Mathematika 64 (2018) 115–123 © 2018 University College London
doi:10.1112/S0025579317000377

MONOCHROMATIC FACTORIZATIONS OF WORDS AND PERIODICITY

CAÏUS WOJCIK AND LUCA Q. ZAMBONI

Abstract. In 2006 Brown asked the following question in the spirit of Ramsey theory: given a non-periodic infinite word $x = x_1x_2x_3\ldots$ with values in a set $\mathbb{A}$, does there exist a finite colouring $\varphi : \mathbb{A}^+ \to C$ relative to which x does not admit a φ-monochromatic factorization, i.e. a factorization of the form $x = u_1u_2u_3\ldots$ with $\varphi(u_i) = \varphi(u_j)$ for all $i, j \geqslant 1$? Various partial results in support of an affirmative answer to this question have appeared in the literature in recent years. In particular it is known that the question admits an affirmative answer for all non-uniformly recurrent words and for various classes of uniformly recurrent words including Sturmian words and fixed points of strongly recognizable primitive substitutions. In this paper we give a complete and optimal affirmative answer to this question by showing that if $x = x_1x_2x_3\ldots$ is an infinite non-periodic word with values in a set $\mathbb{A}$, then there exists a 2-colouring $\varphi : \mathbb{A}^+ \to \{0, 1\}$ such that for any factorization $x = u_1u_2u_3\ldots$ we have $\varphi(u_i) \neq \varphi(u_j)$ for some $i \neq j$.

§1. *Introduction.* Give a set $\mathbb{A}$, let $\mathbb{A}^+$ denote the free semigroup generated by $\mathbb{A}$ consisting of all finite words $u_1u_2\ldots u_n$ with $u_i \in \mathbb{A}$, and $\mathbb{A}^{\mathbb{N}}$ the set of all infinite words $x = x_1x_2x_3\ldots$ with $x_i \in \mathbb{A}$. We say $x \in \mathbb{A}^{\mathbb{N}}$ is *periodic* if $x = u^\omega = uuu\ldots$ for some $u \in \mathbb{A}^+$. The following question was independently posed by Brown in [**3**] and by the second author in [**19**][1]:

Question 1. Let $x \in \mathbb{A}^{\mathbb{N}}$ be non-periodic. Does there exist a finite colouring $\varphi : \mathbb{A}^+ \to C$ relative to which x does not admit a φ-monochromatic factorization, i.e. a factorization of the form $x = u_1u_2u_3\ldots$ with $\varphi(u_i) = \varphi(u_j)$ for all $i, j \geqslant 1$?

The term finite here means that $\mathrm{Card}(C) < +\infty$. A finite colouring $\varphi : \mathbb{A}^+ \to C$ is called a *separating colouring* for x (or a separating $|C|$-colouring for x) if for all factorizations $x = u_1u_2u_3\ldots$ there exist $i, j \geqslant 1$ such that $\varphi(u_i) \neq \varphi(u_j)$. While Brown originally stated it as a question, Question 1 has evolved into a conjecture which states that every non-periodic word admits a separating colouring. We begin by illustrating Question 1 with an example: consider the *Thue–Morse* infinite word

$$x = 01101001100101101001\ldots,$$

Received 19 September 2016, published online 6 February 2018.
MSC (2010): 68R15, 05C55, 05D10 (primary).

[1] The original formulation of the question was stated in terms of finite colourings of the set of all factors of x.

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

where the nth term of x (starting from $n = 0$) is defined as the sum modulo 2 of the digits in the binary expansion of n. The origins of this word go back to the beginning of the last century with the works of Thue [17, 18] in which he proves amongst other things that x is *overlap-free* i.e. x contains no factor of the form uuu' where u' is a non-empty prefix of u.

Consider the 3 colouring $\varphi : \{0, 1\}^+ \to \{0, 1, 2\}$ defined by

$$\varphi(u) = \begin{cases} 0 & \text{if } u \text{ is a prefix of } x \text{ ending with 0,} \\ 1 & \text{if } u \text{ is a prefix of } x \text{ ending with 1,} \\ 2 & \text{if } u \text{ is not a prefix of } x. \end{cases}$$

We claim that no factorization of x is φ-monochromatic. In fact, suppose to the contrary that $x = u_1u_2u_3\ldots$ is a φ-monochromatic factorization of x. Since u_1 is a prefix of x, it follows that $\varphi(u_1) \in \{0, 1\}$, i.e. there exists $a \in \{0, 1\}$ such that each u_i is a prefix of x terminating with a. Pick $i \geqslant 2$ such that $|u_i| \leqslant |u_{i+1}|$. Then as each u_i is a prefix of x, it follows that u_i is a prefix of u_{i+1} and hence au_iu_i is a factor of x. Writing $u_i = va$ (with v empty or in $\{0, 1\}^+$), we have $au_iu_i = avava$ which is an overlap, contradicting that x is overlap-free. This proves that there exists a separating 3-colouring for the Thue–Morse word. Recently Avgustinovich and Parshina [1] proved that it is possible to colour $\{0, 1\}^+$ using only 2 colours in such a way that no factorization of the Thue–Morse word is monochromatic.

Let us remark in the example above that since the Thue–Morse word x is not periodic, each proper suffix of x begins in some factor which is not a prefix of x. This means that each proper suffix x' of x may be written as an infinite concatenation $x' = u_1u_2u_3\ldots$ with $\varphi(u_i) = 2$ for all $i \geqslant 1$, where φ is the 3-colouring of $\{0, 1\}^+$ defined above. Moreover, this monochromatic factorization of x' has an even stronger monochromatic property: the set $\{u_n : n \geqslant 1\}^+$ is also φ-monochromatic (each element has φ colour equal to 2). This is because each element of $\{u_n : n \geqslant 1\}$ is a non-prefix of x and hence the same is true of any concatenation formed by elements from this set. It turns out that a weaker version of this phenomenon is true in greater generality: given any $x \in A^{\mathbb{N}}$ and any finite colouring $\varphi : \mathbb{A}^+ \to C$, one can always find a suffix x' of x which admits a factorization $x' = u_1u_2u_3\ldots$ where $\varphi(u_i \ldots u_j) = \varphi(u_1)$ for all $1 \leqslant i \leqslant j$. This fact may be obtained via a straightforward application of the infinite Ramsey theorem [13] (see [3, 10] or see [15] for a proof by Schützenberger which does not use Ramsey's theorem).

Thus Question 1 belongs to the class of Ramsey type problems in which one tries to show that some abstract form of Ramsey's theorem does not hold in certain settings. For instance, the infinite version of Ramsey's theorem [13] (for colouring of pairs) states that whenever the set $\Sigma_2(\mathbb{N})$ of all 2-element subsets of $\mathbb{N}$ is finitely coloured, there exists an infinite set $X \subseteq \mathbb{N}$ with $\Sigma_2(X)$ monochromatic. Hence the same applies when $\mathbb{N}$ is replaced by $\mathbb{R}$. On the other hand, Sierpiński [16] showed that there exists a finite colouring of $\Sigma_2(\mathbb{R})$ such that there does not exist an uncountable set X with $\Sigma_2(X)$ monochromatic. In other words, Ramsey's theorem does not extend to the

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

uncountable setting in $\mathbb{R}$. Similarly, by a straightforward application of Ramsey's theorem, one deduces that given any finite colouring of $\mathbb{N}$, there exists an infinite $X \subseteq \mathbb{N}$ all of whose pairwise sums $\{n+m : n, m \in X, n \neq m\}$ is monochromatic. Again it follows the same is true with $\mathbb{N}$ replaced by $\mathbb{R}$. On the other hand, Hindman *et al* [6] recently exhibited (using the Continuum Hypothesis CH) the existence of a finite colouring of $\mathbb{R}$ such that there does not exist an uncountable set with all its pairwise sums monochromatic. In other words, this additive formulation of Ramsey's theorem also fails in the uncountable setting in $\mathbb{R}$. A related question in these sorts of problems concerns the least number of colours necessary to avoid the presence of monochromatic subsets of a certain kind. For instance Hindman [5] showed that there exists a 3-colouring of $\mathbb{N}$ such that there does not exist an infinite subset X with $X + X$ monochromatic, and it is an open question of Owings [12] whether the same result may be obtained with only 2 colours. Again, using CH and a few more colours (288 to be precise), it is possible to extend Hindman's result to the reals (see [6, Theorem 2.8]). But again it is not known whether the same result can be obtained with only 2 colours [8]. Thus a stronger version of Question 1 would read: Does every non-periodic word admit a separating 2-colouring?

Various partial results in support of an affirmative answer to Question 1 were obtained in [2, 9–11, 14]. For instance in [9], it is shown that Question 1 admits an affirmative answer for all non-uniformly recurrent words and various classes of uniformly recurrent words including Sturmian words. In [14], Salo and Törmä proved that for every aperiodic linearly recurrent word $x \in \mathbb{A}^{\mathbb{N}}$ there exists a finite colouring of $\mathbb{A}^+$ relative to which x does not admit a monochromatic factorization into factors of increasing lengths. And recently Bernardino *et al* [2] proved that Question 1 admits an affirmative answer for fixed points of strongly recognizable primitive substitutions. In addition to the fact that these partial results apply only to a restricted class of non-periodic words (e.g. Sturmian words or fixed points of certain primitive substitutions), in most cases the number of colours required to colour $\mathbb{A}^+$ to avoid a monochromatic factorization of x is determined to be quite large. For instance in [14], the authors proved that if $x \in \mathbb{A}^{\mathbb{N}}$ is an aperiodic linearly recurrent word, then there exists a constant $K \geqslant 2$ and a colouring $\varphi : \mathbb{A}^+ \to C$ with $\mathrm{Card}(C) = 2 + \sum_{i=0}^{K^5-1} 2K^i(K+1)^{2i}$, such that no factorization of $x = u_1u_2u_3\ldots$, verifying the additional constraint that $|u_i| \leqslant |u_{i+1}|$ for each $i \geqslant 1$, is φ-monochromatic. The constant K above is chosen such that for every factor u of x, every first return w to u satisfies $|w| \leqslant K|u|$ (see for instance [4]). A similar large bound depending on the recognizability index of a substitution is obtained in [2] in the context of fixed points of strongly recognizable substitutions. In contrast, it is shown in [9] that every Sturmian word admits a separating 3-colouring.

In this paper we give a complete and optimal affirmative answer to Question 1 by showing that for every non-periodic word $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$, there exists a 2-colouring $\varphi : \mathbb{A}^+ \to \{0, 1\}$ relative to which no factorization of x is φ-monochromatic. Moreover, this is a characterization of periodicity of infinite words.

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

THEOREM 1. *Let $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$ be an infinite word. Then x is periodic if and only if for every 2-colouring $\varphi : \mathbb{A}^+ \to \{0, 1\}$ there exists a φ-monochromatic factorization of x, i.e. a factorization $x = u_1u_2u_3\ldots$ such that $\varphi(u_i) = \varphi(u_j)$ for all $i, j \geqslant 1$.*

Theorem 1 has several immediate consequences: for instance, fix a symbol $a \in \mathbb{A}$ and suppose an infinite word $x \in \mathbb{A}^{\mathbb{N}}$ admits a factorization $x = u_1u_2u_3\ldots$ where each u_i is a prefix of x *rich* in the symbol $a \in \mathbb{A}$, meaning that for each $i \geqslant 1$ and for each factor v of x with $|v| = |u_i|$ we have $f_a(u_i) \geqslant f_a(v)$, where $f_a(u_i)$ denotes the frequency of a in u_i. Then $\{n \in \mathbb{N} : x_n = a\}$ is a finite union of arithmetic progressions. As another application of Theorem 1 we show that if u_1, $u_2, u_3, \ldots, u_{2k+1} \in A^+$ (with k a positive integer), then any $x \in \mathbb{A}^{\mathbb{N}}$ belonging to $\{u_1, u_2\}^{\mathbb{N}} \cap \{u_2, u_3\}^{\mathbb{N}} \cap \cdots \cap \{u_{2k}, u_{2k+1}\}^{\mathbb{N}} \cap \{u_{2k+1}, u_1\}^{\mathbb{N}}$ is periodic.

Theorem 1 may be reformulated in the language of ultrafilters as follows: let $\beta\mathbb{A}^+$ denote the Stone–Čech compactification of the discrete semigroup $\mathbb{A}^+$ which we regard as the set of all ultrafilters on $\mathbb{A}^+$, identifying the points of $\mathbb{A}^+$ with the principal ultrafilters. As is well known, the operation of concatenation on $\mathbb{A}^+$ extends uniquely to $\beta\mathbb{A}^+$ making $\beta\mathbb{A}^+$ a compact right topological semigroup with $\mathbb{A}^+$ contained in its topological centre (see for instance [7]). In particular by the Ellis–Numakura lemma, $\beta\mathbb{A}^+$ contains an idempotent element i.e. an element p verifying $p \cdot p = p$. We show that an infinite word $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$ is periodic if and only if there exists $p \in \beta\mathbb{A}^+$ such that for each $A \in p$ there exists a factorization $x = u_1u_2u_3\ldots$ with each $u_i \in A$. Moreover p as above may be taken to be an idempotent element of $\beta\mathbb{A}^+$.

§2. *Proof of Theorem 1 and applications.* Before embarking on the proof of Theorem 1, we introduce some notation which will be relevant in what follows. Let $\mathbb{A}$ be a set (the *alphabet*). We do not assume that the cardinality of $\mathbb{A}$ is finite. Let $\mathbb{A}^*$ denote the set of all finite words $u = u_1u_2\ldots u_n$ with $u_i \in \mathbb{A}$. We call n the *length* of u and denote it $|u|$. The empty word is denoted ε and by convention $|\varepsilon| = 0$. We put $\mathbb{A}^+ = \mathbb{A}^* \setminus \{\varepsilon\}$. For $u \in \mathbb{A}^+$ and $a \in \mathbb{A}$, we let $|u|_a$ denote the number of occurrences of a in u. Let $\mathbb{A}^{\mathbb{N}}$ denote the set of all right sided infinite words $x = x_1x_2x_3\ldots$ with values in $\mathbb{A}$. More generally, for $x \in \mathbb{A}^{\mathbb{N}}$ and $A \subseteq \mathbb{A}^+$, we write $x \in A^{\mathbb{N}}$ if $x = u_1u_2u_3\ldots$ with $u_i \in A$ for all $i \geqslant 1$, that is in case x factors over the set A. We say x is *periodic* if $x \in \{u\}^{\mathbb{N}}$ for some $u \in \mathbb{A}^+$.

Proof of Theorem 1. First assume $x = x_0x_1x_2\ldots \in \mathbb{A}^{\mathbb{N}}$ is periodic, i.e. $x \in \{u\}^{\mathbb{N}}$ for some $u \in \mathbb{A}^+$. Then the factorization $x = u_1u_2u_3\ldots$ with each $u_i = u$ is φ-monochromatic for any choice of $\varphi : \mathbb{A}^+ \to \{0, 1\}$. Next assume x is not periodic. We will define a 2-colouring $\varphi : \mathbb{A}^+ \to \{0, 1\}$ with the property that no factorization of x is φ-monochromatic. Pick any total order on the set $\mathbb{A}$ and let $\prec$ denote the induced lexicographic order on $\mathbb{A}^+$ and $\mathbb{A}^{\mathbb{N}}$. For $u, v \in \mathbb{A}^+$ with $|u| = |v|$, we write $u \preccurlyeq v$ if either $u \prec v$ or $u = v$. For each $n \geqslant 1$, let $P_x(n)$ denote the prefix of x of length n, and for each $y \in \mathbb{A}^{\mathbb{N}}$, let $x \wedge y$ denote the longest common prefix of x and y. Define $\varphi : \mathbb{A}^+ \to \{0, 1\}$ by the following rules.

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

If u is not a prefix of x, then set

$$\varphi(u) = \begin{cases} 0 & \text{if } u \prec P_x(|u|), \\ 1 & \text{if } P_x(|u|) \prec u. \end{cases} \tag{1}$$

If u is a prefix of x, say $x = uy$ with $y \in \mathbb{A}^{\mathbb{N}}$, then set

$$\varphi(u) = \begin{cases} 0 & \text{if } x \prec y, \\ 1 & \text{if } y \prec x. \end{cases} \tag{2}$$

Since x is not periodic, for every proper suffix y of x we have either $x \prec y$ or $y \prec x$, whence φ is well defined. We observe that φ has the following key property: if u is any non-empty prefix of x, say $x = uy$ with $y \in \mathbb{A}^{\mathbb{N}}$, then there exists $N \geqslant 0$ such that $\varphi(u) \neq \varphi(v)$ for every prefix v of y with $|v| > N$. In fact, let $N = |x \wedge y|$. Then if v is a prefix of y with $|v| > N$, then $v \neq P_x(|v|)$. So if $\varphi(u) = 0$, then $x \prec y$ and hence $P_x(|v|) \prec v$, whence $\varphi(v) = 1$. Similarly if $\varphi(u) = 1$, then $y \prec x$ and hence $v \prec P_x(|v|)$, whence $\varphi(v) = 0$. The other key property of φ is given by the following lemma.

LEMMA 2. *For all $u \in \mathbb{A}^+$, if $u = u_1 u_2 \dots u_k$ with $u_i \in \mathbb{A}^+$, then there exists $1 \leqslant i \leqslant k$ such that $\varphi(u_i) = \varphi(u)$.*

Proof. Fix $a \in \{0, 1\}$. We will prove by induction on k that if $u = u_1 u_2 \dots u_k$ and $\varphi(u_i) = a$ for all $1 \leqslant i \leqslant k$, then $\varphi(u) = a$. Without loss of generality we may assume $a = 0$ for otherwise we could replace $\prec$ by the reverse order. For $k = 1$ the result is immediate. Next we consider the case $k = 2$, i.e. we assume $u = u_1 u_2$ and $\varphi(u_1) = \varphi(u_2) = 0$ and we will show that $\varphi(u) = 0$. If u_1 is not a prefix of x, then $u_1 \prec P_x(|u_1|)$ and hence $u \prec P_x(|u|)$, whence $\varphi(u) = 0$. Thus we may suppose that u_1 is a prefix of x. Let u_2' be such that $u_1 u_2'$ is a prefix of x and $|u_2| = |u_2'|$. If $u_2 \prec u_2'$, then $u = u_1 u_2 \prec u_1 u_2' = P_x(|u|)$, whence $\varphi(u) = 0$. Thus we may suppose that $u_2' \preceq u_2$. Also since $\varphi(u_1) = 0$ it follows that $P_x(|u_2'|) \preceq u_2'$, whence $P_x(|u_2'|) \preceq u_2' \preceq u_2$. If any one of these two inequalities were strict, then $P_x(|u_2|) \prec u_2$ which would imply that $\varphi(u_2) = 1$, a contradiction. Thus we must have $u_2 = u_2' = P_x(|u_2|)$, i.e. both u and u_2 are prefixes of x. Let $x' \in \mathbb{A}^{\mathbb{N}}$ be such that $x = u_1 x'$ and let $w \in \mathbb{A}^+$ denote the longest common prefix between x and x'. Since $\varphi(u_1) = 0$, there exist symbols $a, b \in \mathbb{A}$ with $a \prec b$ such that $u_1 wb$ and wa are each a prefix of x. Also since u_2 is a prefix of both x and x' we have $|u_2| \leqslant |w|$ and hence we can write $w = u_2 v$ for some $v \in \mathbb{A}^*$. So each of $u_1 u_2 vb = uvb$ and $u_2 va$ is a prefix of x. Finally since $\varphi(u_2) = 0$, we have that $P_x(|va|) \preceq va \prec vb$, whence $\varphi(u) = 0$. This completes the case $k = 2$. Now, let $n > 2$ and suppose the result of the lemma holds for all $k < n$, and suppose $u = u_1 u_2 \dots u_n$ with $\varphi(u_i) = a$ for all $1 \leqslant i \leqslant n$. Then by considering the factorization $u = (u_1 u_2) u_3 \dots u_n$ of length $n - 1$ and the fact that $\varphi(u_1 u_2) = a$ (which comes from the case $k = 2$), we deduce by induction hypothesis that $\varphi(u) = a$ as required. □

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

To complete the proof of Theorem 1, we will show that no factorization of x is φ-monochromatic. Let $x = u_1u_2u_3\ldots$ with $u_i \in \mathbb{A}^+$. Pick $N \geqslant 0$ such that $\varphi(v) \neq \varphi(u_1)$ for all prefixes v of $u_2u_3\ldots$ with $|v| > N$. Pick $k \geqslant 2$ such that $|u_2\ldots u_k| > N$. By the previous lemma, there exists $2 \leqslant i \leqslant k$ such that $\varphi(u_i) = \varphi(u_2\ldots u_k)$. Hence $\varphi(u_i) = \varphi(u_2\ldots u_k) \neq \varphi(u_1)$. □

We end this section with some applications of Theorem 1 or its proof. An infinite word $x \in \mathbb{A}^{\mathbb{N}}$ is called *Lyndon* if there exists an order on $\mathbb{A}$ relative to which x is strictly smaller than each of its proper suffixes. Our first application is the following result originally proved in [9].

COROLLARY 3. *A Lyndon word $x \in \mathbb{A}^{\mathbb{N}}$ does not admit a prefixal factorization, i.e. a factorization of the form $x = u_1u_2u_3\ldots$ where each u_i is a prefix of x.*

Proof. Suppose to the contrary that $x = u_1u_2u_3\ldots$ is a prefixal factorization of x. Let $\varphi : \mathbb{A}^+ \to \{0, 1\}$ be the separating 2-colouring for x defined in (1) and (2). Note that x being Lyndon is not-periodic. Since x is Lyndon we have that $\varphi(u_i) = 0$ for all $i \geqslant 1$, contradicting that φ is a separating 2-colouring for x. □

Let $x \in \mathbb{A}^{\mathbb{N}}$ and $a \in \mathbb{A}$. A factor u of x is said to be *rich* in a if $|u|_a \geqslant |v|_a$ for all factors v of x with $|v| = |u|$. Theorem 6.7 in [9] states that a Sturmian word $x \in \{0, 1\}^{\mathbb{N}}$ does not admit a factorization of the form $x = u_1u_2u_3\ldots$ where each u_i is a prefix of x rich in the same letter $a \in \{0, 1\}$. The following generalizes this result to all binary non-periodic words.

COROLLARY 4. *Let $x \in \{0, 1\}^{\mathbb{N}}$ and $a \in \{0, 1\}$. Suppose x admits a prefixal factorization $x = u_1u_2u_3\ldots$ with each u_i rich in a. Then x is periodic.*

Proof. Suppose to the contrary that x is not periodic. Let $\varphi : \{0, 1\}^+ \to \{0, 1\}$ be the separating 2-colouring for x defined in (1) and (2) relative to the order on $\{0, 1\}$ where a is taken to be the least element. For each $i \geqslant 1$, writing $x = u_iy_i$ with $y_i \in \{0, 1\}^{\mathbb{N}}$, we claim that $x \prec y_i$. Otherwise if $y_i \prec x$, then we can write $x = zbx' = u_izay'$ for some $z \in \{0, 1\}^*$, $x', y' \in \{0, 1\}^{\mathbb{N}}$ and where $\{a, b\} = \{0, 1\}$. But then the factor u_i' of length $|u_i|$ immediately preceding the suffix y' of x would contain one more occurrence of the symbol a than u_i, contradicting that u_i was rich in a. Having established that $x \prec y_i$, it follows that $\varphi(u_i) = 0$ for all $i \geqslant 1$ contradicting that φ is a separating 2-colouring for x. □

COROLLARY 5. *Let $\mathbb{A}$ be an arbitrary set, $x \in \mathbb{A}^{\mathbb{N}}$ and $a \in \mathbb{A}$. Suppose x admits a prefixal factorization $x = u_1u_2u_3\ldots$ with each u_i rich in a. Then $\{n \in \mathbb{N} : x_n = a\}$ is a finite union of (infinite) arithmetic progressions.*

Proof. Consider the morphism $\phi : \mathbb{A} \to \{0, 1\}$ given by $\phi(a) = 1$ and $\phi(b) = 0$ for all $b \in \mathbb{A}\setminus\{a\}$. Then applying ϕ to the prefixal factorization $x = u_1u_2u_3\ldots$ gives a prefixal factorization of $\phi(x)$ in which each $\phi(u_i)$ is rich in 1. Hence by Corollary 4 we have that $\phi(x)$ is periodic and hence the set of occurrences of 1 in $\phi(x)$, which is equal to the set of occurrences of a in x, is a finite union of arithmetic progressions. □

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

COROLLARY 6. *Let $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$ and k be a positive integer. Let $B \subseteq \mathbb{A}^+$ with $\mathrm{Card}(B) \geqslant 2k-1$. Suppose that $x \in A^{\mathbb{N}}$ for every k-element subset A of B. Then x is periodic.*

Proof. Let us assume to the contrary that x is not periodic. By Theorem 1 there exists a 2-colouring $\varphi : \mathbb{A}^+ \to \{0,1\}$ relative to which no factorization of x is φ-monochromatic. Let $\Sigma_k(B)$ denote the set of all k-element subsets of B. By assumption x factors over each $A \in \Sigma_k(B)$. On the other hand, since $\mathrm{Card}(B) \geqslant 2k-1$, it follows that there exists a φ-monochromatic subset $A \in \Sigma_k(B)$. This gives rise to a φ-monochromatic factorization of x, a contradiction. □

COROLLARY 7. *Let $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$ and k be a positive integer. Let $u_1, u_2, u_3, \ldots, u_{2k+1} \in A^+$ and suppose $x \in \{u_1,u_2\}^{\mathbb{N}} \cap \{u_2,u_3\}^{\mathbb{N}} \cap \cdots \cap \{u_{2k}, u_{2k+1}\}^{\mathbb{N}} \cap \{u_{2k+1},u_1\}^{\mathbb{N}}$. Then x is periodic.*

Proof. Suppose to the contrary that x is not periodic. Pick any separating 2-colouring $\varphi : \mathbb{A}^+ \to \{0,1\}$ for x. Then $\varphi(1) = \varphi(2i+1)$ for each $1 \leqslant i \leqslant k$. Thus $x \notin \{u_{2k+1}, u_1\}^{\mathbb{N}}$, a contradiction. □

§3. *A reformulation of Theorem 1 in the language of ultrafilters.* In this section we give a reformulation of Theorem 1 in the language of ultrafilters. Given a set S, let βS denote the Stone–Čech compactification of S (relative to the discrete topology on S) which we regard as the collection of all ultrafilters p on S. Theorem 1 may be reformulated in the language of ultrafilters as follows.

THEOREM 8. *Let $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$ be an infinite word. Then x is periodic if and only if there exists $p \in \beta\mathbb{A}^+$ such that for each $A \in p$ there exists a factorization $x = u_1u_2u_3\ldots$ with each $u_i \in A$.*

Proof. Suppose x is periodic, i.e. $x \in \{u\}^{\mathbb{N}}$ for some $u \in \mathbb{A}^+$. Then clearly x factors over each set A belonging to the principal ultrafilter $e(u) = \{A \subseteq \mathbb{A}^+ : u \in A\}$. Conversely, suppose there exists $p \in \beta\mathbb{A}^+$ such that for each $A \in p$ there exists a factorization $x = u_1u_2u_3\ldots$ with each $u_i \in A$. Let $\varphi : \mathbb{A}^+ \to \{0,1\}$ be any 2-colouring of $\mathbb{A}^+$ and consider the partition $\mathbb{A}^+ = \varphi^{-1}(0) \cup \varphi^{-1}(1)$. Since $\mathbb{A}^+ \in p$, it follows that $\varphi^{-1}(a) \in p$ for some $a \in \{0,1\}$. Thus there exists a factorization $x = u_1u_2u_3\ldots$ with each $u_i \in \varphi^{-1}(a)$. In other words, x admits a φ-monochromatic factorization. By Theorem 1 we conclude that x is periodic. □

Remark 9. It can be shown that Theorem 8 is actually an equivalent reformulation of Theorem 1. More precisely, given an infinite word $x = x_1x_2x_3\ldots \in \mathbb{A}^{\mathbb{N}}$, the following statements are easily seen to be equivalent:

(a) there exists $p \in \beta\mathbb{A}^+$ such that for each $A \in p$ we can write $x = u_1u_2u_3\ldots$ with each $u_i \in A$;
(b) for each finite colouring $\varphi : \mathbb{A}^+ \to C$ there exists a φ-monochromatic factorization of x.

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

The operation of concatenation on $\mathbb{A}^+$ extends uniquely to $\beta\mathbb{A}^+$ making $\beta\mathbb{A}^+$ a compact right topological semigroup with $\mathbb{A}^+$ contained in its topological centre (see for instance [7]). The operation $\cdot$ on $\beta\mathbb{A}^+$ is defined as follows: for $p, q \in \beta\mathbb{A}^+$

$$p \cdot q = \{A \subseteq \mathbb{A}^+ : \{u \in \mathbb{A}^+ : u^{-1}A \in q\} \in p\}, \tag{3}$$

where $u^{-1}A = \{v \in \mathbb{A}^+ : uv \in A\}$. As a consequence of the Ellis–Numakura lemma, $\beta\mathbb{A}^+$ contains an idempotent element, i.e. an element p verifying $p \cdot p = p$. We now show that p in Theorem 8 may be taken to be an idempotent element of $\beta\mathbb{A}^+$. For $x = x_1x_2x_3 \ldots \in \mathbb{A}^{\mathbb{N}}$, we set

$$\mathscr{F}(x) = \{A \subseteq \mathbb{A}^+ : x \in A^{\mathbb{N}}\}$$

and

$$\mathcal{U}(x) = \{p \in \beta\mathbb{A}^+ : p \subseteq \mathscr{F}(x)\}.$$

Thus $p \in \mathcal{U}(x)$ if and only if x factors over every $A \subseteq \mathbb{A}^+$ belonging to p.

THEOREM 10. *Let $x = x_0x_1x_2 \ldots \in \mathbb{A}^{\mathbb{N}}$. Then the following are equivalent:*

(i) *x is periodic;*
(ii) *$\mathcal{U}(x)$ is a closed sub-semigroup of $\beta\mathbb{A}^+$;*
(iii) *$\mathcal{U}(x)$ contains an idempotent element;*
(iv) *and the set* $\mathrm{Pref}(x)$ *consisting of all prefixes of x is an IP-set.*

Proof. We first note that for each infinite word x, the set $\mathcal{U}(x)$ is a closed subset of $\beta\mathbb{A}^+$. In fact, suppose $p \in \beta\mathbb{A}^+ \setminus \mathcal{U}(x)$, then there exists $A \in p$ with $A \notin \mathscr{F}(x)$. Then $\overline{A}$ is an open neighbourhood of p and any $q \in \overline{A}$ contains the set A and hence is not in $\mathcal{U}(x)$. To see that (i) $\Longrightarrow$ (ii), suppose that x is periodic. By Theorem 8 we have $\mathcal{U}(x) \neq \emptyset$. It remains to show that $p \cdot q \in \mathcal{U}(x)$ whenever p, $q \in \mathcal{U}(x)$. Pick the shortest $u \in \mathbb{A}^+$ such that $x \in \{u\}^{\mathbb{N}}$ and set $\mathcal{A} = \{u^j : j \in \mathbb{N}\}$. Then $p \in \mathcal{U}(x)$ if and only if $\mathcal{A} \in p$. Let $A \in p \cdot q$ with $p, q \in \mathcal{U}(x)$. Then by (3) we have $\{s \in \mathbb{A}^+ : s^{-1}A \in q\} \in p$. Since $p \in \mathcal{U}(x)$ it follows that $\{s \in \mathbb{A}^+ : s^{-1}A \in q\} \cap \mathcal{A} \neq \emptyset$. Pick $n \in \mathbb{N}$ such that $u^n \in \{s \in \mathbb{A}^+ : s^{-1}A \in q\}$. Then $\{t \in \mathbb{A}^+ : u^nt \in A\} \in q$ and since $q \in \mathcal{U}(x)$ it follows that there exists $m \in \mathbb{N}$ with $u^m \in \{t \in \mathbb{A}^+ : u^nt \in A\}$. In other words $u^{n+m} \in A$ and hence x factors over A. Thus $p \cdot q \in \mathcal{U}(x)$. The implication (ii) $\Longrightarrow$ (iii) follows from the Ellis–Numakura lemma. To see (iii) $\Longrightarrow$ (iv) pick an idempotent element $p \in \mathcal{U}(x)$. We note that $\mathrm{Pref}(x)$ belongs to every $q \in \mathcal{U}(x)$. In fact, suppose that $\mathrm{Pref}(x) \notin q$, for some $q \in \mathcal{U}(x)$. Then $\mathbb{A}^+ \setminus \mathrm{Pref}(x) \in q$ which implies that x factors over $\mathbb{A}^+ \setminus \mathrm{Pref}(x)$. But this is a contradiction since in any factorization of x, the first term occurring in the factorization belongs to $\mathrm{Pref}(x)$. Thus in particular $\mathrm{Pref}(x) \in p$. Since p is an idempotent, it follows that $\mathrm{Pref}(x)$ is an IP-set. Finally, to see that (iv) $\Longrightarrow$ (i) assume that $\mathrm{Pref}(x)$ is an IP-set. Then $\mathrm{Pref}(x)$ contains $FP(\langle s_n\rangle_{n\in\mathbb{N}})$ for some infinite sequence $\langle s_n\rangle_{n\in\mathbb{N}}$ of prefixes of x. This means that for each $n \geqslant 2$, both $s_1s_2 \ldots s_n$ and $s_2 \ldots s_n$ are prefixes of x. Thus $x = s_1x$ which implies that x is periodic. □

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press

Acknowledgement. We are extremely grateful to the anonymous referee for suggestions on improving the exposition of this paper.

References

1. S. Avgustinovich and O. Parshina, private communication.
2. A. Bernardino, R. Pacheco and M. Silva, Coloring factors of substitutive infinite words. *Discrete Math.* **340** (2017), 443–451.
3. T. C. Brown, Colorings of the factors of a word. *Preprint*, 2006, Department of Mathematics, Simon Fraser University, Canada.
4. F. Durand, B. Host and C. Skau, Substitutional dynamical systems, Bratteli diagrams and dimension groups. *Ergodic Theory Dynam. Systems* **19** (1999), 953–993.
5. N. Hindman, Partitions and sums of integers with repetition. *J. Combin. Theory Ser. A* **27** (1979), 19–32.
6. N. Hindman, I. Leader and D. Strauss, Pairwise sums in colourings of the reals. *Abh. Math. Semin. Univ. Hambg.* (to appear).
7. N. Hindman and D. Strauss, *Algebra in the Stone–Čech Compactification: Theory and Applications*, 2nd edn., de Gruyter (Berlin, 2012).
8. I. Leader, private communication.
9. A. de Luca, E. Pribavkina and L. Q. Zamboni, A coloring problem for infinite words. *J. Combin. Theory Ser. A* **125** (2014), 306–332.
10. A. de Luca and L. Q. Zamboni, On some variations of coloring problems of infinite words. *J. Combin. Theory Ser. A* **137** (2016), 166–178.
11. A. de Luca and L. Q. Zamboni, On prefixal factorizations of words. *European J. Combin.* **52**(part A) (2016), 59–73.
12. J. Owings, Problem E2494. *Amer. Math. Monthly* **81** (1974), 902.
13. F. P. Ramsey, On a problem of formal logic. *Proc. Lond. Math. Soc.* **30** (1930), 264–286.
14. V. Salo and I. Törmä, Factor colorings of linearly recurrent words. *Preprint*, 2015, arXiv:1504.0582.
15. M. P. Schützenberger, Quelques problèmes combinatoires de la théorie des automates, Cours professé à l'Institut de Programmation en 1966/67, notes by J.-F. Perrot, http://igm.univ-mlv.fr/∼berstel/Mps/Cours/PolyRouge.pdf.
16. W. Sierpiński, Sur un problème de la théorie des relations. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2** (1933), 285–287.
17. A. Thue, Über unendliche Zeichenreihen. *Skr. K. Nor. Vidensk. Selsk.* **7** (1906), 1–22.
18. A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Skr. K. Nor. Vidensk. Selsk.* **1** (1912), 1–67.
19. L. Q. Zamboni, A note on coloring factors of words, in Oberwolfach Report 37/2010, Mini-workshop: combinatorics on words, August 22–27, 2010, 42–44.

Caïus Wojcik,
Institut Camille Jordan,
Université Lyon 1,
F69622 Villeurbanne,
France
E-mail: caius.wojcik@gmail.com

Luca Q. Zamboni,
Institut Camille Jordan,
Université Lyon 1,
F69622 Villeurbanne,
France
E-mail: zamboni@math.univ-lyon1.fr

https://doi.org/10.1112/S0025579317000377 Published online by Cambridge University Press