ASYMPTOTIC ORBITS OF PRIMITIVE SUBSTITUTIONS

MARCY BARGE, BEVERLY DIAMOND AND CHARLES HOLTON

Abstract. A primitive, aperiodic substitution on d letters has at most d^2 asymptotic orbits; this bound is sharp. Since asymptotic arc components in tiling spaces associated with substitutions are in 1-1 correspondence with asymptotic words, this provides a bound for those as well.

1. Introduction and Terminology

A *primitive aperiodic substitution* φ on a finite alphabet $\mathcal{A} = \{1, 2, \ldots, d\}$ has an associated minimal substitutive system (W_φ, σ) consisting of the shift map σ on the collection of bi-infinite words W_φ satisfying the following property: a word $w \in W_\varphi$ if and only if for each finite subword w' of w, there are $i \in \mathcal{A}$ and $n \in \mathbb{N}$ such that w' is a subword of $\varphi^n(i)$. One can ask whether there is a bound, in terms of the size of the alphabet $\mathcal{A}$, on the number of orbits asymptotic to another under the shift, that is, *right asymptotic orbits* (or equivalently, under its inverse, in which case we have *left asymptotic orbits*) (see below for more precise definitions).

The investigation of asymptotic orbits in dynamical systems goes back to Gottschalk and Hedlund–it follows from 10.36 of [GH] that at least one pair of right asymptotic orbits and one pair of left asymptotic orbits exist for any infinite minimal substitutive system. Also, the proof of Theorem V. 21 in [Qu] implies that there are only finitely many asymptotic orbits.

For the case in which φ is primitive, aperiodic and *proper* (i.e., φ has only one *φ-periodic word*), the general form of asymptotic orbits is described in [BD]. It follows from this general form that φ has no more than $d^2 - d$ left (or right) asymptotic orbits. The following result provides a bound of d^2 for the total number of asymptotic orbits for any primitive, aperiodic substitution and an improved bound for the case in which the substitution is proper.

Theorem 1. *A primitive, aperiodic substitution φ on d letters has at most d^2 asymptotic orbits. If φ is proper, then φ has at most $4(d-1)$ asymptotic orbits.*

Since asymptotic arc components in tiling spaces associated with substitutions are in 1-1 correspondence with asymptotic words, this provides a bound for those as well.

In §2, we obtain a bound for the number of left asymptotic orbits for the case in which the substitution has no prefix problem. We prove in §3 that this bound also holds for the general case. Combining this with an analogous bound for right asymptotic orbits, we obtain Theorem 1.

We explore briefly the connection between asymptotic orbits and the *complexity* of sequences arising from substitutions in §4.

We introduce terminology necessary for the proof of the theorem. Let $\mathcal{A} = \{1, 2, \ldots, d\}$ be a finite alphabet and $\mathcal{A}^*$ the collection of finite nonempty words formed from the alphabet $\mathcal{A}$. A *substitution* φ is a map $\varphi : \mathcal{A} \to \mathcal{A}^*$; φ extends naturally to a map $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ by concatenation. The map φ has an associated incidence matrix $A_\varphi = A = (a_{ij})_{i,j \in \mathcal{A}}$ in which a_{ij} is the number of occurrences of i in the word $\varphi(j)$; φ is *primitive* if there is n so that for each $i, j \in \mathcal{A}$, j appears in $\varphi^n(i)$ and *has no prefix problems* if for $i \neq j$, $\varphi(i)$ is not a prefix of $\varphi(j)$.

A word w is *allowed for* φ if and only if for each finite subword w' of w, there are $i \in \mathcal{A}$ and $n \in \mathbb{N}$ such that w' is a subword of $\varphi^n(i)$. Let W_φ denote the set of *allowed bi-infinite words* for φ, W_φ^+ the set of *allowed right infinite words* for φ, and W_φ^- the set of *allowed left infinite words* for φ. We identify the 0^{th} coordinate in a bi-infinite word w by either an indexing, as in $w = \ldots w_{-1}w_0w_1\ldots$, or by use of a decimal point (or both). For $w \in W_\varphi$, define the *orbit of* w to be the equivalence class of bi-infinite words

$$[w] = \{w' \in W_\varphi : w' \text{ is a shift of } w\}.$$

The substitution $\varphi : \mathcal{A} \to \mathcal{A}^*$ extends to $\varphi : W_\varphi \to W_\varphi$ where

$$\varphi(\ldots w_{-1}w_0w_1\ldots) = \ldots \varphi(w_{-1}) \;.\; \varphi(w_0)\varphi(w_1)\ldots$$

as well as to a map on equivalence classes

$$\varphi([w]) = [\varphi(w)]$$

which is 1-1 and onto ([Mo]). The word w is *periodic* for φ, or $\varphi-$periodic, if for some $m \in \mathbb{N}$,

$$\varphi^m(w) = \ldots \varphi^m(w_{-1}).\varphi^m(w_0)\varphi^m(w_1)\ldots = \ldots w_{-1}.w_0w_1\ldots.$$

Each primitive substitution φ has at least one allowed φ-periodic bi-infinite word which is necessarily uniformly recurrent under the shift. (For instance, if ij is a subword of $\varphi(k)$ for some i, j, k, then as $n \to \infty$, the finite words $\varphi^n(i).\varphi^n(j)$ converge to a cycle of allowed $\varphi-$periodic, bi-infinite words that are uniformly recurrent under the shift.) A primitive substitution φ is *aperiodic* if at least one (equivalently, each) $\varphi-$periodic bi-infinite word is not periodic under the natural shift map, in which case W_φ (with the shift map) is an infinite minimal dynamical system. If φ is periodic (that is, not aperiodic), then W_φ is finite.

In the remainder of this paper, we assume that the substitution φ is primitive and aperiodic.

Two bi-infinite words $w, w' \in W_\varphi$ are *left asymptotic* provided that there is some $k \in \mathbb{Z}$ so that $w_i = w'_i$ for $i \leq k$. Two orbits $[w]$, $[w']$ are *left asymptotic* if there are $v \in [w]$, $v' \in [w']$ so that v and v' are left asymptotic. A single orbit $[w]$ is a *left asymptotic orbit* if there is $w' \in W_\varphi$ such that $[w] \cap [w'] = \emptyset$

and $[w], [w']$ are left asymptotic orbits. The notion of *right asymptotic* is defined similarly, and *asymptotic* is either left or right asymptotic.

In obtaining the bound on the asymptotic orbits of φ in terms of the size of the alphabet $\mathcal{A}_\varphi$, we work only with the set of left asymptotic orbits; similar arguments apply to the right asymptotic orbits.

Suppose that $[w], [w']$ are left asymptotic orbits. Then $\varphi^n([w])$ and $\varphi^n([w'])$ are left asymptotic orbits for each $n \in \mathbb{N}$. Since there are only finitely many such orbits (Theorem V. 21 of [Qu]), for some $k \geq 0, l \geq 1$, $\varphi^k([w]) = \varphi^{k+l}([w])$ and $\varphi^k([w']) = \varphi^{k+l}([w'])$. If $k \geq 1$, then $\varphi(\varphi^{k-1}([w])) = \varphi(\varphi^{k+l-1}([w]))$ and $\varphi(\varphi^{k-1}[(w'])) = \varphi(\varphi^{k+l-1}([w']))$. But φ is 1-1 on orbits, hence $\varphi^{k-1}([w]) = \varphi^{k+l-1}([w])$ and $\varphi^{k-1}([w]') = \varphi^{k+l-1}([w'])$. That is, $[w]$ and $[w']$ are periodic under φ.

Passing to a power if necessary, we assume the following:

(i) if $w \in W_\varphi^+$ or $w \in W_\varphi^-$ and $\varphi^k(w) = w$ for some $k \geq 1$, then $\varphi(w) = w$;

(ii) if $\varphi(i) = j \ldots$, then $\varphi(j) = j \ldots$;

(iii) if $[w]$ is a left asymptotic orbit of W_φ, then $\varphi([w]) = [w]$.

2. Asymptotic orbits–no prefix problem

Left asymptotic orbits are found most easily when the substitution φ has no prefix problems. Suppose that $v = \ldots v_{-1}.v_0 v_1 \ldots \in [w], v' = \ldots v'_{-1}.v'_0 v'_1 \in [w]$ are such that $v_i = v'_i$ for all $i < 0$ and $v_0 \neq v'_0$. Then $\varphi(v) = \ldots \varphi(v_{-1}).\varphi(v_0)\varphi(v_1) \ldots$ and $\varphi(v') = \ldots \varphi(v'_{-1}).\varphi(v'_0)\varphi(v'_1) \ldots$ agree to the left of the decimal point, and possibly for some coordinates to the right of the decimal point. Suppose that $\varphi(v_0)$ and $\varphi(v'_0)$ differ in their initial letter. Since $\varphi(v) \in [v]$, $\varphi(v') \in [v']$, and $\varphi(v)$ and $\varphi(v')$ agree precisely in the left tails $\ldots \varphi(v_{-1})$ and $\ldots \varphi(v'_{-1})$, $\varphi(v) = v$ and $\varphi(v') = v'$. If $\varphi(v_0)$ and $\varphi(v'_0)$ do not differ in their initial letter, and if φ has no prefix problems, $\varphi(v_0) = uix$ and $\varphi(v'_0) = ujy$, where u is nonempty (x and y may be empty), and $i \neq j$. Since $\varphi(v) \in [v]$, $\varphi(v') \in [v']$, and $\varphi(v)$ and $\varphi(v')$ agree precisely in the left tails $\ldots \varphi(v_{-1})u$ and $\ldots \varphi(v'_{-1})u$,

$$v = \ldots \varphi(v_{-1})u.ix\varphi(v_1) \ldots,$$

hence $i = v_0$, and

$$v' = \ldots \varphi(v'_{-1})u.jy\varphi(v_1) \ldots,$$

hence $j = v'_0$. Applying the same argument to $\varphi^2(v)$ and $\varphi^2(v')$, we see that

$$v = \ldots \varphi^2(v_{-1})\varphi(u)u.v_0 x\varphi(x)\varphi^2(v_1) \ldots$$

and

$$v' = \ldots \varphi^2(v'_{-1})\varphi(u)u.v'_0 y\varphi(y)\varphi^2(v'_1) \ldots.$$

Continuing,

$$v = \ldots \varphi^{n-1}(u) \ldots \varphi(u)u.v_0 x\varphi(x) \ldots \varphi^{n-1}(x) \ldots$$

and

$$v' = \ldots \varphi^{n-1}(u) \ldots \varphi(u)u.v'_0 y\varphi(y) \ldots \varphi^{n-1}(y) \ldots.$$

We summarize with the next lemma.

Lemma 2. *For a primitive aperiodic substitution* φ *with no prefix problems, there are two ways in which left asymptotic orbits* $[w], [w']$ *arise:*

(I) $[w] = [u.v]$, $[w'] = [u.v']$, *with* $u.v, u.v' \in W_\varphi$, $v \neq v'$, *and* $\varphi(u.v) = u.v$, $\varphi(u.v') = u.v'$*; and*

(II) there are $i \neq j \in \mathcal{A}$ *and* $u \in \mathcal{A}^*$ *with* $\varphi(i) = uix$, $\varphi(j) = ujy$, *and*

$$\ldots \varphi^2(u)\varphi(u)u.ix\varphi(x)\varphi^2(x)\ldots \in [w],$$

$$\ldots \varphi^2(u)\varphi(u)u.jy\varphi(y)\varphi^2(y)\ldots \in [w'].$$

(In the case that x *is the empty word,* $[w] = [\ldots \varphi^2(u)\varphi(u)ui.av]$ *for some fixed word* $av \in W_\varphi^+$ *for which* ia *is allowable. A similar statement applies if* y *is the empty word.)*

A left asymptotic orbit may be of both types. Consider the substitution φ given by $\varphi(1) = 12131, \varphi(2) = 12132, \varphi(3) = 31$. The orbit $[w] = [\ldots 12131.12131\ldots]$ is both Type I and Type II.

Let $l = \#\{w \in W_\varphi^- : \varphi(w) = w\}$ and $r = \#\{w \in W_\varphi^+ : \varphi(w) = w\}$. Recall that we may assume that all φ-periodic words are fixed.

Clearly the number of (left asymptotic) orbits of Type I is bounded above by rl. Corollary 5 will then follow immediately from Proposition 4, in which we prove that the number of Type II orbits is bounded above by $2(d - r)$.

Let $\{a : \varphi(b) = a\ldots \text{ for some } b \in \mathcal{A}\} = \{a_1, \ldots, a_r\}$. For $1 \leq j \leq r$, define

$$S(j) = \{(a, u) : a \in \mathcal{A}, u = a_j \ldots \in \mathcal{A}^*, \text{and there is } b \neq a,$$

$$b \in \mathcal{A}, \text{ such that } \varphi(a) = ua\ldots, \varphi(b) = ub\ldots\}.$$

For $1 \leq j \leq r$, let

$$\mathcal{U}(j) = \{u \in \mathcal{A}^* : \text{ for some } a \in \mathcal{A}, (a, u) \in S(j)\},$$

and for each $u \in \mathcal{U}(j)$, let

$$S_u(j) = S_u = \{a \in \mathcal{A} : (a, u) \in S(j)\}.$$

In the following, for words u, v, $u < v$ will denote that u is a prefix of v.

Lemma 3. *For* $k \geq 2$*, there do not exist distinct letters* $b_0, \ldots, b_{k-1}$ *and distinct words* $u_0, \ldots u_{k-1}$ *such that* $b_i, b_{i+1} \in S_{u_i}$ *for* $i = 0, \ldots, k-1$*. (Here* $b_k \equiv b_0$*.)*

Proof: Suppose that there are such b_i, u_i. For some j, $|u_j|$ is minimal. Without loss of generality, $|u_0| \leq |u_j|$ for $0 \leq j \leq k-1$. Now $\varphi(b_1) = u_0b_1\ldots$ and $\varphi(b_1) = u_1b_1\ldots$, which implies that $u_0 < u_1$. If $u_0 \neq u_1$, then $u_0b_1 < u_1$.

Claim: $u_0b_1 < u_i$ for each $i \in 1, \ldots, k-1$.

Proof of claim: Inductively suppose that $u_0b_1 < u_{i-1}$. Now $\varphi(b_i) = u_{i-1}b_i\ldots$ and $\varphi(b_i) = u_ib_i\ldots$, which implies that either $u_{i-1} < u_i$ or $u_i < u_{i-1}$. If the former holds, then $u_0b_1 < u_i$ as desired. If $u_i < u_{i-1}$, then $u_0b_1 < u_i$, as we want, or $u_i < u_0b_1$. In the latter case, $|u_i| = |u_0|$, by the minimality of $|u_0|$, so $u_0 = u_i$, contrary to assumption.

So $u_{k-1} = u_0b_1\dots$ and $\varphi(b_{k-1}) = u_{k-1}b_{k-1}\dots$. Now $\varphi(b_0) = \varphi(b_k) = u_{k-1}b_k\dots = u_0b_1\dots b_k\dots$. Also, $b_0 \in S_{u_0}$, so $\varphi(b_0) = u_0b_0\dots$. But $b_0 \neq b_1$, and the lemma is proved. $\Box$.

Proposition 4. *A primitive, aperiodic substitution with no prefix problem has no more than* $2d - 2r$ *left asymptotic orbits of Type II.*

Proof: Construct a graph G as follows. The vertices of G are the letters of $\mathcal{A}$, and for $a, b \in \mathcal{A}$, $a \neq b$, there is an edge between a and b if and only if $a, b \in S_u(j)$ for some $u \in \mathcal{A}^*$, $j \in \mathcal{A}$. It follows from Lemma 3 that if $j \leq r$ and $u, v \in \mathcal{U}(j)$ with $u \neq v$, then $|S_u(j) \cap S_v(j)| \leq 1$. That is, there is an edge between a, b if and only if a, b are associated with precisely one pair of Type II left asymptotic orbits. Note that r is a lower bound for the number of components of G, since $S_u(j) \cap S_v(i) = \emptyset$ if $i \neq j$.

It is clear from the definition of G that for $u \in \mathcal{U}(j)$, the subgraph of G on vertices $S_u(j)$ is a nontrivial complete graph. Also, it follows from Lemma 3 that any nontrivial complete subgraph of G has vertices contained in $S_u(j)$ for some $u \in \mathcal{A}^*$, $j \in \mathcal{A}$. Let L denote the cardinality of the set of left asymptotic orbits of Type II. According to the above comments, $L = \Sigma_{i=1}^{m}|G_i|$, where $\{G_1, G_2, \dots, G_m\}$ is a listing of the nontrivial maximal complete subgraphs of $G_{\mathcal{A}}$ and $|G|$ equals the number of vertices of G.

We prove inductively that for any k element subset $\mathcal{A}'$ of $\mathcal{A}$, the subgraph of G on $\mathcal{A}'$, $G_{\mathcal{A}'}$, has the property that if $\{G_1, G_2, \dots, G_{m'}\}$ is a listing of the nontrivial maximal complete subgraphs of $G_{\mathcal{A}'}$, then $\Sigma_{i=1}^{m'}|G_i| \leq 2|\mathcal{A}'| - 2c$, where c is the number of components of $G_{\mathcal{A}'}$. The proposition then follows.

This clearly holds for $|\mathcal{A}'| = 2$. Suppose that this holds for any subset of $\mathcal{A}$ of cardinality n', where $2 \leq n' \leq k - 1$, and that $\mathcal{A}' \subseteq \mathcal{A}$ with $|\mathcal{A}'| = k$. Without loss of generality, $G_{\mathcal{A}'}$ is connected. Choose $v \in \mathcal{A}'$, and consider $G_{\mathcal{A}'\setminus\{v\}}$. Let c' denote the number of components of $G_{\mathcal{A}'\setminus\{v\}}$.

It follows from Lemma 3 and comments above that v is contained in exactly c' nontrivial maximal complete subgraphs of $G_{\mathcal{A}'}$. Let $G_1, G_2, \dots, G_m$ be a listing of all nontrivial maximal complete subgraphs of $G_{\mathcal{A}'}$, indexed so that for $1 \leq i \leq c' \leq m$, v is a vertex of G_i. There is a nonnegative integer $c'' \leq c'$ and an indexing of $\{G_i\}_{i \leq m}$ so that in addition, $1 \leq i \leq c''$ if and only if $|G_i \cap G_{\mathcal{A}'\setminus\{v\}}| = 1$.

For $1 \leq i \leq c''$, $|G_i| = 2$. For $c'' + 1 \leq i \leq c'$,

$$|G_i| = |G_i \cap G_{\mathcal{A}'\setminus\{v\}}| + 1,$$

and for $c' \leq i \leq m$, $|G_i| = |G_i \cap G_{\mathcal{A}'\setminus\{v\}}|$. Finally, the nontrivial maximal complete subgraphs of $G_{\mathcal{A}'\setminus\{v\}}$ are exactly

$$\{G_i \cap G_{\mathcal{A}'\setminus\{v\}} : c'' + 1 \leq i \leq m\}.$$

Then

$$\begin{aligned}
\Sigma_{i=1}^{m}|G_i| &= \Sigma_{i=1}^{c''}|G_i| + \Sigma_{i=c''+1}^{c'}|G_i| + \Sigma_{i=c'+1}^{m}|G_i| \\
&\leq 2c'' + (c' - c'') + 2(k-1) - 2c' \\
&= 2(k-1) - c' + c'' \\
&\leq 2k - 2 \\
&= 2|G_{\mathcal{A}'}| - 2(1) \\
&= 2|G_{\mathcal{A}'}| - 2(\#\text{ of components of } G_{\mathcal{A}'}). \quad \square
\end{aligned}$$

Corollary 5. *If φ is a primitive aperiodic substitution on d letters with no prefix problems, then the number of left asymptotic orbits of φ is bounded above by $2(d-r)+rl$.*

Before completing the proof in the general case, we provide an example to show that the bound can be attained.

Example. Let φ be the substitution $\varphi(1) = 112131$, $\varphi(2) = 221232$, $\varphi(3) = 331323$. Then φ has $3^2 = 9$ fixed words of the form $\dots i.j \dots$ for each $i, j \in \{1, 2, 3\}$, each of which is left asymptotic (and right asymptotic) to other fixed words.

It is easy to see that this construction can be generalized to d letters so that for each $i \in \mathcal{A}$, $\varphi(i) = i \dots i$ and for each pair $i, j \in \mathcal{A}_\varphi$, ij is allowed. Then φ has exactly d^2 right asymptotic composants and d^2 left asymptotic composants.

3. Asymptotic orbits-the general case

We now show that the bound of Corollary 5 also holds in the general case.

Two substitutions φ and φ' are *strong shift equivalent* if there are morphisms ρ_i, ψ_i, $i = 1, \dots, k$ so that $\varphi = \rho_1 \circ \psi_1$, $\psi_1 \circ \rho_1 = \rho_2 \circ \psi_2, \dots, \psi_j \circ \rho_j = \rho_{j+1} \circ \psi_{j+1}, \dots, \psi_k \circ \rho_k = \varphi'$.

Lemma 6. *If φ and φ' are strong shift equivalent, then W_φ and $W_{\varphi'}$ have the same number of asymptotic orbits. Also, φ and φ' have the same numbers of fixed (or periodic) left and right infinite words. If φ is primitive and aperiodic, so is φ'.*

Proof: Suppose that $\varphi = \rho \circ \psi$ and $\varphi' = \psi \circ \rho$. Then the functions $[w] \to [\psi(w)]$ and $[v] \to [\rho(v)]$ take orbits of W_φ to orbits of $W_{\varphi'}$ and orbits of $W_{\varphi'}$ to orbits of W_φ, respectively. Moreover, since $[w] \to [\rho \circ \psi(w)] = [\varphi(w)]$ and $[v] \to [\psi \circ \rho(v)] = [\varphi'(v)]$ are bijections on orbits, so also are $[w] \to [\psi(w)]$ and $[v] \to [\rho(v)]$. Since a periodic substitution has only a finite number of bi-infinite words, while an aperiodic substitution has infinitely many, φ is aperiodic if and only if φ' is aperiodic. Clearly, if $w = \dots w_{-1}.w_0 w_1 \dots$ and $u = \dots u_{-1}.u_0 u_1 \dots$ are asymptotic words, then $\eta(w) = \dots \eta(w_{-1}).\eta(w_0)\eta(w_1) \dots$ and $\eta(u) = \dots \eta(u_{-1}).\eta(u_0)\eta(u_1) \dots$ are asymptotic words for any morphism η. It is easily checked that if $\dots w_{-2}w_{-1}.$ is fixed (or periodic) for φ, then so is $\dots\psi(w_{-2})\psi(w_{-1}).$ for φ', etc. □

The substitution φ is *1-1 on letters* if $\varphi(j) \neq \varphi(k)$ for any $j \neq k$.

Lemma 7. *If φ is any primitive aperiodic substitution, then there is a primitive aperiodic substitution $\tilde{\varphi}$ such that $\tilde{\varphi}$ is 1-1 on letters, the alphabet of $\tilde{\varphi}$ is no larger than that of φ, and $\tilde{\varphi}$ is strong shift equivalent to φ.*

Proof: If φ is 1-1 on letters, let $\tilde{\varphi} = \varphi$. Otherwise, let $\varphi(i_1), \dots, \varphi(i_k)$ be the distinct words from the list $\varphi(1), \dots, \varphi(d)$, where $d = |\mathcal{A}_\varphi|$. Let $\eta : \{1, \dots, k\} \to \{1, \dots, d\}^*$ and $\gamma : \{1, \dots, d\} \to \{1, \dots, k\}^*$ be the morphisms given by $\eta(j) = i_j$ and $\gamma(l) = j$ provided $\varphi(l) = \varphi(i_j)$. Then $\varphi = (\varphi \circ \eta) \circ \gamma$. Let $\varphi^{(1)} = \gamma \circ (\varphi \circ \eta)$. Then $\varphi^{(1)}$ is a substitution on $k < d$ letters. If $\varphi^{(1)}$ is 1-1 on letters, let $\tilde{\varphi} = \varphi^{(1)}$. Otherwise continue, creating $\varphi^{(2)} = (\varphi^{(1)})^{(1)}, \dots$. For some $m < d$, $\varphi^{(m)}$ is 1-1 on letters; let $\tilde{\varphi} = \varphi^{(m)}$. Since $\tilde{\varphi}$ and φ are strong shift equivalent, so are their incidence matrices, and it follows that $\tilde{\varphi}$ is also primitive (see, for example, [LM]). □

Given $1 \leq j \neq k \leq n$, let σ_{jk} denote the substitution:

$$i \to i, \quad i \neq k,$$
$$k \to jk.$$

We will refer to σ_{jk} as an *elementary substitution.*

Lemma 8. *If φ is a primitive aperiodic substitution on d letters, then either*

(i) there is a primitive aperiodic substitution φ' on $k \leq d$ letters that has no prefix problem and is strong shift equivalent to φ, or

(ii) there are primitive aperiodic substitutions ψ and σ on $k \leq d$ letters so that ψ is strong shift equivalent to φ, σ is a composition of elementary substitutions, ψ and σ have the same allowed words (that is, $W_\psi = W_\sigma$ as sets), and ψ and σ have the same number of fixed left infinite words.

In case (i), we say the the prefix problem for φ *is solved.*

Proof: We attempt to eliminate all prefix problems of φ by a sequence of strong shift equivalences. Success will give us conclusion (i), failure conclusion (ii). Passing to a power if necessary, we assume that any left or right infinite word which is periodic for φ is fixed.

If φ has no prefix problems, let $\varphi' = \varphi$. Otherwise, suppose $\varphi(j) = p$ and $\varphi(k) = pu$ for some nonempty word p and $j \neq k$. If u is empty, let $\varphi_1 = \tilde{\varphi}$ as in Lemma 7. Otherwise, let τ_1 be the substitution

$$\tau_1(i) = \varphi(i), \quad i \neq k$$
$$\tau_1(k) = u$$

and let σ_1 be the elementary substitution

$$\sigma_1 = \sigma_{jk}.$$

Then $\varphi = \tau_1 \circ \sigma_1$; let $\varphi_1 = \sigma_1 \circ \tau_1$. If φ_1 has no prefix problems, let $\varphi' = \varphi_1$. Otherwise, let φ_2 be either $\tilde{\varphi}_1$ or $\sigma_2 \circ \tau_2$ with $\tau_2 \circ \sigma_2 = \varphi_1$, as above. Continuing in this way, we either arrive at φ_k with no prefix problems, in which case $\varphi' = \varphi_k$, or we generate an infinite sequence of substitutions $\{\varphi_k\}_{k\in\mathbb{N}}$ with prefix problems. In the latter case, there must be an N so that the size of the alphabet of φ_k is the same for all $k \geq N$. Then the incidence matrices A_{φ_k} are all of the same size for $k \geq N$ and in the same strong shift equivalence class. It follows that the Perron-Frobenius eigenvalues of the matrices A_{φ_k}, for $k \geq N$, are equal (see [LM]). Since there are only finitely many non-negative integer matrices of a given size possessing a given dominant eigenvalue, the collection of matrices $\{A_{\varphi_k}\}_{k\geq N}$ is finite, and so also is the collection of substitutions $\{\varphi_k\}_{k\geq\mathbb{N}}$. There must then be $K \geq N$ and $k \geq 1$ so that $\varphi_{K+k} = \varphi_K$. Since $\varphi_{K+1} = \sigma_{K+1} \circ \tau_{K+1}$ where $\varphi_K = \tau_{K+1} \circ \sigma_{K+1}$, etc., we have

$$\begin{aligned}
&\sigma_{K+k} \circ \ldots \circ \sigma_{K+1} \circ \tau_{K+1} \circ \ldots \circ \tau_{K+k} \\
&= \varphi_{K+k}^k = \varphi_K^k \\
&= \tau_{K+1} \circ \ldots \circ \tau_{K+k} \circ \sigma_{K+k} \circ \ldots \circ \sigma_{K+1}.
\end{aligned}$$

Let $\sigma = \sigma_{K+k} \circ \ldots \circ \sigma_{K+1}$ and $\tau = \tau_{K+1} \circ \ldots \circ \tau_{K+k}$ so that $\varphi_K^k = \sigma \circ \tau = \tau \circ \sigma$, and let $\psi = \varphi_K$. We show that σ is primitive and that ψ and σ have the same set of allowed words as well as the same fixed left infinite words.

Since ψ is strong shift equivalent with φ, ψ is primitive and any ψ-periodic word is fixed. Also, the set of allowed words for ψ is the same as that for ψ^l for all $l \geq 1$.

Since σ is a composition of elementary substitutions, $\sigma(i) = \ldots i$ for each $i \in \mathcal{A}_\psi$. Let $i \in \mathcal{A}_\psi$ be such that $\psi(i) = \ldots i$ (this occurs for at least one i), and let X denote the ψ-fixed left infinite word $X = lim_{n\to\infty}\psi^n(i)$. Since $(\sigma \circ (\psi^k)^n)(i) = ((\psi^k)^n \circ \sigma)(i) = (\psi^k)^n(\ldots i) = \ldots (\psi^k)^n(i)$ for all n, σ also fixes X, hence X is allowed for σ. Any fixed (or periodic) word for ψ is of the above form, hence left infinite words fixed for ψ are fixed for σ.

Since ψ^k is primitive, and $|\sigma(i)| > 1$ for some i, $|(\psi^k)^n \circ \sigma(j)| = |\sigma \circ (\psi^k)^n(j)| > |(\psi^k)^n(j)|$ for large enough n, hence $|\sigma(j)| > 1$ for all j. Were σ not primitive, then for some i, j, X would contain subwords of the form $\sigma^n(i)$ with no occurrences of j. But X is uniformly recurrent while $|\sigma^n(i)| \to \infty$, so this is not possible.

It follows that X completely determines the allowed bi-infinite words for σ and ψ. In particular, the sets of allowed words W_ψ and W_σ are identical.

It remains to show that if X is a σ-fixed left infinite word, X is also ψ-fixed. For such a word X, extend X to a σ-periodic bi-infinite word Y with period m. Then $\sigma^m(\psi^k(Y)) = \psi^k(\sigma^m(Y)) = \psi^k(Y)$. That is, $\psi^k(Y)$ is also periodic for σ, as is $\psi^{nk}(Y)$ for each $n \in \mathbb{N}$. Since σ has only finitely many periodic words, there is a smallest $n_1 > 0$ for which there is $0 \leq n_2 < n_1$ for which $\psi^{n_1 k}(Y) = \psi^{n_2 k}(Y)$. If $n_2 \neq 0$, then $\psi^k(\psi^{(n_1-1)k}(Y)) = \psi^k(\psi^{(n_2-1)k}(Y))$. The fact that ψ is 1-1 on bi-infinite words implies that $\psi^{(n_1-1)k}(Y) = \psi^{(n_2-1)k}(Y)$, contradicting the definition of n_1 and n_2. That is, $n_2 = 0$, and Y is periodic for ψ^k. But any word periodic for ψ^k is periodic, hence fixed, for ψ, so $\psi(X) = X$. $\Box$

Proposition 9. *If φ is a primitive aperiodic substitution on d letters, then the number of left asymptotic orbits of φ is bounded above by $2(d-r)+rl$.*

Proof: Let φ be a primitive aperiodic substitution on d letters. If φ has no prefix problem, then the theorem follows from Proposition 5. If φ has a prefix problem, then by Lemma 8, either (i) the prefix problem can be solved for φ by φ' on $k \leq d$ letters, and the result follows from Proposition 5 and Lemma 6, or (ii), there are primitive aperiodic substitutions ψ and σ on $k \leq d$ letters so that ψ is strong shift equivalent to φ, σ is a composition of elementary substitutions, ψ and σ have the same allowed words, and ψ and σ have the same number of fixed left infinite words. According to Lemma 6, φ and ψ have the same number of left asymptotic orbits, as do ψ and σ.

It then suffices to show that if σ is a primitive aperiodic substitution on d letters that is a composition of elementary substitutions, σ has no more than d^2 left asymptotic orbits. We argue by induction. If σ is a substitution on 1 letter, σ is periodic and the conclusion holds vacuously. Suppose that the conclusion holds for primitive aperiodic substitutions on fewer than d letters that are a composition of elementary substitutions, and let σ be primitive, aperiodic, on d letters, and a composition of elementary substitutions. For some m_0, σ^{m_0} has the property that any σ^{m_0}-periodic infinite word is fixed. Write $\sigma^{m_0} = \sigma_{jk}\sigma'$, where σ' is also a composition of elementary substitutions. Then the two letter word kk is not allowed for σ^{m_0}. We will 'rewrite' σ^{m_0} (in the spirit of [Dur]) using the $d-1$ 'stopping rules' $\{i : i \neq k, 1 \leq i \leq d\}$. That is, consider the alphabet

$$\mathcal{A} = \{i : i \neq k, 1 \leq i \leq d\} \cup \{ki : ki \text{ is allowed for } \sigma^{m_0}\};$$

let $\mathcal{A} = \{u_1, \ldots, u_m\}$ where $m = |\mathcal{A}|$. Each allowed word for σ^{m_0} can be factored uniquely as a concatenation of elements of $\mathcal{A}$. Let $\hat{\sigma} : \{1, \ldots, m\} \to (\{1, \ldots, m\})^*$ be the substitution defined by $\hat{\sigma}(i) = i_1 \ldots i_l$ provided $\sigma^{m_0}(u_i) = u_{i_1} \ldots u_{i_l}$. It is easy to see that $\hat{\sigma}$ is also primitive and aperiodic, and that the asymptotic orbits of $\hat{\sigma}$ are in 1-1 correspondence with those of σ^{m_0} and σ.

Claim: $\hat{\sigma}$ has exactly $d-1$ fixed left infinite words and at most $d-1$ fixed right infinite words (and no other one-sided infinite words periodic under $\hat{\sigma}$).

Proof of claim: If $w = \ldots w_{-1}.$ is fixed by $\hat{\sigma}$, then the word $u_{w_{-1}} \in \mathcal{A}$ is either a singleton i, where $i \neq k$, or is of the form ki, where $i \neq k$. Since $\sigma^{m_0}(i)$ has length greater than 1 (σ is primitive), the fixed left infinite words for $\hat{\sigma}$ are in 1-1 correspondence with those for σ (and σ^{m_0}) of the form $\ldots i.$, where $i \neq k$. If $w = .w_0w_1\ldots$ is a periodic right infinite word for $\hat{\sigma}$, then $.u_{w_0}u_{w_1}\ldots$ is a periodic, hence fixed, right infinite word for σ^{m_0}, of which there are at most $d-1$.

Applying Lemma 8 to $\hat{\sigma}$, either (i) there is a substitution χ that is strong shift equivalent to $\hat{\sigma}$, primitive, aperiodic, and has no prefix problems, or (ii) conclusion (ii) holds. In case (i), the size m of the alphabet of χ is bounded above by $2d-2$, and χ has exactly $d-1$ fixed left infinite words, r fixed right infinite words, where $0<r<d$, and no other periodic infinite words (Lemma 6). According to Proposition 5, the number of left asymptotic orbits of χ, and hence of φ, is bounded above by

$$2m+r(l-2)\leq 2(2d-2)+(d-1)(d-3)=d^2-1$$

and the theorem is proved. In case (ii), we obtain a composition of elementary substitutions σ' with the same number of asymptotic orbits as $\hat{\sigma}$, and hence φ, and with exactly $d-1$ fixed left infinite words. Since a composition of elementary substitutions has exactly as many fixed left infinite words as the size of its alphabet, σ' is on $d-1$ letters and, by the inductive assumption, the result is proved. □

Note that at any stage of the process described in the proof above, asymptotic orbits found for a substitution resulting from either a step in solving the prefix problem or a rewriting of a cycle of elementary substitutions can be carried back via the appropriate morphisms to asymptotic orbits for the original substitution. For an example of finding asymptotic orbits for a substitution with a suffix problem by solving the suffix problem and carrying back asymptotic words from the resulting system, see Example 3.17 of [BD].

We complete the proof of the main theorem.

Proof of Theorem 1: By arguments analogous to the above, a substitution φ on d letters has at most $2(d-l)$ right asymptotic orbits of Type II while the total number of (left or right) asymptotic orbits of Type I is bounded above by rl. That is, the total number of asymptotic orbits is bounded above by $rl+2(d-r)+2(d-l)$. It is easy to show that the maximum value of this expression occurs at $r=l=d$.

If φ is proper, then all asymptotic orbits are of Type II, and the bound is of the form $4(d-1)$. □

4. COMPLEXITY

Various authors have considered the notion of *complexity* for several classes of sequences. For an infinite word w, the *complexity function* for w is defined as: $p_w(n)=p(n)=\#$ distinct subwords of w of length n. It is known that for sequences arising from primitive substitutions, the complexity is sublinear.

Suppose that φ is a primitive aperiodic substitution, and let $S=\{u_i\}_{i=1}^k$, the set of *stems*, denote the finite collection of allowed left infinite words *generating left asymptotic orbits*. That is, for $u\in S$, there are at least two right infinite words v,v' with uv and uv' both allowed and v,v' differing in their first letter. We say that the left asymptotic orbits $[uv],[uv']$ are *equivalent*; let E denote the number of equivalence classes.

Proposition 10.

$$\liminf_{n\to\infty}(p(n+1)-p(n)) \geq \#\ \textit{left asymptotic orbits} - E$$

Remarks: Since the above inequality must also hold for right asymptotic orbits, and one can have different numbers of left and right asymptotic orbits, equality for both the left and right cases need not hold. The reader can verify that for $\varphi(1) = 11121221$, $\varphi(2) = 12$, equality does not hold for either the left or right case. On the other hand, for the Morse-Thue sequence obtained as a fixed point for the substitution $\varphi(1) = 12$, $\varphi(2) = 21$, and for the example given following Lemma 2, equality holds. (See Chaper 5 of [BFMS] for a detailed discussion of the complexity of the Morse-Thue sequence.) We do not fully understand the relationship between complexity and asymptotic orbits.

Proof of Proposition 10: For each stem $u \in S$, let $br(U)$ denote the cardinality of a maximal collection of right infinite words $\{v_i : uv_i$ is allowed and the first letters of v_i, v_j differ for $i \neq j\}$. For each stem u, there is a contribution of $br(u) - 1$ to $p(n+1) - p(n)$, and these contributions are distinct for different stems for sufficiently large n. Thus $p(n+1) - p(n) \geq \Sigma_{u\in S}(br(u) - 1)$ for large n. On the other hand, $(\Sigma_{u\in S} br(u)) - |S| = \#$ left asymptotic orbits $-E$. $\square$

References

[BD] M. Barge and B. Diamond, A complete invariant for the topology of one-dimensional substitution tiling spaces, *Ergodic Theory and Dynamical Systems* **21** (2001), 1333-1358.

[BFMS] V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel (editors), *Introduction to Finite Automata and Substitution Dynamical Systems*, to appear in Lectures Notes in Mathematics, Springer-Verlag, http://iml.univ-mrs.fr/editions/preprint00/book/prebookdac.html.

[Dur] F. Durand, A characterization of substitutive sequences using return words, *Discrete Math* **179**(1998), 89–101.

[GH] W. Gottschalk and G. A. Hedlund, *Topological Dynamics*, American Mathematical Society Colloquium Publications, Vol. 36. American Mathematical Society, Providence, R. I., 1955.

[LM] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, Cambridge, 1995.

[Mo] B. Mossé, Puissances de mots et reconnaissabilité des points fixes d'une substitution, *Theoretical Computer Science* **99** (1992), 327–334.

[Qu] M. Queffélec, *Substitution Dynamical Systems- Spectral Analysis*, Springer-Verlag, Berlin, 1987.

Department of Mathematics, Montana State University, Bozeman, MT 59717
barge@math.montana.edu

Department of Mathematics, College of Charleston, Charleston, SC 29424
diamondb@cofc.edu

Department of Mathematics, University of California, Berkeley, CA, 94720-3840 cholton@math.berkeley.edu