those monoid presentations in which no relation word is empty. By Lemma 3.4 (for the sum length stratification) or Lemma 3.6 (for the maximum length stratification), X' is generic in X, and obviously, $Y' = Y$ is generic in Y.

Define $f : X' = Y \to Y' = Y$ to be the identity function; then f preserves the sum length and maximum length stratifications. Letting $d = 1$ and $k_n = 1$ for all n, we see that the conditions of Lemma 4.4 are satisfied, so P is generic in X if and only if $f(P \cap X')$ is generic in Y. Since f is the identity function on X', a semigroup presentation $\mathscr{P}$ lies in $f(P \cap X')$ exactly if $\mathscr{P}$ interpreted as a monoid presentation lies in P. Since $\mathscr{P}$ has no empty relation words, it is easy to see that the monoid presented by $\mathscr{P}$ is isomorphic to S^1, where S is the semigroup presented by $\mathscr{P}$. Thus, $\mathscr{P} \in f(P \cap X')$ if and only if $S^1 \in \mathscr{C}$, that is, if and only if $\mathscr{P} \in Q$. Hence, $f(P \cap X') = Q$, and so P is generic in X if and only if Q is generic in Y, as required. □

COROLLARY 4.9. *For every integer* $m \geq 1$, *positive integer* k *and alphabet* A *of size at least* 2, *the generic* A*-generated* k*-relation semigroup* (*with respect to either the sum length stratification or the maximum length stratification*) *satisfies the small overlap condition* $C(m)$.

An *unordered monoid presentation* consists of a set A of generators and an (unordered) *set* R of relations, each of which is an *unordered* pair of words from A^*. Equivalence of words is defined exactly as for ordered presentations (see Section 3), as are the sum length and maximum length stratifications on the sets of A-generated presentations with some fixed number k of relations. There is an obvious map from the ordered to the unordered presentations over a given alphabet A, which simply "forgets" the ordering of the relations and the ordering of the pair of words in each relation, and discards any duplicate relations. Unordered semigroup presentations can of course be defined analogously.

THEOREM 4.10. *Let $\mathscr{C}$ be a class of monoids, A an alphabet and k a positive integer. Then, the generic* [*negligible*] *A-generated k-relation monoid* (*with respect to either the sum length stratification or the maximum length stratification*) *belongs to $\mathscr{C}$ if and only if the generic* [*respectively, negligible*] *a-generator k-relation unordered monoid* (*with respect to the corresponding stratification*) *belongs to $\mathscr{C}$. The corresponding statement for semigroups also holds.*

PROOF. We prove the result for monoids; that for semigroups can be proved in exactly the same way. Let X be the set of ordered k-relation monoid presentations over A and Y the set of unordered k-relation monoid presentations over A. Let $P \subseteq X$ and $Q \subseteq Y$ be the sets of presentations in X and Y, respectively, such that the monoid presented belongs to $\mathscr{C}$.

Let $X' \subseteq X$ be the set of ordered presentations which do not feature the same relation twice, or two relations of the form (u, v) and (v, u) for some distinct words u and v. We have seen that $C(2)$ presentations do not feature the same relation word twice, so X' certainly contains all the $C(2)$ presentations. It follows by Theorem 3.5 (for the sum relation length stratification) or Theorem 3.8 (for the maximum relation length stratification) that X' is generic in X. Let $Y' = Y$; then certainly Y' is generic in Y.

Define $f : X' \to Y' = Y$ to be the obvious map from ordered to unordered presentations which forgets the ordering of the relations and of the pair of words in each relation. By definition, an ordered presentation in X' has no duplicate relations (even up to reordering the pairs), so this really does define a map to Y. Since every unordered presentation can be written in some order, this map is surjective. And since f takes each ordered presentation to an unordered presentation of the same monoid, it is clear that f respects P and maps $P \cap X'$ onto Q. It is easily seen that f preserves both the sum length and the maximum length stratifications. Moreover, f clearly has fibre size bounded above by $k!\ 2^k$.

It follows that the conditions of Corollary 4.7 are satisfied, so that P is generic in X if and only if $f(P) = Q$ is generic in Y. □

We thus allow ourselves to speak of a generic finitely presented monoid or semigroup, without worrying about whether the presentation is defined to have a set or a sequence of relations.

5. Properties of Generic Finitely Presented Monoids and Semigroups

In this section, we explore some of the consequences of our results for generic finitely presented monoids and semigroups. Recall that a monoid or semigroup is called $\mathscr{J}$*-trivial* if distinct elements always generate distinct principal ideals.

PROPOSITION 5.1. *Any* $C(3)$ *semigroup or monoid is torsion free and* $\mathscr{J}$*-trivial.*

PROOF. Let S be a semigroup or monoid with a $C(3)$ presentation $\langle A \mid R \rangle$. It follows from a result of (Remmers 1971, Corollary 4.14) (see also Higgins 1992, Corollary 5.2.16) that only finitely many words over the alphabet A represent the same element of S.

Suppose first that S it is not $\mathscr{J}$-trivial, and choose $a, b \in S$ be distinct elements generating the same ideal. Then in particular, a is in the ideal generated by b, so we have $a = pbq$ for some $p, q \in S$. But also b is in the ideal generated by a, so that and $b = ras = rpbqs$ for some $r, s \in S$. Now choose words $\hat{b}, \hat{p}, \hat{q}, \hat{r}, \hat{s} \in A^*$ representing $b, p, q, r, s \in S$, respectively. Certainly at least one of $\hat{r}$ and $\hat{s}$ is non-empty, since otherwise we would have $r = s = 1$ so that $b = ras = a$. But now it is easily seen that $(\hat{r}\hat{p})^i\hat{b}(\hat{q}\hat{s})^i$ represents b for every $i > 0$, contradicting Remmers' result.

Similarly, suppose $a \in S$ is non-identity torsion element. Then, there is a non-empty word $\hat{a} \in A$ representing a. But now it is easy to see that infinitely many powers of $\hat{a}$ must represent the same element, again contradicting Remmers' result. □

Combining with our theorem we have the following.

Theorem 5.2. *Let A be an alphabet of size at least 2 and let k be a positive integer. Then, the monoid defined by the generic A-generated k-relation presentation (with respect to either the sum length stratification or the maximum length stratification) is non-trivial, torsion-free and $\mathscr{J}$-trivial. In particular, it is not a group, an inverse monoid or a regular monoid. The corresponding statements for semigroups also hold.*

Proof. By Theorem 3.5 (respectively, Theorem 3.8 for the other stratification) the generic A-generated k-relation presentation satisfies $C(3)$, and so by Proposition 5.1 the semigroup presented is torsion-free and $\mathscr{J}$-trivial. If it were trivial then every word over the alphabet would have to represent the identity, contradicting once more Remmers' result mentioned in the proof of the previous proposition. □

By a recent result of the author (Kambites 2009a, Theorem 2), the uniform word problem for $C(4)$ semigroups is solvable in time linear in the word lengths and polynomial in the presentation size. Hence, we obtain

Theorem 5.3. *Let A be an alphabet of size at least 2 and let k be a positive integer. Then, the generic A-generated k-relation presentation (with respect to either the sum length stratification or the maximum length stratification) has word problem solvable in linear time. The corresponding statement for semigroups also holds.*

Since there is also an algorithm to decide, in (worst-case) polynomial time whether a given presentation satisfies the condition $C(4)$ (Kambites 2009a, Corollary 5) we also obtain

Theorem 5.4. *Let A be an alphabet of size at least 2 and k be a positive integer. Then, the uniform word problem for A-generated, k-relation monoid presentations is in* **GenP**. *The corresponding statement for semigroups also holds.*

Further work of the author (Kambites 2009b) has established a number of automata-theoretic properties of monoids which admit finite presentations satisfying the condition $C(4)$. It follows from Theorem 3.5 and Theorem 3.8 that the "generic" monoid and semigroup will enjoy all these properties. The following theorem summarises these properties; for brevity, we omit definitions of terms which can be found in Kambites (2009b).

THEOREM 5.5. *Let A be an alphabet of size at least 2 and let k be a positive integer. Then, the monoid defined by the generic A-generated k-relation presentation (with respect to either the sum length stratification or the maximum length stratification) is rational in the sense of Sakarovitch (1987), asynchronous automatic and word hyperbolic in the sense of Duncan & Gilman (2004). It also has the property that its rational subsets form a boolean algebra, coincide with its recognisable subsets and have uniformly decidable membership problem.*

Acknowledgements

This research was supported by an RCUK Academic Fellowship. The author would like to thank A. V. Borovik and V. N. Remeslennikov for their many suggestions; he also thanks the organisers and participants of the *AIM Workshop on Generic Complexity*, held in Palo Alto in August 2007, where he had many helpful conversations, and the American Institute of Mathematics for funding his attendance there.

References

M. Bóna (2002). *A walk through combinatorics*. World Scientific Publishing Co. Inc., River Edge, NJ. ISBN 981-02-4900-4, xviii+406.

G. B. Dantzig (1951). Maximization of a linear function of variables subject to linear inequalities. In *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13, 339–347. John Wiley & Sons Inc., New York, N.Y.

A. Duncan & R. H. Gilman (2004). Word hyperbolic semigroups. *Math. Proc. Cambridge Philos. Soc.* **136**(3), 513–524. ISSN 0305-0041.

R. Gilman, A. G. Miasnikov, A. D. Myasnikov & A. Ushakov (2007). Report on Generic Case Complexity. *Herald of Omsk University* 103–110. Available online at http://www.acc.stevens.edu/Files/GC/gc_survey.pdf.

M. Gromov (1987). Hyperbolic groups. In *Essays in Group Theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, 75–263. Springer, New York.

Y. Gurevich (1991). Average case complexity. In *Automata, languages and programming (Madrid, 1991)*, volume 510 of *Lecture Notes in Comput. Sci.*, 615–628. Springer, Berlin.

P. M. Higgins (1992). *Techniques of semigroup theory*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York. ISBN 0-19-853577-5, x+258. With a foreword by G. B. Preston.

M. Kambites (2009a). Small overlap monoids I: the word problem. *J. Algebra* **321**, 2187–2205.

M. Kambites (2009b). Small overlap monoids II: automatic structures and normal forms. *J. Algebra* **321**, 2302–2316.

I. Kapovich, A. Myasnikov, P. Schupp & V. Shpilrain (2003). Generic-case complexity, decision problems in group theory, and random walks. *J. Algebra* **264**(2), 665–694. ISSN 0021-8693.

V. Klee & G. J. Minty (1972). How good is the simplex algorithm? In *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)*, 159–175. Academic Press, New York.

R. C. Lyndon & P. E. Schupp (1977). *Combinatorial Group Theory*. Springer-Verlag.

A. Markov (1947). On the impossibility of certain algorithms in the theory of associative systems. *C. R. (Doklady) Acad. Sci. URSS (N.S.)* **55**, 583–586.

A. Yu. Ol′shanskiĭ (1992). Almost every group is hyperbolic. *Internat. J. Algebra Comput.* **2**(1), 1–17. ISSN 0218-1967.

E. L. Post (1947). Recursive unsolvability of a problem of Thue. *J. Symbolic Logic* **12**, 1–11. ISSN 0022-4812.

J. H. Remmers (1971). *Some algorithmic problems for semigroups: a geometric approach.* Ph.D. thesis, University of Michigan.

J. H. Remmers (1980). On the geometry of semigroup presentations. *Adv. in Math.* **36**(3), 283–296. ISSN 0001-8708.

D. Ruinskiy, A. Shamir & B. Tsaban (2007). Length-based cryptanalysis: the case of Thompson's group. *J. Math. Cryptol.* **1**, 359–372.

J. Sakarovitch (1987). Easy Multiplications I. The Realm of Kleene's Theorem. *Inform. and Comput.* **74**, 173–197.

V. Shpilrain & A. Ushakov (2005). Thompson's group and public key cryptography. *Lecture Notes in Computer Science* **3531**.

V. Shpilrain & G. Zapata (2006). Combinatorial group theory and public key cryptography. *Appl. Algebra Engrg. Comm. Comput.* **17**(3–4), 291–302. ISSN 0938-1279.

Manuscript received 4 November 2008

Mark Kambites
School of Mathematics,
University of Manchester,
Manchester M13 9PL,
England, UK.
Mark.Kambites@manchester.ac.uk