Mathematika 60 (2014) 444–462 © University College London
doi:10.1112/S0025579313000296

CONTRACTIVE FAMILIES ON COMPACT SPACES

LUKA MILIĆEVIĆ

Abstract. A family $f_1, \dots, f_n$ of operators on a complete metric space X is called contractive if there exists $\lambda < 1$ such that for any x, y in X we have $d(f_i(x), f_i(y)) \leqslant \lambda d(x, y)$ for some i. Stein conjectured that for any contractive family there is some composition of the operators f_i that has a fixed point. Austin gave a counterexample to this, and asked whether Stein's conjecture is true if we restrict to compact spaces. Our aim in this paper is to show that, even for compact spaces, Stein's conjecture is false.

§1. *Introduction.* Let (X, d) be a (non-empty) complete metric space. Given n continuous functions $f_1, f_2, \dots, f_n : X \to X$, and a real number $\lambda \in (0, 1)$, we call $\{f_1, f_2, \dots, f_n\}$ a λ-*contractive* family if for every pair of points x, y in X there is i such that $d(f_i(x), f_i(y)) \leqslant \lambda d(x, y)$. Further, we say that $\{f_1, f_2, \dots, f_n\}$ is a *contractive* family if it is a λ-contractive family for some $\lambda \in (0, 1)$. In particular, when f is a function on X and $\{f\}$ is a contractive family we say that f is a *contraction*. Recall the well-known theorem of Banach [2] which says that any contraction on a complete metric space has a unique fixed point.

In [5], Stein conjectured the following generalization of Banach's theorem. Let $\{f_1, f_2, \dots, f_n\}$ be a λ-contractive family on a complete metric space. Then some composition of $f_1, f_2, \dots, f_n$ (i.e. some word in $f_1, \dots, f_n$) has a fixed point.

In [1], Austin constructed a counterexample to this statement.

THEOREM 1.1 [1]. *There is a complete metric space (X, d) with a contractive family of operators $\{f, g\}$, such that no word in f, g has a fixed point.*

Furthermore, Austin asked if this is possible in a compact space.

QUESTION 1.2 [1]. Does every contractive family of operators on a compact space have a composition with a fixed point?

The purpose of this paper is to show that even with the additional assumption of compactness, there still need not be a fixed point.

THEOREM 1.3. *There is a compact metric space (X, d) with a contractive family of operators $\{f, g\}$, such that no word in f, g has a fixed point.*

We remark that our construction provides the counterexample for any given $\lambda \in (0, 1)$.

Received 27 April 2013, published online 14 May 2014.
MSC (2010): 47H09 (primary), 54E45 (secondary).

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

§2. *Outline of the construction of the counterexample.* Since our construction of the counterexample which proves Theorem 1.3 is fairly complicated, we provide some motivation for our steps. Suppose that we are in the setting of Theorem 1.3. Given any set S of points in our space, since X is compact, we know that S is bounded. Let D be its diameter. Then it is not hard to show that the diameter of one of $f(S)$ or $g(S)$ is at most $4\lambda D$. To see this, pick any point x in S, consider the set $S_f = \{y \in S : \lambda d(x, y) \geqslant d(f(x), f(y))\}$ and let $S_g = S \setminus S_f$. If $S_f = S$, we are done, so suppose that $S_g \neq \emptyset$. If we can find $y \in S_f$ such that $\lambda d(y, z) \geqslant d(f(y), f(z))$ for all $z \in S_g$, then, by looking at the distance from $f(y)$, the diameter of $f(S)$ does not exceed $4\lambda D$. On the other hand, if there is no such y, then each point of $g(S_f)$ has distance at most λD from some point in $g(S_g)$ which has distance at most λD from $g(x)$, so $g(S)$ has diameter at most $4\lambda D$.

This simple observation leads us to the idea that instead of considering distances between each pair of points in the space of interest, thinking about diameters of sets should be much more convenient in our problem. With this in mind, we develop the notion of "diameter spaces", which will play a key role throughout our construction. Due to their importance, we include the proper definition.

Let X be a non-empty set. Given a collection $\mathcal{D}$ of subsets of X, we call $(X, \mathcal{D})$ a *diametrizable space* provided the following conditions are met.

(1) Given $x, y \in X$ there is $U \in \mathcal{D}$ with $x, y \in U$.
(2) If $U, V \in \mathcal{D}$ and $U \cap V \neq \emptyset$ then $U \cup V \in \mathcal{D}$.

We refer to elements of X as points and elements of $\mathcal{D}$ as *diametrizable sets*. Further, if a function $\operatorname{diam} : \mathcal{D} \to \mathbb{R}_{\geqslant 0}$ is such that, given $U, V \in \mathcal{D}$ with a non-empty intersection,

$$\operatorname{diam}(U) + \operatorname{diam}(V) \geqslant \operatorname{diam}(U \cup V)$$

holds, then we call diam the *diameter* and $(X, \mathcal{D}, \operatorname{diam})$ a *diameter space*. We refer to this inequality as the *triangle inequality for the diameter spaces*.

In a very natural way, one can use a diameter space to induce a pseudometric on the underlying space, by simply finding the infimum of diameters of all the diametrizable sets containing any two given distinct points. Furthermore, by imposing suitable conditions on the diametrizable sets, one can get nice properties to hold for the pseudometric space.

In order to continue, we must first specify the underlying set. Hence, let us look for the space that should, in some vague sense, be the minimal counterexample. One of the possible ways to approach this issue is to fix a point x_0, and then look what other points can we obtain. It is not hard to see that completion of the set of all images that one can get by applying f and g to x_0 is itself a compact metric space, and that f, g form a λ-contractive map on this subspace of X as well. Now, starting from x_0, we must include all the points described, and we can actually map all these points that are obtained using f, g from x_0 to finite words over a two-letter alphabet (this

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

is a bijection if no two results of distinct compositions of f, g applied to x_0 coincide). Therefore, our construction will start from an underlying set X of all finite words over $\{a, b\}$, with obvious functions f, g each of which adds one of the characters to the beginning of the word given as input. Then, provided we have a metric on X, we will take its completion, and hope that the metric space that we get, along with f, g, satisfies all the properties of Theorem 1.3. This is where diameter spaces come into play. We will describe the properties of a collection of diametrizable sets $\mathcal{D}$ on X that guarantee that the completion of the induced pseudometric space, along with f, g (being the concatenation functions described), is a counterexample. Being a relatively long list, we refer the reader to Corollary 3.9 for an idea of what conditions we impose on $\mathcal{D}$, which will in fact be a sequence of subsets of X, denoted by $S_0, S_1, S_2, \ldots$.

§3. *Inducing a counterexample from a diameter space.*

3.1. *Diameter spaces and their connection with metric spaces.* In this subsection we show how one can obtain a pseudometric space from diameter spaces. The following proposition tells us how to induce a pseudometric on the underlying set.

PROPOSITION 3.1. *Let* $(X, \mathcal{D}, \mathrm{diam})$ *be a diameter space. Define a function* $d : X^2 \to \mathbb{R}$ *by* $d(x, y) = \inf \mathrm{diam}(U)$ *when* x *and* y *are distinct, where the infimum is taken over all diametrizable sets that contain points* x *and* y*, and* $d(x, y) = 0$ *otherwise. Then* d *is well defined and* (X, d) *is a pseudometric space.*

Proof. Firstly, suppose that we are given two distinct points x and y. Then S, the set of all values that the diameter of a diametrizable set containing x, y can take is non-empty, and bounded from below, so $\inf S$ exists, and d is well defined.

To prove that d is a pseudometric, we just need to show that the triangle inequality holds, since $d(x, x) = 0$ holds for all points x and d is symmetric by construction. Therefore let x, y, z be three points. If any of these are equal, we are done. Otherwise, given $\epsilon > 0$ we can find sets $U, V \in \mathcal{D}$ such that $x, y \in U$, $y, z \in V$, $d(x, y) \leqslant \mathrm{diam}\, U \leqslant d(x, y) + \epsilon/2$, $d(y, z) \leqslant \mathrm{diam}\, V \leqslant d(y, z) + \epsilon/2$. But U and V intersect hence their union forms a diametrizable set, and further $\mathrm{diam}(U \cup V) \leqslant \mathrm{diam}(U) + \mathrm{diam}(V) \leqslant d(x, y) + d(y, z) + \epsilon$, so $d(x, z) \leqslant d(x, y) + d(y, z) + \epsilon$. But this holds for any positive ϵ proving the triangle inequality, and therefore the proposition. □

The pseudometric that we constructed from the diameter space will be referred to as the *pseudometric induced by diam*, or when diam is clear from the context, as the *induced pseudometric*.

Remark that, given any metric space (X, d), we can construct a diameter space $(X, \mathcal{D}, \mathrm{diam})$ by taking diametrizable sets to be the finite subsets of X, and the diameter function diam to have the usual meaning, that is, for any $U \in \mathcal{D}$, we set $\mathrm{diam}(U)$ to be $\max(x, y)$ where maximum is taken over all pairs of points

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

in U. Then, the metric d coincides with the pseudometric induced by diam. Thus in this way, we can consider every metric space to be a diameter space at the same time.

In addition, we note that any finite measure space is a diameter space as well.

3.2. *Required properties of the diameter space.* In the previous subsection we saw how to obtain a pseudometric space from a diameter space. We will use this procedure to construct a counterexample to the conjecture, but as the proposition only guaranties that we get a pseudometric space, we need to add additional properties of a diameter space to ensure that we reach our goal. First of all, we work with specific underlying set X and the diameter function which are consistent with the nature (or more precisely the geometry) of the problem. As we are considering essentially compositions of functions applied to an element of the given metric space satisfying the assumptions of the conjecture, we will work with points of X being finite words over the alphabet consisting of two letters $\Sigma = \{a, b\}$, including the empty word. Before we proceed further, let us introduce some notation.

Notation. If u, v are two words of X, we write uv for the word obtained by writing first u then v. Suppose that u is a *prefix* or *initial segment* of v. If there is another word w with $v = uw$, write $u \leqslant v$. The *length* of a word u, denoted by $l(u)$, is the number of letters in u. Characters are considered to be words of length one at the same time as being characters. When S is a subset of X we write $uS = \{us : s \in S\}$ and $Su = \{su : s \in S\}$ when $u \in X$. Given a positive integer n and word u, write u^n for $uu \dots u$, where u appears n times. The empty word is denoted by $\emptyset$. Finally, we allow infinite words occasionally (although these are not elements of X), and write u^∞ for the infinite word obtained by writing consecutive copies of u infinitely many times.

We will take our functions $f, g : X \to X$ to be given by $f(u) = au$ and $g(u) = bu$ for all words $u \in X$. Then every word is actually the same as the corresponding composition of functions applied to the empty word. On the other hand, having in mind the contraction property of the family of the functions that we want to hold, we take $\mathcal{D}$ to be a sequence of subsets of X, namely $\mathcal{D} = \{S_0, S_1, S_2, \dots\}$, and set $\operatorname{diam}(S_k) = \lambda^k$ for some fixed $\lambda \in (0, 1)$. Now we just need to specify what needs to hold for $\mathcal{D}$ so that we get a counterexample.

Let us start by ensuring that $(X, \mathcal{D}, \operatorname{diam})$ is a diameter space. We accomplish this by requiring the following property.

A1 For any non-negative integers $i < j$, if $S_i \cap S_j \neq \emptyset$ then $S_j \subset S_i$ and $S_0 = X$.

Then we see that $(X, \mathcal{D})$ is a diametrizable space and that, given any two diametrizable sets S_i, S_j that intersect, we have $S_i \cup S_j = S_i$ or $S_i \cup S_j = S_j$ so the triangle inequality holds, thus this is indeed a diameter space.

Now consider (X, d), where d is the induced pseudometric by diam. To make d non-degenerate, we introduce another property.

A2 Each point belongs to only finitely many diametrizable sets.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

Since diam is always positive, and the infimum defining $d(x, y)$ for $x \neq y$ is actually a minimum taken over finitely many positive values, we get that d is non-degenerate, and thus a metric†.

As far as compactness is concerned, the fact that we can obtain a compact space from a totally bounded one by taking its completion is what motivates our next step. Hence, another condition is introduced.

A3 For any positive integer N, there are integers $i_1, i_2, \ldots, i_n$ greater than N such that $X \setminus (S_{i_1} \cup S_{i_2} \cup \cdots \cup S_{i_n})$ is finite.

PROPOSITION 3.2. *The completion* $(\overline{X}, d)$ *of the metric space induced by* $(X, \mathcal{D}, \text{diam})$ *is compact, provided* $\mathcal{D}$ *satisfies* ***A1–A3****.*

Proof. By the comments above, we just need to show that (X, d) is totally bounded. Let $\epsilon > 0$ be given, and choose N for which $\lambda^N < \epsilon/2$ holds. By **A3**, there are $i_1, i_2, \ldots, i_n > N$ for which the union of $S_{i_1}, S_{i_2}, \ldots, S_{i_n}$ covers all but finitely many points, denoted by $y_1, y_2, \ldots, y_m$. Then each S_{i_k} is contained in $B_{x_k}(\epsilon)$ for some $x_k \in S_{i_k}$ and $y_k \in B_{y_k}(\epsilon)$, so X is covered by finitely many open balls of radius ϵ, and so the metric space is totally bounded. □

Let us say that a Cauchy sequence $(x_n)_{n \geqslant 1}$ is proper if there is no N with $x_N = x_{N+1} = \cdots$. The three conditions described so far give us a nice characterization of the proper Cauchy sequences in $(\overline{X}, d)$, whose elements lie in X.

PROPOSITION 3.3. *Suppose that* $\mathcal{D}$ *satisfies properties* ***A1–A3****. Then a sequence of points in* X *is proper Cauchy with respect to the induced metric if and only if for any given positive integer* M *there is* $m > M$ *such that* S_m *contains all but finitely many points of the sequence.*

Proof. "Only if" direction. Let $(x_n)_{n \geqslant 1}$ be a proper Cauchy sequence in X and let a positive integer M be given. Take a positive $\epsilon < \lambda^M$. Then, as the given sequence is Cauchy, we have N such that $m, n > N$ implies $d(x_n, x_m) < \epsilon$.

Now fix any $m > N$ and let $I = \{i > M : \exists n > N, x_n, x_m \in S_i\}$. By the definition of d, we know that this set is non-empty, therefore has a minimal element i_0. If $n > N$ then $d(x_n, x_m) < \epsilon < \lambda^M$, so there is $j > M$ with x_n, x_m both belonging to S_j. But $x_m \in S_{i_0}$ so S_{i_0}, S_j intersect and by the choice of i_0 and property **A1** we have $x_n \in S_j \subset S_{i_0}$, so almost all points of the sequence are contained in S_{i_0}.

"If" direction. Given $\epsilon > 0$, take M such that $\lambda^M < \epsilon$. Then there is $m > M$ with S_m containing almost all points in the sequence, and the distance between two points in S_m is at most $\lambda^{M+1} < \epsilon$, so the sequence is Cauchy. If it were not proper, the point which is equal to almost all of its members would belong to infinitely many of the sets S_i, which is impossible. □

PROPOSITION 3.4. *Under the same assumptions as in the previous proposition, no proper Cauchy sequence in* $(\overline{X}, d)$ *converges to a point in* X*.*

† It is easy to see that this is in fact an ultrametric.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

Proof. Suppose that a point $x \in X$ is a limit of a proper Cauchy sequence (x_n) in $\overline{X}$. First of all, if $x_n \in \overline{X} \setminus X$, substitute x_n by a point $y \in X$ such that $d(x_n, y) < \frac{1}{2}d(x_n, x)$. The newly obtained sequence now lies in X, and is still proper Cauchy, with unchanged limit. As $x \in X$ belongs to S_i for only finitely many i, we have $B_x(\lambda^n) \cap X = \{x\}$, for sufficiently large n. This is, however, a contradiction, due to Proposition 3.3. □

LEMMA 3.5. *Let $(\overline{X}, d)$ be the metric space considered so far. Suppose that F is a function on X that preserves Cauchy sequences, that is, given a Cauchy sequence $(x_n)_{n\geqslant 1}$, the sequence of images $(F(x_n))_{n\geqslant 1}$ is Cauchy as well. Then, the extension of F to the completion of the space given by $F(x) = \lim_{n\to\infty} F(x_n)$, where (x_n) is any Cauchy sequence in X tending to x in $\overline{X} \setminus X$, is continuous.*

Proof. Firstly, we should prove that such extension of F is well defined. Take arbitrary $x \in \overline{X} \setminus X$; since $\overline{X}$ is the completion of x, there must be a Cauchy sequence in X whose limit is x. But the image of this sequence under F is Cauchy as well, so it has limit in $\overline{X}$, so we just need to show its uniqueness. Therefore, suppose that $(x_n)_{n\geqslant 1}, (y_n)_{n\geqslant 1}$ are two sequences in X tending to x. Merging these two sequences into $(t_n)_{n\geqslant 1}$, where $t_{2n-1} = x_n, t_{2n} = y_n, n \geqslant 1$, implies that $(t_n)_{n\geqslant 1}$ is Cauchy, hence $(F(t_n))_{n\geqslant 1}$ is also Cauchy, so $(F(x_n))_{n\geqslant 1}$ and $(F(y_n))_{n\geqslant 1}$ have the same limit, as required.

Secondly, we should prove that F is continuous in $\overline{X}$. Let $(x_n)_{n\geqslant 1}$ be a sequence tending to some $x \in \overline{X}$. If $x \in X$, then the sequence is eventually constant and equal to x, hence $(F(x_n))_{n\geqslant 1}$ trivially tends to $F(x)$. Otherwise, $x \notin X$, so consider a new sequence $(t_n)_{n\geqslant 1}$ given as follows. If $x_n \in X$, set $t_n = x_n$, and if this does not hold, there is a Cauchy sequence $(y_m)_{m\geqslant 1}$ in X whose limit is x_n. By assumption, $(F(y_m))_{m\geqslant 1}$ is Cauchy in X, and as we have shown previously, it tends to $F(x_n)$. Hence, for sufficiently large m, we have that $d(x_n, y_m), d(F(x_n), F(y_m)) < 1/n$, so set $t_n = y_m$. Thus, since $d(x_n, t_n) < 1/n$ for all n, we have that $(t_n)_{n\geqslant 1}$ is Cauchy in X and tends to x, so its image under F is a Cauchy sequence with limit $F(x)$, but $d(F(x_n), F(t_n)) < 1/n$ holds for all n, thus $\lim_{n\to\infty} F(x_n) = F(x)$, as required, implying the continuity of F. □

This lemma suggests the fourth property of diametrizable sets.

A4 If $i_1 < i_2 < \cdots$ are indices such that $S_{i_1} \supset S_{i_2} \supset \cdots$ then, given any N, we can find $n_a, n_b > N$ for which S_{n_a} contains all but finitely many elements of aS_{i_m} for some m, and S_{n_b} contains all but finitely many elements of bS_{i_p} for some p.

PROPOSITION 3.6. *If $\mathcal{D}$ satisfies **A1–A4**, then $f, g : \overline{X} \to \overline{X}$, defined before and then extended as in the previous lemma, are continuous with respect to the induced metric.*

Proof. We will show the claim for f; proof for g follows the same lines. We only need to show that f preserves Cauchy sequences in X; in fact it is sufficient to prove that if $(x_n)_{n\geqslant 1}$ is proper Cauchy, then $(f(x_n))_{n\geqslant 1}$ is Cauchy. Thus,

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

suppose that we are given a proper Cauchy sequence $(x_n)_{n\geqslant 1}$ in X, so there are indices $i_1 > 1, i_2 > 2, \dots$ (without loss of generality $i_1 < i_2 < \cdots$) such that S_{i_k} covers all but finitely many elements of the sequence for every k. Due to intersections and **A1** we have $S_{i_1} \supset S_{i_2} \supset \cdots$. Let $\epsilon > 0$ be given and take N such that $\lambda^N < \epsilon$. Further, by **A4** there exist $n > N$ and some m for which S_n contains all but finitely many elements of aS_{i_m}. Exploiting the fact that almost all elements of $(f(x_n))_{n\geqslant 1}$ are contained in aS_{i_m} yields that there is M such that $k, l > M$ implies $f(x_k), f(x_l) \in S_n$ hence $d(x_k, x_l) \leqslant \lambda^N < \epsilon$, as required. □

To make $\{f, g\}$ a contractive family, we define the following property.

A5 For any $i \in \{0, 1, 2, \dots\}$, there is $j > i$ such that $cS_i \subset S_j$ for some character $c \in \{a, b\}$.

PROPOSITION 3.7. *If $\mathcal{D}$ satisfies **A1–A5**, then f, g form a λ-contractive family in $(\overline{X}, d)$.*

Proof. Let us first consider $x, y \in X$. Taking the largest possible n for which $x, y \in S_n$, we have $m > n$ by **A5** such that without loss of generality $aS_n \subset S_m$, thus $d(f(x), f(y)) = d(ax, ay) \leqslant \lambda^m \leqslant \lambda\lambda^n = \lambda d(x, y)$, as desired. (If we had character b instead, we would get a contraction when applying g.)

In the general case, $x, y \in \overline{X}$, we can find sequences in X, $(x_n)_{n\geqslant 1}$ tending to x, and $(y_n)_{n\geqslant 1}$ tending to y (if one of these is already in X, then take a trivial sequence). Then, by the previous case, one of f, g contracts infinitely many pairs (x_n, y_n), f say. Let the indices of those pairs be $i_1 < i_2 < \cdots$. Then $\lambda d(x, y) = \lambda \lim_{n\to\infty} d(x_{i_n}, y_{i_n}) \geqslant \lim_{n\to\infty} d(f(x_{i_n}), f(y_{i_n})) = d(f(x), f(y))$, since f is continuous, as required. □

PROPOSITION 3.8. *Suppose that a function F, which is a word in f, g, has a fixed point. Let w be the non-empty word which corresponds to F, that is, $F(x) = wx$ for all x. Then there are infinitely many i such that there is u_i in S_i with wu_i also being a member of S_i.*

Proof. Suppose that F fixes u. Then $u \notin X$, so take $(x_n)_{n\geqslant 1}$ in X converging to u. Hence $(wx_n)_{n\geqslant 1}$ converges to $u = F(u)$ as well, so merging these two sequences together, we get a Cauchy sequence, and the result follows from Proposition 3.3. □

Therefore, let us introduce the final property of the diametrizable sets that we need to hold. The following corollary does this and sums up our work so far.

COROLLARY 3.9. *Let $\lambda \in (0, 1)$. Consider a collection $\mathcal{D}$ of diametrizable subsets of X that satisfies the following.*

A1 *The set S_0 is the whole of X and, given any non-negative integers $i < j$, $S_i \cap S_j \neq \emptyset$ implies $S_j \subset S_i$.*

A2 *Any point in X belongs to only finitely many diametrizable sets.*

A3 *For any positive integer N, there are integers $i_1, i_2, \dots, i_n$ greater than N such that $X \setminus (S_{i_1} \cup S_{i_2} \cup \cdots \cup S_{i_n})$ is finite.*

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

A4 *If $i_1 < i_2 < \cdots$ are indices such that $S_{i_1} \supset S_{i_2} \supset \cdots$ then, given any N, we can find $n_a, n_b > N$ for which S_{n_a} contains all but finitely many elements of aS_{i_m} for some m, and S_{n_b} contains all but finitely many elements of bS_{i_p} for some p.*

A5 *For any $i \in \{0, 1, 2, \dots\}$, there is $j > i$ such that $cS_i \subset S_j$ for some character $c \in \Sigma$.*

A6 *Given a non-empty word w in X there are only finitely many diametrizable sets S_i with $u_i \in S_i$ for which $wu_i \in S_i$.*

Then, with constructions described above, $(\overline{X}, d)$ is a compact metric space with continuous functions $f, g : \overline{X} \to \overline{X}$ that form a λ-contractive family, but no word in f, g has a fixed point.

§4. *Choosing the diametrizable sets.* Our main task now is to find a collection of diametrizable sets $\mathcal{D}$ which has the properties **A1**–**A6**. The most natural subsets of X to be examined are $W_w = wX$, where w is any word in X. We immediately observe that these, ordered by the length of w and then alphabetically (so that $S_0 = W_\emptyset$, $S_1 = W_a, \dots$), satisfy all the properties except for **A6**. Hence, we will use these as the pillar of our construction, albeit slightly modified to make **A6** hold. The issue with the sets described is that, given any non-empty word $w \in X$, we allow W_w to contain all the initial segments of w^∞. With this in mind, we say that a non-empty word w in X is forbidden if there is another finite word u such that w is a prefix of u^∞ and $l(w) > l(u)^2$. Otherwise, say that w is available. For example, $\emptyset, a, abab$ are available, while $ababa$ is forbidden.

PROPOSITION 4.1. *Given a word $w \in X$, either aw or bw is available.*

Proof. If $w = \emptyset$, aw is available. Suppose that $l(w) \geqslant 1$ and that the claim is false, so aw is an initial segment of w_1^∞ for some non-empty w_1, bw is an initial segment of w_2^∞ for some non-empty w_2 and $l(w) \geqslant l(w_1)^2, l(w_2)^2$. We can permute w_1 cyclically to v_1 so that $av_1^\infty = w_1^\infty$ holds, and we can correspondingly transform w_2 to v_2. Observe that the last character of v_1 is a and of v_2 is b. This way, w becomes a prefix of v_i^∞ for $i = 1, 2$. But $l(w) \geqslant l(v_1)^2$, $l(v_2)^2$, hence $l(w) \geqslant l(v_1)l(v_2)$ and so $v_1^{l(v_2)} = v_2^{l(v_1)}$, by comparing them as initial segments of w. However, this implies equality of the last characters of v_1 and v_2, which is impossible. The claim now follows. □

In the same spirit, we prove the following statement.

PROPOSITION 4.2. *Given a word $w \in X$, either wa or wb is available.*

Proof. Suppose, to the contrary, that both are forbidden. Let t_1, t_2 be words such that $wa \leqslant t_1^\infty$, $wb \leqslant t_2^\infty$, $l(wa) > l(t_1)^2$, $l(wb) > l(t_2)^2$. Observe that both $t_1^{l(t_2)}$ and $t_2^{l(t_1)}$ are prefixes of w, so, being of the same length, they coincide. But then $t_1^\infty = t_2^\infty$, which is a contradiction, as otherwise $wa \neq wb$ would be two prefixes of the same length of this infinite word. □

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

COROLLARY 4.3. *Let w be a non-empty word. Let u be an initial segment of w^∞, and suppose that $l(u) \geqslant l(w)^2$. Take a character s such that $v = us$ is not a prefix of w^∞. Then v is available.*

We say that a word w is minimal if it is non-empty and, given non-empty u for which $u^\infty = w^\infty$, we have $l(w) \leqslant l(u)$.

PROPOSITION 4.4. *A non-empty word w is not minimal if and only if there is a word u such that $w = u^k$, for some $k \geqslant 2$.*

Proof. If $w = u^k$ with $k > 1$ then $l(w) > l(u)$ and $w^\infty = u^\infty$, therefore w is not minimal. Suppose now that we have a non-empty w, for which there exists u such that $u^\infty = w^\infty$, but $l(u) < l(w)$. Write $d = \gcd(l(u), l(w))$, so $l(u) = qd, l(w) = pd$, for some positive integers p, q, in particular $p \geqslant 2$. Further, $u^p = w^q = v_1 v_2 \dots v_{pq}$, where $v_1, v_2, \dots, v_{pq}$ are of length d. Considering successive copies of u, we have that $v_{i+q} = v_i$ when $i \leqslant pq - q$, and similarly, by looking at copies of w, we have $v_{i+p} = v_i$ when $i \leqslant pq - p$. So $v_i = v_{p+i} = \dots = v_{p(q-1)+i}$ for $i \in [p]$ (where for a positive integer N, $[N]$ denotes the set $\{1, 2, \dots, N\}$). Observe that $v_q = v_{2q} = \dots = v_{pq}$ and, as p and q are coprime, $q, 2q, \dots, pq$ take all possible values modulo p, hence $v_1 = v_2 = \dots = v_{pq}$, allowing us to conclude $w = v_1^p$, $p > 1$. □

Having established these results about words, we are ready to choose the diametrizable subsets $\mathcal{D}$ of X. Consider the following subsets.

For all available words w, including $\emptyset$, we include W_w in $\mathcal{D}$ (we refer to these as *W-type sets*, so we say that S_i is of W-type if $S_i = W_w$, for some available w).

For all minimal words w, all integers p, r such that $p \in \{2^i : i \in \mathbb{N}_0\}$ and $0 \leqslant r \leqslant p - 1$, we set $A_{w,p,r} = \{u \in X : u \text{ initial segment of } w^\infty, l(u) \in \{r + ip : i \in \mathbb{N}\}, l(u) > l(w)^2\}$ (call these *A-type* sets).

Finally, for each minimal word w and $k \in \mathbb{N}$, we define $B_{w,k}$, which we refer to as *B-type* sets. For these, we need additional notation.

First of all, for each $k \in \mathbb{N}$ we define an infinite arithmetic progression I_k. We set $I_1 = \mathbb{N}$, $I_2 = I_1 \setminus \{\min I_1\} = \mathbb{N} \setminus \{1\}$. For each integer $m \geqslant 2$, if k is an integer such that $2^m - 1 \leqslant k \leqslant 2^m + 2^{m-1} - 2$, we set $I_k = \{s + i \cdot 2^{m-1} : i \in \mathbb{N}_0\}$, where $s = \min I_{(k-1)/2+2^{m-2}}$, when k is odd; and $I_k = \{s + i \cdot 2^{m-1} : i \in \mathbb{N}_0\}$, where $s = \min I_{(k-2)/2+2^{m-2}} + 2^{m-2}$, if k is even. On the other hand, if k is an integer such that $2^m + 2^{m-1} - 1 \leqslant k \leqslant 2^{m+1} - 2$, then we set $I_k = I_{k-2^{m-1}} \setminus \{\min I_{k-2^{m-1}}\}$.

Note that, given a minimal word w and $n \in \mathbb{N}$, by Corollary 4.3 we get a unique word w_n of the form $v_n s_n$, where $v_n \leqslant w^\infty$ and s_n is a character, w_n is available, and $l(w_n) = l(w)^2 + n$. Finally, we define $B_{w,k} = \bigcup_{i \in I_k} W_{w_i}$.

Thus, we set $\mathcal{D} = \{W_w : w \in X, w \text{ is available}\} \cup \{A_{w,p,r} : w \in X, p, r \in \mathbb{Z}, w \text{ is minimal}, p \in \{2^i : i \in \mathbb{N}_0\}, 0 \leqslant r \leqslant p - 1\} \cup \{B_{w,k} : w \in X, k \in \mathbb{N}, w \text{ is minimal}\}$.

To illustrate the definition of B-type sets, we list a few examples: $I_7 = \{4, 8, 12, 16, \dots\}$, $I_{13} = \{9, 13, 17, 21, \dots\}$, $B_{a,7} = W_{aaaab} \cup W_{aaaaaaaab} \cup W_{aaaaaaaaaaaab} \cup \dots$, $B_{ba,13} = W_{bababababaa} \cup W_{bababababababaa} \cup W_{babababababababababaa} \cup \dots$.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

Let us make a few easy remarks about $\mathcal{D}$. Fix a minimal word w. Then $B_{w,1} = W_{w^{l(w)}} \setminus (A_{w,1,0} \cup \{w^{l(w)}\})$. Also, if $m \geqslant 2$ is an integer, and k is an odd integer that satisfies $2^m - 1 \leqslant k \leqslant 2^m + 2^{m-1} - 2$, then $I_{(k-1)/2+2^{m-2}} = I_k \cup I_{k+1}$. Furthermore, if $I_{k_1} \cap I_{k_2} \neq \emptyset$ then $I_{k_1} \subset I_{k_2}$ or vice versa.

For any $U \in \mathcal{D}$, observe that there is a unique word in U of shortest length, which we will denote by $\sigma(U)$.

Let us now establish a few claims about the structure of $\mathcal{D}$, which will be exploited in the rest of the paper.

PROPOSITION 4.5. *If $U \neq V$ are two diametrizable sets and they intersect, then one is contained in the other. Furthermore, if they are not identical, then $U \triangle V$ is infinite.*

Proof. We deal with the possible types of U and V.

Case 1: U and V are W-type. Suppose that $U = W_{w_1}$, $V = W_{w_2}$; without loss of generality, $l(w_1) \leqslant l(w_2)$. If w_1 is not an initial segment of w_2 then U and V do not intersect, so we must have $W_{w_1} \supset W_{w_2}$, that is $U \supset V$. If $U \neq V$, then $w_1 \neq w_2$, and hence $W_{w_1 s} \subset (U \setminus V)$, where s is a character for which $w_1 s$ is not a prefix of w_2. We will use this case for showing the rest.

Case 2: U and V are A-type. Suppose that $U = A_{w_1,p_1,r_1}$, $V = A_{w_2,p_2,r_2}$, where w_1, w_2 are some minimal words, and p_1, p_2, r_1, r_2 are suitable integers. There is a word $w \in U \cap V$, so w is the initial segment of w_1^∞ and w_2^∞, while $l(w) \geqslant l(w_1)^2, l(w_2)^2$, from which we deduce that $w_1^{l(w_2)} = w_2^{l(w_1)}$, being the initial segment of w of length $l(w_1)l(w_2)$. Therefore $w_1^\infty = w_2^\infty$ and, due to minimality, $w_1 = w_2$. Now, due to the definition of A-type sets for a fixed minimal word, we get $U \subset V$ or vice versa and infinite symmetric difference.

Case 3: U is A-type, V is B-type. Let $w \in U \cap V$. This makes w a prefix of some t^∞, where t is minimal and $l(t)^2 < l(w)$, as $U = A_{t,p,r}$, p, r being suitable integers. Also $w \in W_v \subset V$, for some available v, so as v is an initial segment of w, hence t^∞, we have $l(v) \leqslant l(t)^2$, since v is available. But then $U \subset A_{t,1,0} \subset W_{t^{l(t)}} \subset W_v \subset V$, which proves the first claim.

On the other hand, U has no available words, but V has infinitely many, which gives the second part.

Case 4: U is A-type, V is W-type. The proof in this case is the same as in Case 3.

Case 5: U and V are B-type. Suppose that $U = B_{w_1,k_1}$, $V = B_{w_2,k_2}$, for some minimal words w_1, w_2 and positive integers k_1, k_2. Let $w \in U \cap V$, thus $w \in W_{v_1} \subset B_{w_1,k_1}$ and $w \in W_{v_2} \subset B_{w_2,k_2}$. By Case 1, without loss of generality, $W_{v_1} \subset W_{v_2}$ holds.

If $v_1 = v_2$, letting u_i be v_i without the last character gives us a prefix of w_i^∞, $i \in [2]$, and $l(w_1)^2, l(w_2)^2 \leqslant l(u_1)$, so as before $w_1^\infty = w_2^\infty$ and, due to minimality, $w_1 = w_2$. So $U \subset V$ or $V \subset U$, due to the construction of B-type sets for a fixed minimal word, and also the second part of the claim follows.

On the other hand, if $v_1 \neq v_2$, then $v_2 \leqslant u_1$, u_1 being v_1 after omitting the last character, as before. Further, u_1 is a prefix of w_1^∞, so v_2 is too, but v_2 is

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

available, hence $l(v_2) \leqslant l(w_1)^2$, hence $U \subset W_{v_2} \subset V$. For the second part of the claim, consider any other W-type set contained in V, distinct from W_{v_2}.

Case 6: U is B-type, V is W-type. Let $V = W_v$, for some available word v, and $w \in U \cap V$, so $w \in W_u \subset U$. If $W_v \subset W_u$, then $V \subset U$, so we are done, and second part follows as in the previous case. Otherwise, by Case 1, $W_u \subsetneq W_v$, so $v \neq u$ and $v \leqslant u$. We have $U = B_{t,k}$, for some minimal t and integer k, so $v \leqslant t^\infty$ and available so $l(v) \leqslant l(t)^2$, hence $U \subset W_v = V$. Also, $(W_v \setminus U) \supset W_{t^{l(t)}} \setminus B_{t,1} = A_{t,1} \cup \{t^{l(t)}\}$, which is infinite. □

PROPOSITION 4.6. *If U is any of the diametrizable sets, then there are unique V_1, V_2 (up to ordering) which are proper subsets of U in $\mathcal{D}$, such that $V_1 \cup V_2$ has almost all elements of U.*

Proof. Firstly, let us show the existence of such sets. Of course, we go through the possible types of set U. The only non-trivial case, which does not directly follow from construction of $\mathcal{D}$, is Case 3.

Case 1: U is of A-type. Suppose that $U = A_{w,p,r}$, for some minimal w and integers p, r. Then we can take $V_1 = A_{w,2p,r}$, $V_2 = A_{w,2p,p+r}$.

Case 2: U is of B-type. Suppose that $U = B_{w,k}$, for some minimal w and integer k. Then, by construction, either $B_{w,k} = B_{w,k_1} \cup B_{w,k_2}$, for two integers k_1, k_2, or $B_{w,k} = B_{w,k+1} \cup W_v$, for some available v, giving us the required sets V_1, V_2.

Case 3: U is of W-type. Suppose that $U = W_w$, for some available w. If both wa and wb are available, we can take $V_1 = W_{wa}$ and $V_2 = W_{wb}$.

Otherwise, ws is not available for some character s. Therefore, there is a word t such that ws is a prefix of t^∞, and $l(t)^2 < l(ws) = l(w) + 1$. But w is also a prefix of t^∞ and available, so $l(t)^2 \geqslant l(w)$, thus $l(w) = l(t)^2$, that is, $w = t^{l(t)}$.

The next thing to do is to establish the minimality of t. Let u be any word such that $l(u) \leqslant l(t)$ and $t^\infty = u^\infty$. Then, as before, ws is an initial segment of u^∞ and $l(u)^2 < l(ws) = l(w) + 1$, so by the same arguments we get $l(w) = l(u)^2$, thus $l(u) = l(t)$, as desired. But then $W_w = A_{t,1,0} \cup B_{t,1} \cup \{w\}$, proving the existence part of the claim.

Suppose now that U' is any set strictly contained in U. Then, by Proposition 4.5, $U \setminus U'$ is infinite, so U' cannot contain both V_1, V_2. Also, from the same proposition we get that V_1 and V_2 are disjoint, as otherwise one must contain the other and thus be equal to U. So, U' intersects at precisely one of V_1 and V_2, since all diametrizable sets are infinite. Without loss of generality, U' intersects V_1. If $V_1 \subsetneq U'$, then $U' \setminus V_1$ is infinite, and thus intersects V_2, which is impossible. Hence, by Proposition 4.5, $U' \subset V_1$. So, if we had any other V_1', V_2' with the property assumed, we would have $V_1' \subset V_1$ and $V_2' \subset V_2$, reordering if necessary (we cannot have both sets included in the same V_i), so unless these are both equalities $V_1 \cup V_2 \setminus (V_1' \cup V_2')$ would be infinite, yielding a contradiction, and concluding the proof. □

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

COROLLARY 4.7. *If U is a diametrizable set with proper diametrizable subsets V_1, V_2, U', such that $U \setminus (V_1 \cup V_2)$ is finite, then V_1, V_2 are disjoint and one of them contains U'.*

PROPOSITION 4.8. *Given a word w, there are only finitely many diametrizable sets containing it.*

Proof. If $w \in W_u$, for some u, then u is a prefix of w, hence there are only finitely many such sets containing w.

If $w \in A_{t,p,r}$, then $l(w) \geqslant p, l(t)$, so there are only finitely many choices for t, p, r.

Finally, suppose that $w \in B_{t,k}$. The length of w must be greater than $l(t)^2$, which gives us finitely many choices for the minimal word t. Fix t. Recalling the definition of B-type sets, we have $B_{t,k} = \cup_{i \in I_k} W_{w_i}$, where $l(w_i) = l(t)^2 + i$. But if m is a non-negative integer, then $\min I_k \geqslant m$ for $k \geqslant 2^m - 1$. Thus, for such k, we get $l(\sigma(B_{t,k})) \geqslant l(t)^2 + m$, so $w \in B_{t,k}$ for only finitely many k, as desired. □

These claims help us to understand better the structure of $\mathcal{D}$. In particular, we can view $\mathcal{D}$ as a binary tree whose nodes are the diametrizable sets, whose root is $W_\emptyset$. Given a set $U \in \mathcal{D}$, the sons V_1, V_2 of the root are given by Corollary 4.7. What is not clear, however, is whether the tree so defined actually contains all the diametrizable sets. But, given any such a set $U \in \mathcal{D}$, either $U = W_\emptyset$ or U is contained in one of the sons of the root, by Corollary 4.7. Proceeding further in this fashion, either we reach U, or we get an infinite collection of diametrizable sets all containing U. But that implies that if we do not reach U, its elements belong to infinitely many members of $\mathcal{D}$, which contradicts Proposition 4.8. Hence this binary tree has precisely $\mathcal{D}$ for its set of nodes. Moreover, to say that a diametrizable set U_1 is a subset of another such set U_2 is equivalent to having U_1 as an ancestor of U_2 in this binary tree. To illustrate what has just been discussed, Figure 1 shows the first few layers of the tree. We refer to this tree as $\mathcal{T}$.

§5. *Ordering $\mathcal{D}$.* In order to finish constructing the counterexample, we must make $\mathcal{D}$ well ordered, so that then we know which diametrizable set is actually which S_n. Defining such an order, and proving that it is in fact what we need, is the purpose of this section.

Consider the relation $<$ on $\mathcal{D}$, given as follows: if $U \neq V$, we say that $U < V$ if any of the following properties holds:

O1 $l(\sigma(U)) < l(\sigma(V))$;

O2 $l(\sigma(U)) = l(\sigma(V))$ and $U \supset V$;

O3 $l(\sigma(U)) = l(\sigma(V))$, neither is contained in the other, and we have that either U is of A-type but V is not, or U is of B-type and V is of W-type;

O4 $l(\sigma(U)) = l(\sigma(V))$, neither is contained in the other, they are of the same type and $\sigma(U)$ is alphabetically before $\sigma(V)$.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

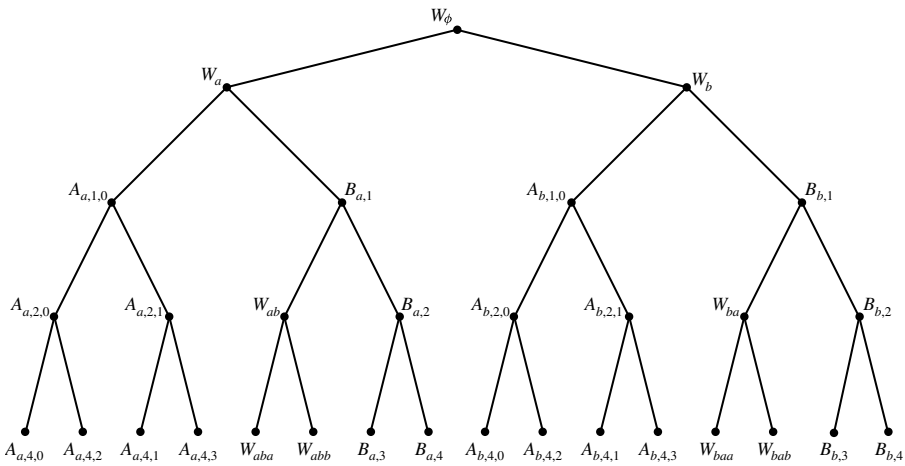


Figure 1: Structure of the collection of diametrizable sets.

PROPOSITION 5.1. *If $U < V$ by **O2** then:*

(i) *both U and V are A-type; or*
(ii) *both U and V are B-type; or*
(iii) *U is B-type and V is W-type.*

Proof. Regard $\mathcal{D}$ as a binary tree $\mathcal{T}$ as described in the concluding remarks of the previous section. Then $U \subset V$ tells us that U is an ancestor of V, and as $U < V$ by **O2**, we have $\sigma(U) = \sigma(V)$. Thus, the shortest word must be the same for all sets on the path from U to V in $\mathcal{T}$. Now, we analyse the splits, that is, given a node, what its sons are. Returning to the choice of diametrizable sets, we see that A-type sets always split into two A-type sets, and B-type sets split into B-type and W-type sets. We conclude that the claim will be true if when a W-type set splits into an A-type and a B-type set, the shortest word of the father is not in any of the sons. But suppose that we have such a situation, a set $T = W_w$, whose children are $A_{t,1,0}$ and $B_{t,1}$. From Case 6 of the proof of Proposition 4.5, recall that $w = t^{l(t)}$. But then $w \notin A_{t,1,0}$ and $w \notin B_{t,1}$, as desired. □

PROPOSITION 5.2. *The relation $<$ on the chosen sets is strict total order, and further collection of chosen sets is well ordered under $<$.*

Proof. If U, V are two distinct sets among the chosen ones, we want that precisely one of $U < V$, $V < U$ is true.

Suppose that neither of these holds. Hence $\sigma(U) = \sigma(V)$. But then their intersection is non-empty, hence one is contained in the other, so $U < V$ or vice versa by **O2**.

Now assume that both hold. Therefore, $l(\sigma(U)) = l(\sigma(V))$, neither is contained in the other (otherwise $U \subset V$ and $V \subset U$, thus $U = V$), they are of the same type, and $\sigma(U) = \sigma(V)$. But as before, these must intersect, leading us to a contradiction.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

Having proved the trichotomy, we now go on to establish the transitivity of the relation. Suppose that we have three chosen sets U, V, T such that $U \prec V \prec T$, from which $U \neq V$, $V \neq T$ follows. Further, we cannot have $T = U$ as this would imply $U \prec V \prec U$. Hence, we can assume that all three sets are distinct. If **O1** holds for $U \prec V$ or $V \prec T$, then it also holds for U, T. So, assume this is not true. Then we have the following cases.

*Case 1: $U \prec V$ by **O2**, $V \prec T$ by **O2**.* Then $U \supsetneq V \supsetneq T$, so $U \prec T$ by **O2** as well.

*Case 2: $U \prec V$ by **O2**, $V \prec T$ by **O3**.* By Proposition 5.1, we see that (as V cannot be of W-type), U and V are of the same type. As $V \subset U$ and $V \not\subset T$, then $U \not\subset T$, so either $T \subset U$ and $U \prec T$ by **O2**, or $U \prec T$ by **O3**.

*Case 3: $U \prec V$ by **O2**, $V \prec T$ by **O4**.* As above, we get $U \not\subset T$. If $T \subset U$ then $T \succ U$. Hence, assume that $T \not\subset U$. Suppose that U is B-type and V is W-type, thus $U \prec T$ by **O3**. Otherwise, by Proposition 5.1, U and V are of the same type, hence so is T, and $\sigma(U) = \sigma(V)$ (because of inclusion), so $U \prec T$ by **O4**.

*Case 4: $U \prec V$ by **O3**, $V \prec T$ by **O2**.* If $T \supset U$, then $U \subset V$, which is impossible, so $U \not\subset T$. If $T \subset U$, we are done, otherwise from Proposition 5.1 we obtain $U \prec T$ by **O3**.

*Case 5: $U \prec V$ by **O3**, $V \prec T$ by **O3**.* Thus U is of A-type, V is of B-type, T is of W-type. By Proposition 5.1, we conclude that $U \not\subset T$, $T \not\subset U$, so $U \prec T$ by **O3**.

*Case 6: $U \prec V$ by **O3**, $V \prec T$ by **O4**.* By Proposition 5.1, $U \not\subset T$, so either $T \subset U$ giving $U \prec T$ by **O2**, or neither is contained in the other and $U \prec T$ by **O3**.

*Case 7: $U \prec V$ by **O4**, $V \prec T$ by **O2**.* Having U as a subset of T implies $U \subset V$, which is impossible. So $U \not\subset T$ and if $T \subset U$, then $U \prec T$ by **O2**. Otherwise, unless U and T are of the same type, we have $U \prec T$ by **O3**, due to Proposition 5.1. Finally, from inclusion we deduce $\sigma(T) = \sigma(U)$ and thus $U \prec T$ by **O4**.

*Case 8: $U \prec V$ by **O4**, $V \prec T$ by **O3**.* If $U \subset T$, by Proposition 5.1, we reach a contradiction. Hence $U \supset T$, so $U \prec T$ by **O2**, or $U \prec T$ by **O3** otherwise.

*Case 9: $U \prec V$ by **O4**, $V \prec T$ by **O4**.* If $T \supset U$ or vice versa, we have $\sigma(T) = \sigma(U)$ which is impossible. As all three sets are of the same type, we get $U \prec T$ by **O4**.

Finally, given a subset P of $\mathcal{D}$, consider its subset P' of those sets $U \in P$ such that $l(\sigma(U)) = \min\{l(\sigma(V)) : V \in P\}$. This is finite by Proposition 4.8, hence we can find the minimal element of P' with respect to $\prec$, which is smaller than any member of $P \setminus P'$ by **O1**, making $\mathcal{D}$ well ordered under $\prec$. □

Hence, as $\mathcal{D}$ is countable, we can take $S_0 = \min \mathcal{D}$ and, for $k \geqslant 1$, $S_k = \min(\mathcal{D} \setminus \{S_0, S_1, \ldots, S_k\})$. Observe that for any given $U \in \mathcal{D}$ there are only

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

finitely many words of length at most $l(\sigma(U))$. Proposition 4.8 then tells us that there are only finitely many $V \in \mathcal{D}$ such that $l(\sigma(V)) \leqslant l(\sigma(U))$. Hence, there are only finitely many $V \in \mathcal{D}$ such that $V < U$, so $U = S_k$ for some $k \in \mathbb{N}$. Therefore, $\{S_0, S_1, S_2, \dots\} = \mathcal{D}$. Note also that $S_0 = X = W_\emptyset$, being the only set with shortest word whose length is non-positive.

5.1. *Proof that $\mathcal{D}$ satisfies the required properties.* All that remains is to show that $S_0, S_1, \dots$ satisfy **A1**–**A6**. Having done most of the work already, the proof of the following claims will either be rather short or the obvious case examination.

PROPOSITION 5.3. $(S_n)_{n\geqslant 0}$ *satisfy* ***A1***.

Proof. Let S_i, S_j be such that $i < j$ and they intersect. Thus, $S_i < S_j$, $l(\sigma(S_i)) \leqslant l(\sigma(S_j))$ and also $S_i \subset S_j$ or $S_i \supset S_j$, by Proposition 4.5. But if $S_i \subset S_j$, then we must have $l(\sigma(S_i)) = l(\sigma(S_j))$ and hence $S_j < S_i$ by **O2**, which is contradiction. Also $S_0 = X$. □

PROPOSITION 5.4. $(S_n)_{n\geqslant 0}$ *satisfy* ***A2***.

Note that this is Proposition 4.8, but we include it here for completeness.

PROPOSITION 5.5. $(S_n)_{n\geqslant 0}$ *satisfy* ***A3***.

Proof. Let N be given. By Corollary 4.7, given S_i, we can find S_j, S_k, disjoint subsets of S_i, which cover all but finitely many elements of S_i, therefore $i < j, k$ by **O1** or **O2**. So we can start with S_0, and perform such splits until we are left with sets $S_{m_1}, S_{m_2}, \dots, S_{m_k}$, with $m_1, m_2, \dots, m_k > N$, which cover almost all elements of X, as in each split we lose only finitely many elements. □

PROPOSITION 5.6. $(S_n)_{n\geqslant 0}$ *satisfy* ***A4***.

Proof. Suppose that $S_{i_1} \supset S_{i_2} \supset \cdots$ for some $i_1 < i_2 < \cdots$. As usual, we consider different cases.

Case 1: there is a set of A-type among these. Let S_{i_k} be such a set. As an A-type set splits into A-type sets, we have that when $n \geqslant k$, $S_{i_n} = A_{w,p_n,r_n}$, for some minimal word w, and integers p_n, r_n, and $p_{n+1} > p_n$, so in fact $p_{n+1} \geqslant 2p_n$, as these are powers of 2. In particular, we deduce that for sufficiently large n, $p_n > nl(w)$ and thus w^n is a prefix of all words in S_{i_n} implying $S_{i_n} \subset W_{w^n}$. By Proposition 4.1, one of aw^n, bw^n is available (suppose without loss of generality the former), hence $aS_{i_n} \subset W_{aw^n} = S_{j_n}$, some j_n.

Consider a cyclic permutation u of w, such that $sw = us$, s being the last character of w. Let v be a word such that $v^\infty = u^\infty$ and $l(u) \geqslant l(v)$. But then, for the cyclic permutation t of v such that for character s' we have $s't = vs'$, we have $s't^\infty = v^\infty = sw^\infty$, so $l(u) = l(t) \geqslant l(w) = l(v)$, proving the minimality of v. If $s = a$, we would have that aw^n is not available as it would be an initial segment of u^∞, but $l(u) = l(w)$. Hence $s = b$, and $bS_{i_n} \subset A_{v,p_n,r} = S_{j_n}$, where $r = r_n + 1$, unless $r_n = p_n - 1$ and then $r = 0$, j_n a suitable index. As j_n tends to infinity as n does, **A4** holds in this case.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

Case 2: there are no A-type sets, but there infinitely many of W-type. Denote the W-type sets among these by $W_{w_1} \supsetneq W_{w_2} \supsetneq \cdots$. As w_i is a prefix of w_{i+1}, for all i, these define an infinite word w whose initial segments are the words w_i. By Proposition 4.1, without loss of generality, aw_i is available infinitely often, so we can take $W_{aw_{i_1}}$ for suitable $i_1 < i_2 < \cdots$, to establish prefix a part of the **A4** property.

Similarly, if there are infinitely many initial segments of bw available, choosing these and their corresponding W-type sets establishes the claim. Suppose the contrary, that is, there is m such that prefixes of bw of length greater than or equal to m are all forbidden. Denote by u_n an initial segment of bw of length n, and let t_n be the shortest word such that $l(t_n)^2 < l(u_n)$ and u_n is the initial segment of t_n^∞. Now suppose that there is no $n \geqslant m$ such that $l(t_{n+1}) > l(t_n)$, hence $l(t_m) \geqslant l(t_{m+1}) \geqslant \cdots$, so $bw = t_n^\infty$, for some n. But this means that the w_i are forbidden from some point, resulting in a contradiction. Thus there must be such an $n \geqslant m$, so pick the smallest possible. Hence u_{n+1} is not the initial segment of t_n^∞, but u_n is both the initial segment of t_n^∞ and t_{n+1}^∞. Further, $l(u_n) \geqslant l(t_n)^2, l(t_{n+1})^2$ so, as before, $t_n^\infty = t_{n+1}^\infty$, but u_{n+1} is a prefix of the later, but not the former, giving a contradiction.

Case 3: almost all sets are of B-type. Given $S_{i_n} = B_{w,k}$, let s be the last character of w, so $sw^{l(w)}$ is forbidden. Let u be the cyclic permutation of w such that $sw = us$, and so u is minimal as well, by arguments in Case 1 of this proof. Then $sB_{w,k} = B_{u,l}$ or $sB_{w,k} = B_{u,l} \setminus W_v$ for suitable index l and word v, so take $S_{j_n} = B_{u,l}$. Let s' be a character not equal to s and u be $\sigma(B_{w,k})$ without the last character. Then as su is forbidden, $s'u$ is available and $s'B_{w,k} \subset W_{s'u} = S_{j_n}$. As the j_n tend to infinity as n does, we are done. □

PROPOSITION 5.7. $(S_n)_{n\geqslant 0}$ *satisfy* ***A5***.

Proof. Let S_n be given. If it is of A or W-type, $S_n \subset W_{\sigma(S_n)}$, so choose character s such that $s\sigma(S_n)$ is available and hence $sS_n \subset W_{s\sigma(S_n)}$ and $S_n < W_{s\sigma(S_n)}$, since $l(\sigma(S_n)) < l(\sigma(W_{s\sigma(S_n)}))$. If S_n is of B-type, then let u be the shortest word in S_n after erasing the last character, so $S_n \subset W_u$ and once again choose $s \in \{a, b\}$ for which su is available, hence $sS_n \subset W_{su}$. If W_{su} intersects W_u, it must be its subset and u prefix of su, but then $u = s^{l(u)}$, so su is forbidden, as $l(u) \geqslant 1$. So W_{su} and W_u are disjoint, hence W_{su} and S_n are also. Combining this with $l(\sigma(S_n)) = l(\sigma(W_{su})$ and comparing the types gives $S_n < W_{su}$ by **O3**. □

PROPOSITION 5.8. $(S_n)_{n\geqslant 0}$ *satisfy* ***A6***.

Proof. Let $w \neq \emptyset$ be given. We will divide the proof into three cases, each showing the claim for a particular set type.

Case 1: suppose, for some word u, that we have $u, wu \in S_n = A_{t,p,r}$, *where t is minimal.* Hence $l(u) > l(t)^2$. Suppose further that $l(w) < l(t)$. So w is an initial segment of u, as u is an initial segment of wu. But then w^2 is also an initial

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

segment of u, etc. up to $w^{l(t)}$ as $l(u) > l(t)^2 > l(w)l(t)$. Hence $t^{l(w)} = w^{l(t)}$, but t is minimal so $l(w) \geqslant l(t)$. Hence we have only finitely many choices for A-type S_n, as $l(w) \geqslant l(t)$ and $p \leqslant l(w)$ must hold.

Case 2: If S_n is of W-type, say W_v, we have $u, wu \in W_v$, v available. As there are only finitely many v such that $l(v) \leqslant l(w)$, without loss of generality $l(v) > l(w)$. We have $u = vr_1$, for some word r_1. Hence $wvr_1 \in W_v$, so $wv \in W_v$. But then $wv = vr_2$, for some r_2, so $v = wv_1$, for some v_1, as $l(w) < l(v)$, and $wwv_1 = wv_1r_2$ implies $wv_1 = v_1r_2$. We can iterate this until $v = w^k v_k$, $l(v_k) < l(w)$. But $wv_k = v_k r_{k+1}$, for some r_{k+1}, making v an initial segment of w^∞, but it is available hence $l(v) \leqslant l(w)^2$, therefore we have finitely many choices for S_n of W-type.

Case 3: finally, suppose that S_n is B-type, and $u, wu \in S_n$. Then we have two available words w_1s_1, w_2s_2, where s_1, s_2 are characters, w_1, w_2 prefixes of t^∞, some minimal word t, $u \in W_{w_1s_1}$, $wu \in W_{w_2s_2}$, $l(w_1), l(w_2) \geqslant l(t)^2$. So $u = w_1s_1r_1$, $wu = w_2s_2r_2$ holds for some words r_1, r_2.

Suppose that $l(w) \leqslant l(t)$. Let wu' be the initial segment of wu of length $l(t)^2$, thus it is a prefix of w_2 hence of t^∞. Thus $wt^{l(t)-1}$ is a prefix of $t^{l(t)}$. Hence, $w^k t^{l(t)-k}$ is a prefix of $t^{l(t)}$ for all $k \in [l(t)]$. In particular, $w^{l(t)}$ is a prefix of $t^{l(t)}$, thus $w^{l(t)} = t^{l(w)}$, implying $w^\infty = t^\infty$, but t is minimal, so $w = t^\alpha$, for some $\alpha \geqslant 1$, thus $w = t$ as $l(w) \leqslant l(t)$. Now we have $wu = w_2s_2r_2 = ww_1s_1r_1 = tw_1s_1r_1$, which differ from t^∞ for the first time at s_1 and s_2, hence $w_2 = tw_1$. But for only finitely many k can we have $B_{w,k}$ having such $W_{w_1s_1}$, $W_{w_2s_2}$ as subsets due to the construction of B-type sets, and the fact that the common difference of arithmetic progression I_k is at least $k/4$.

Now suppose that $l(w) > l(t)$. If the claim is to be false, we can assume without loss of generality that for infinitely many k we have some $u \in B_{t,k}$ and $wu \in B_{t,k}$. Hence we can assume that there are w_1, w_2 prefixes of t^∞, with some characters s_1, s_2, such that w_1s_1, w_2s_2 are available, $u \in W_{w_1s_1}$, $wu \in W_{w_2s_2}$, w_1, w_2 arbitrarily long, thus, say, $l(w_1), l(w_2) > l(w)^2$. So $u = w_1s_1r_1$, $wu = w_2s_2r_2$ for some words r_1, r_2, thus $wu = ww_1s_1r_1 = w_2s_2r_2$. Hence $t^{l(w)}$ is a prefix of w_2 as $l(w_1) > l(w)^2 > l(t)^2$, and similarly $wt^{l(w)-1}$ is an initial segment of ww_1 so both are prefixes of wu, hence $t^{l(w)}$ is a prefix of $wt^{l(w)-1}$, which is then a prefix of $w^2t^{l(2)-2}$, etc. and a prefix of $w^{l(w)}$, so $w^\infty = t^\infty$ and, due to minimality, $w = t^k$, for some $k \geqslant 2$. Hence $w_2s_2r_2 = wu = t^k w_1s_1r_1$, so the character where wu first differs from t^∞ is at the same time at s_1 and s_2, hence $w_2 = t^k w_1 = ww_1$, implying $l(w_2) = l(w_1) + l(w)$, but, as previously explained, this can occur just for finitely many $B_{t,k}$, which proves the claim. □

Having shown the required statements we are ready to conclude the following theorem.

THEOREM. *Given $\lambda \in (0, 1)$, there is a compact (pseudo)metric space (X, d) on which we have continuous functions $f, g : X \to X$ such that, given $x, y \in X$, either $d(f(x), f(y)) \leqslant \lambda d(x, y)$ or $d(g(x), g(y)) \leqslant \lambda d(x, y)$ holds, but no word in f, g has a fixed point.*

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

§6. *Further remarks.* Even though we have given an answer to Austin's question and Stein's conjecture, some natural questions remain to be answered. Indeed, what kind of additional assumptions are necessary to ensure that Stein's conjecture is true? Here we mention two theorems related to our problem that could indicate what approach should be taken.

THEOREM 6.1 [4]. *Let f be an operator on a complete metric space such that $\{f, f^2, \dots, f^n\}$ is a contractive family for some n. Then f has a unique fixed point.*

THEOREM 6.2 [1]. *Let f, g be two operators on a complete metric space such that $\{f, g\}$ is a contractive family and f, g commute. Then, f and g have a unique common fixed point.*

Theorem 6.1 was proved by Merryfield and Stein in [4], and later Merryfield, Rothschild and Stein removed the assumption of continuity in [3]; this statement is also known as the generalized Banach contraction theorem. In [1], Austin proved Theorem 6.2 and conjectured that the result holds for arbitrarily many commuting operators.

Theorem 1.3 tells us that one should not look for a topological condition, and that an algebraic condition as in the theorem of Austin or in the generalized Banach contraction theorem is more convenient. The latter type of the condition directly defines the geometry of the problem, and in general the assumption should, vaguely speaking, provide us with a finite-dimensional problem, and forbid compositions of functions behaving very badly, that is, prevent the tree structure used in the counterexample from occurring. An appropriate assumption could be that $|\{f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_m} : i_1, \dots, i_m \in \{1, 2, \dots, n\}\}|$ is of subexponential or polynomial growth in m.

To conclude, we raise the following question and conjecture which are extensions of a conjecture of Austin in [1].

QUESTION 6.3. What algebraic conditions should be imposed on a finite contractive family $\{f_1, f_2, \dots, f_n\}$ of operators on a complete metric space to ensure the existence of a composition of the functions f_i with a fixed point?

CONJECTURE 6.4. Let $\{f_1, f_2, \dots, f_n\}$ be a finite contractive family of operators on a complete metric space. Suppose that there are $C > 0, k \in \mathbb{N}$ such that $|\{f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_m} : i_1, \dots, i_m \in \{1, 2, \dots, n\}\}| \leqslant Cm^k$. Then $f_1, \dots, f_n$ have a unique common fixed point.

Acknowledgement. My thanks go to Trinity College whose funding and support made this project possible.

References

1. T. Austin, On contractive families and a fixed-point question of Stein. *Mathematika* **52**(1–2) (2005), 115–129.

2. S. Banach, Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fund. Math.* **3** (1922), 133–181.

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press

3. J. Merryfield, B. Rothschild and J. D. Stein Jr, An application of Ramsey's theorem to the Banach contraction mapping principle. *Proc. Amer. Math. Soc.* **130**(4) (2002), 927–933.
4. J. Merryfield and J. D. Stein Jr, A generalization of the Banach contraction mapping theorem. *J. Math. Anal. Appl.* **273**(1) (2002), 112–120.
5. J. D. Stein Jr, A systematic generalization procedure for fixed-point theorems. *Rocky Mountain J. Math.* **30**(2) (2000), 735–754.

Luka Milićević,
Trinity College,
Cambridge CB2 1TQ,
U.K.
E-mail: lm497@cam.ac.uk

https://doi.org/10.1112/S0025579313000296 Published online by Cambridge University Press