Known theorems about substitutions, and homemade proofs for them

Ville Salo and Ilkka Törmä

[1] TUCS – Turku Center for Computer Science
University of Turku
vosalo@utu.fi
[2] University of Turku
iatorm@utu.fi

Abstract. We give new proofs for two known results concerning substitutions. More precisely, we prove that a periodic point of a primitive aperiodic substitution cannot contain arbitrarily large powers, and that a uniform primitive aperiodic substitution is recognizable.

1 Introduction

In the field of combinatorics of words, substitutions play an important role. They are defined as functions that, given a word w, replace each letter of w by some other word. By applying a (non-pathological) substitution σ to a single letter repeatedly, one obtains an infinite word w which is a fixed point of σ, that is, $\sigma(w) = w$. Words constructed in this manner are very regular and enjoy many useful stuctural properties, so an extensive theory has been developed on substitutions. The class of *primitive substitutions* is particularly well-behaved, giving rise to uniformly recurrent fixed points.

Consider again the equation $\sigma(w) = w$. By the definition of a substitution, w is an infinite concatenation of the σ-images of its own letters. Suppose now that we are given a factor v of w, and asked in which way the concatenation factors v into subwords. If there is a unique answer for long enough factors v apart from its borders, that is, if we can say from which letters the central part of v has 'come from', we say σ is a *recognizable substitution*. The result that all aperiodic primitive substitutions are recognizable is one of the cornerstones of their theory.

Another classical result on primitive substitutions, which is also used to prove their recognizability, is that they cannot nontrivially generate arbitrarily large powers. In other words, for an aperiodic fixed point w of a primitive substitution, there exists an exponent N such that no word of the form u^N is a factor of w. This result tells us something fundamental about which kinds of infinite words are generated by substitutions.

Both of these results were proved by Mossé in [3], but the article in question is written in French. Because of the fundamental nature of these results in the theory of substitutions, there has been some desire in the community to make

them more accessible. For example, the proof of Mossé has been translated to English in [1], where it was used as a lemma in a slightly modified form.

In this article, we present new proofs for the following results: a primitive substitution cannot nontrivially generate arbitrarily large powers in the sense explained above, and a *uniform* primitive substitution with an aperiodic fixed point is recognizable. A substitution is uniform if the images of the letters are of the same length. The proof of the former result uses the classical theorem of Fine and Wilf, and the second is proved with a pigeonhole argument and is somewhat geometric in nature.

2 Definitions

Let A be a finite set, called the *alphabet*. We denote by A^* the set of finite words over A, by $A^{\mathbb{N}}$ the set of infinite words over A, and by A^ω the union of these sets. The empty word is denoted by λ, and we denote $A^+ = A^* - \{\lambda\}$. For a word $u \in A^\omega$, denote by $|u|$ the length of u and by $u_{[\![i]\!]}$ the ith letter of u, the indexing starting at 0. We also use the shorthand notation $u_{[\![i,j]\!]} = u_{[\![i]\!]}u_{[\![i+1]\!]} \cdots u_{[\![j]\!]}$. For two words $u \in A^*$ and $v \in A^\omega$, we say u is a *factor* of v, if $v_{[\![i,i+|u|-1]\!]} = u$ for some i. For $h, m \in \mathbb{N}$, we say u is an h mod m*-factor* of v, if $i \equiv h \bmod m$. The set of factors of v of length n is denoted $\mathcal{F}_n(v)$, and the set of all factors by $\mathcal{F}(v)$. If $v = uw$ for some $w \in A^\omega$, we denote $w = u^{-1}v$. The *shift map* is the function $\tau : A^{\mathbb{N}} \to A^{\mathbb{N}}$ defined by $\tau(x)_{[\![i]\!]} = x_{[\![i+1]\!]}$.

For $u, v \in A^*$, we say v is u*-periodic* if $v = u^n$ for some $n \in \mathbb{N}$. We say v is *primitive* if v is only u-periodic for $u = v$. If $u = u_1u_2$ and $v = u_2u_1$ for some words u_1, u_2, then we say u and v are *conjugate*. We say $x \in A^{\mathbb{N}}$ is *eventually periodic*, if there exists $p \in \mathbb{N}$ such that $x_{[\![i]\!]} = x_{[\![i+p]\!]}$ for all sufficiently large $i \in \mathbb{N}$. A word $u \in A^+$ *occurs in* x *with gap* G, if $u \in \mathcal{F}(v)$ for every $v \in \mathcal{F}_G(x)$. We say that u occurs in x *with bounded gaps*, if such a G exists. We say x is *uniformly recurrent* if every $u \in \mathcal{F}(x)$ occurs in x with bounded gaps. For a function $f : A^{\mathbb{N}} \to A^{\mathbb{N}}$, we say x is f*-periodic*, if there exists $p \in \mathbb{N}$ such that $f^p(x) = x$.

A *substitution* is a function $\sigma : A^* \to A^*$ such that $\sigma(uv) = \sigma(u)\sigma(v)$ for all $u, v \in A^*$. Thus a substitution is defined by the images of the letters in A. We assume that $\sigma(a) \neq \lambda$ for all $a \in A$, so that we can extend all substitutions to A^ω. We say σ is *primitive* if there exists $k \in \mathbb{N}$ such that for all $a, b \in A$, we have that $b \in \mathcal{F}(\sigma^k(a))$. It is well-known that in this case every σ-periodic point is uniformly recurrent. We say σ is *aperiodic* if no σ-periodic point is eventually periodic. Otherwise we say σ is *periodic*. We say σ is *uniform* if for all $a, b \in A$, we have $|\sigma(a)| = |\sigma(b)|$.

For the rest of this article, we fix σ to be a primitive substitution, with $k \in \mathbb{N}$ as in the definition.

Let x be any σ-periodic point. The *orbit closure* $X(\sigma)$ of σ is the set of words $y \in A^{\mathbb{N}}$ with $F(y) \subset F(x)$. We also denote $\mathcal{F}_n(\sigma) = \mathcal{F}_n(x)$ and $\mathcal{F}(\sigma) = \mathcal{F}(x)$. These notions do not depend on the choice of x by the primitivity of σ. For

$y \in X(\sigma)$, we let

$$E(y) = \{0\} \cup \{|\sigma(y_{[\![0,n]\!]})| \mid n \geq 0\}.$$

This is the set of indices at which the images of letters partitioning the word $\sigma(y)$ begin. We say σ is *recognizable* if there exists $N \in \mathbb{N}$ such that whenever $y \in X(\sigma)$, $|w| \geq N$ and $\sigma(y)_{[\![i,i+|w|-1]\!]} = \sigma(y)_{[\![j,j+|w|-1]\!]} = w$ for some $i, j \in \mathbb{N}$, we have

$$i \in E(y) \iff j \in E(y).$$

Recognizability means that a long enough factor of a fixed point x of σ is either always or never aligned with the image of a letter, so we can 'decode' factors of x using a local rule.

A *partition* of a word $x \in A^{\mathbb{N}}$ is a sequence $P = (w_i)_{i \in \mathbb{N}}$ of words in A^+ such that $x = w_0 w_1 w_2 \cdots$. We denote $w \in P$, if $w = w_i$ for some $i \in \mathbb{N}$. We define $\sigma(P) = (\sigma(w_i))_{i \in \mathbb{N}}$, so that $\sigma(P)$ is a partition of $\sigma(x)$.

3 Aperiodic substitutions cannot produce arbitrarily large powers

In this section, we prove the following theorem.

Theorem 1. *Let σ be aperiodic. Then there exists $N \in \mathbb{N}$ such that $\mathcal{F}(\sigma)$ does not contain the word u^N for any $u \in A^+$.*

Note that Theorem 1 does not directly follow from uniform recurrence:

Example 1. Let $A = \{0, 1\}$. We inductively define $(a_1, b_1) = (0, 1)$ and $(a_n, b_n) = (a_{n-1}^n b_{n-1}, a_{n-1}^{n-1} b_{n-1})$ for all $n > 1$. Let $a_\infty \in A^\omega$ be the limit point of the a_n (it exists and is unique, since each a_n is a prefix of a_{n+1}). Now we claim that a_∞ is aperiodic but uniformly recurrent and contains arbitrarily large powers. The last claim is clear and the second follows from the fact that every a_n occurs with bounded gaps. For the aperiodicity, it can be shown by induction that the only way in which two words $c, d \in \{a_n, b_n\}$ can overlap is when one is a suffix of the other. One can then derive a contradiction from a periodicity hypothesis.

For the rest of this section, fix $l \in \mathbb{N}$ to be $\max_{a \in A} |\sigma(a)|$. Before proving the main theorem, we establish a series of lemmas on how the repeated application of σ affects the lengths and recurrence properties of words in $\mathcal{F}(\sigma)$.

Lemma 1. *There exists $r \in \mathbb{R}$ such that for all $t \in \mathbb{N}$ and $a, b \in A$ we have*

$$\frac{|\sigma^t(a)|}{|\sigma^t(b)|} < r.$$

Proof. Since for all $a, b \in A$, the letter a appears in $\sigma^k(b)$, we have $|\sigma^t(a)| < |\sigma^{t+k}(b)| \leq l^k |\sigma^t(b)|$, so we can choose $r = l^k$. □

Lemma 2. *Let $u, v \in A^+$. Then $\frac{|\sigma^t(w)|}{|w|} < \frac{r|\sigma^t(u)|}{|u|}$.*

Proof. The quantity $q = \frac{|\sigma^t(u)|}{|u|}$ is just the average increase in size in the letters of u when σ^t is applied. By the previous lemma, any letter of w grows by at most a factor r more than any letter of u, so the average growth for letters of w is at most rq when σ^t is applied. □

Lemma 3. *Let $u \in A^+$ and $G, t \in \mathbb{N}$. If u occurs in σ with gap $G|u|$, then $\sigma^t(u)$ occurs with gap $2rG|\sigma^t(u)|$.*

Proof. Let $x \in X(\sigma)$, and let $P = (w_i)_{i\in\mathbb{N}}$ be the partition of x where $|w_i| = G|u|$ for all $i \in \mathbb{N}$. Let $w \in P$ and denote $v = \sigma^t(w)$, and note that $\sigma^t(u)$ is a factor of v. By Lemma 2, we have $\frac{|v|}{|w|} < \frac{r|\sigma^t(u)|}{|u|}$, and thus

$$|v| < \frac{r|w||\sigma^t(u)|}{|u|} = rG|\sigma^t(u)|.$$

But this means that every factor of $\sigma^t(x)$ of length $2rG|\sigma^t(u)|$ must contain a word of $\sigma^t(P)$, and thus the word $\sigma^t(u)$. □

Definition 1. *Let $u \in \mathcal{F}(\sigma)$. A* reappearance of u *is a word $vu \in \mathcal{F}(\sigma)$ such that $uvu \in \mathcal{F}(\sigma)$. In this case uvu is a* complete reappearance of u. *If vu contains exactly one u, then we have a* first reappearance *and a* complete first reappearance*, respectively.*

For $x \in X(\sigma)$, define the partition $P_{u,x} = (w_i)_{i\in\mathbb{N}}$ of x as follows: for all $i \in \mathbb{N}$, w_i is the shortest prefix of $(w_0 w_1 \cdots w_{i-1})^{-1}x$ ending in u. Then all but perhaps the first w_i are first reappearances of u.

Note that the above definition is different from the usual notion of *return to* u, defined as a nonempty word $v \in \mathcal{F}(\sigma)$ such that $vu \in \mathcal{F}(\sigma)$ and u is a prefix of vu.

Lemma 4. *Let $u \in \mathcal{F}(\sigma)$ and $N \in \mathbb{N}$ be such that $u^N \in \mathcal{F}(\sigma)$. There exists $G_{u,N} \in \mathbb{N}$ such that for every word $w \in \mathcal{F}(\sigma)$ of length at least $G_{u,N}|u|$, the word u^N and all complete first reappearances of u are factors of w.*

Proof. By the primitivity of σ, there are finitely many complete first reappearances of u. Namely, if u occurs with gap G, and w is a first reappearance of u with $|w| > G$, then u is a factor of $w_{[\![0,G-1]\!]}$. But since u is also a suffix on w, there are at least two (possibly overlapping) occurrences of u in w, a contradiction. Thus any complete first reappearance of u has length at most G.

There are finitely many words of a given length that appear in σ, and all of them occur with bounded gaps in $X(\sigma)$, so the claim follows. □

The last lemma we need is a classical result in combinatorics of words, proved first in [2].

Lemma 5 (Fine and Wilf). *Let $u, v \in A^+$. If for the infinite words $x = u^\infty$ and $y = v^\infty$, we have $x_{[\![0,|u|+|v|-1]\!]} = y_{[\![0,|u|+|v|-1]\!]}$, then u and v are powers of the same word.*

We are now ready to prove the main theorem:

Proof (Proof of Theorem 1). Assume on the contrary that words of the form u^N appear in $\mathcal{F}(\sigma)$ for every $N \in \mathbb{N}$, and let $x \in A^{\mathbb{N}}$ be an aperiodic σ-periodic word. We can clearly assume that u is primitive, if desired. By taking a power of σ (which does not affect $\mathcal{F}(\sigma)$), we may also assume $\sigma(x) = x$.

Now let $N = 3l$, let u be a word of minimal length such that $u^N \in \mathcal{F}(x)$, and let $G = G_{u,N}$ be given by Lemma 4. Let $M = \max(2rG, N)$. Then there exists a primitive word v such that $v^M \in \mathcal{F}(x)$, and if v is conjugate to $\sigma^t(u)$ for some $t \in \mathbb{N}$, then they are equal. Namely, we can start with $v^{M+1} \in \mathcal{F}(x)$, and then replace v by any of its conjugates if necessary. We also have $|v| \geq |u|$, since u was of minimal length.

Since the length of u increases by at most a factor of l when we apply σ to it, there exists $t \in \mathbb{N}$ such that $|\sigma^t(u)| \leq |v| \leq l|\sigma^t(u)|$.

Let $Q = \sigma^t(P_{u,x})$. All complete first reappearances of u and the word u^N appear in every $w \in \mathcal{F}_{G|u|}(x)$, and thus by Lemma 3, all of their σ^t-images appear in every $w \in \mathcal{F}_{2rG|\sigma^t(u)|}(x)$, and in particular every $w \in \mathcal{F}_{2rG|v|}(x)$. Therefore, $\mathcal{F}(v^M)$ contains $\sigma^t(u)^N$, and $\sigma^t(u)w$ for every word w in Q except perhaps the first.

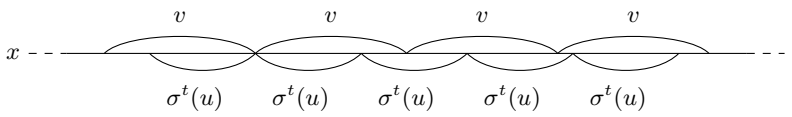

Fig. 1. The word $\sigma^t(u)^N$ occurs in v^M, so v^2 occurs in $\sigma^t(u)^N$.

By the definition of t, we have $3|v| \leq 3l|\sigma^t(u)|$, so v^2 must be a factor of $\sigma^t(u)^N = \sigma^t(u)^{3l}$ (see Figure 1). Then some conjugate u' of $\sigma^t(u)$ has the property that $(u')^\infty$ and v^∞ agree on their first $2|v| \geq |v| + |u'|$ symbols, so u' and v are in fact powers of the same word by Lemma 5. Since v is primitive, this word must be v, and since $|u'| \leq |v|$, we have $u' = v$, and thus $v = \sigma^t(u)$ by our choice of v. In particular, $\sigma^t(u)$ is primitive.

Since σ is aperiodic, there exists a word w in Q other that the first one which is not $\sigma^t(u)$-periodic, and $v^M = \sigma^t(u)^M$ contains the word $\sigma^t(u)w$. But then the primitive subwords $\sigma^t(u)$ at the beginning and end of $\sigma^t(u)w$ must align with the $\sigma^t(u)$-period of v^M, which implies that w is $\sigma^t(u)$-periodic. This contradiction finishes the proof. □

4 An aperiodic uniform substitution is recognizable

We now present an application of Theorem 1 by proving the following result.

Theorem 2. *If σ is aperiodic and uniform, then it is recognizable.*

Note that this is a special case of the result in [3] which characterizes recognizable subtitutions, but we present a new proof using a geometric pigeonhole argument. Let $|\sigma(a)| = m$ for all $a \in A$, and let $x \in A^{\mathbb{N}}$ be a σ-periodic word. It is easy to see that if σ^t is recognizable for some $t \in \mathbb{N}$, then so is σ, so we may assume $\sigma(x) = x$.

Recognizability of uniform substitutions is essentially simpler than the general case, since for all $y \in X(\sigma)$ we have $E(y) = m\mathbb{N}$. The following lemma makes use of this fact.

Lemma 6. *Suppose thatm for all $h \in (0, m)$, there exists $u_h \in A^+$ which is a* 0 mod *m-factor, but not an h* mod *m-factor, of x. Then σ is recognizable.*

Proof. For all 0 mod m-factors w of x, denote

$$I_w = \{(n, h) \mid n \in [0, m), h \in (0, m), u_h \text{ is an } n \bmod m\text{-factor of } w\}.$$

We claim that there exists an $N \in \mathbb{N}$ so large that for all $v \in \mathcal{F}_N(x)$ and $w \in \mathcal{F}(x)$ that are 0 mod m-factors of x, we have $I_w \subset I_v$. Namely, the set of minimal words $u \in \mathcal{F}(x)$ with $u_h \in \sigma(u)$ is finite for each h, and the claim follows from the facts that the $\sigma(u)$ are 0 mod m-factors of x, and each u occurs with bounded gaps.

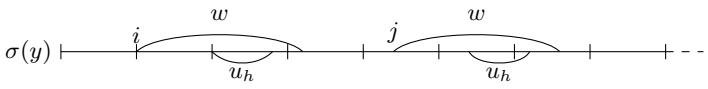

Fig. 2. The word w occurring in $\sigma(y)$. The tick marks denote the coordinates $m\mathbb{N}$.

Let then $y \in X(\sigma)$, $|w| \geq N$, and $\sigma(y)_{[\![i,i+|w|-1]\!]} = \sigma(y)_{[\![j,j+|w|-1]\!]} = w$ for some $i, j \in \mathbb{N}$. Suppose on the contrary that $i \in E(y)$ but $j \notin E(y)$, or equivalently, $i \equiv 0 \bmod m$ but $j \not\equiv 0 \bmod m$. Let $h \in (0, m)$ be such that $h \equiv j \bmod m$. Now we have $(0, h) \in I_w$ (since w starts at i in $\sigma(y)$), and thus u_h is a j mod m-factor of x (since w also starts at j), which implies $(h, h) \in I_w$ by the definition of w. But then, since w starts at i, u_h is an h mod m-factor of $\sigma(y)$, and thus of x, a contradiction. See Figure 2 for a visualization. □

We are now ready to prove Theorem 2.

Proof (Proof of Theorem 2). Suppose on the contrary that σ is not recognizable, so that Lemma 6 gives an $h \in (0, m)$ such that every 0 mod m-factor of x is also an h mod m-factor. Let $w_0 \in \mathcal{F}(x)$ with $|w_0| = 2$ be arbitrary. For all $i \geq 1$, define $w_i = \sigma(w_{i-1})$. Denote by $N(i)$ the number of coordinates n such that $(w_i)_{[\![n,n+m^i-1]\!]} = \sigma^i(a)$ for some $a \in A$.

We now claim that $N(i) \geq i + 1$. For $i = 0$, this is clear, so suppose that $i \geq 1$. By the induction hypothesis, there are at least i coordinates n with

$(w_{i-1})_{[\![n,n+m^{i-1}-1]\!]} = \sigma^{i-1}(a)$ for some $a \in A$. By the definition of σ, for each such n, we have $(w_i)_{[\![mn,mn+m^i-1]\!]} = \sigma^i(a)$. Furthermore, since w_i is the σ-image of a word in $\mathcal{F}(x)$, it is a $0 \bmod m$-factor of x. But then w_i is also an $h \bmod m$-factor, and since $|w_i| = 2m^i$, some $\sigma^i(a)$ is a $(-h) \bmod m$-factor of w_i (see Figure 3). Since the $\sigma^i(a)$ inherited from w_{i-1} are $0 \bmod m$-factors of w_i, we have $N(i) \geq i + 1$.

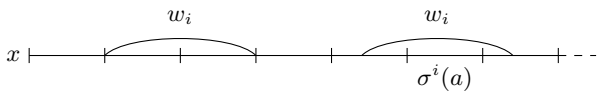

Fig. 3. The word w_i occurring in x. The tick marks denote the coordinates $m^i\mathbb{N}$.

Let now $N \in \mathbb{N}$ be arbitrary. We show that $\mathcal{F}(\sigma)$ contains a word of the form u^N, contradicting Theorem 1. Let $i = N|A| - 1$, and consider the word w_i. Denote

$$I_a = \{i \in [0, m^i - 1] \mid (w_i)_{[\![n,n+m^i-1]\!]} = \sigma^i(a)\}$$

for each $a \in A$. By the above argument, we have $\left|\bigcup_{a\in A} I_a\right| \geq N|A|$, and the pigeonhole principle implies that $|I_a| \geq N$ for some $a \in A$. For this a, there exist $n_1, n_2 \in I_a$ with $0 < k = |n_1 - n_2| \leq \frac{m^i}{N}$. But then $\sigma^i(a)_{[\![0,m^i-k-1]\!]} = \sigma^i(a)_{[\![k,m^i-1]\!]}$, and thus $\sigma^i(a)$ is periodic with period k. Thus $(\sigma^i(a)_{[\![0,k-1]\!]})^N$ is a prefix of $\sigma^i(a)$, and so belongs to $\mathcal{F}(\sigma)$. □

References

1. Fabien Durand. A characterization of substitutive sequences using return words. *Discrete Math.*, 179(1-3):89–101, 1998.
2. N. J. Fine and H. S. Wilf. Uniqueness theorems for periodic functions. *Proc. Amer. Math. Soc.*, 16:109–114, 1965.
3. Brigitte Mossé. Puissances de mots et reconnaissabilité des points fixes d'une substitution. *Theoret. Comput. Sci.*, 99(2):327–334, 1992.