Languages under substitutions and balanced words

par Alex HEINIS

Résumé. Cet article est constitué de trois parties. Dans la première on prouve un théorème général sur l'image d'un language K sous une subsitution. Dans la seconde on applique ce théorème au cas spécial prenant pour K le language des mots balancés et la troisième partie concerne les mots bi-infinis récurrents de croissance de complexité minimale ("minimal block growth").

Abstract. This paper consists of three parts. In the first part we prove a general theorem on the image of a language K under a substitution, in the second we apply this to the special case when K is the language of balanced words and in the third part we deal with recurrent **Z**-words of minimal block growth.

Definitions and notation

In this paper a word is a mapping $w : I \to \{a, b\}$ where I is a subinterval of **Z**. We identify words which are shifts of eachother: hence we identify $w_1 : I_1 \to \{a, b\}, w_2 : I_2 \to \{a, b\}$ if there exists an integer k with $I_1 + k = I_2$ and $w_1(i) = w_2(i + k)$ for all $i \in I_1$. If $I = \mathbf{Z}$ we call w a **Z**-word or a bi-infinite word.

If I is finite we call x finite and its length $|x|$ is defined as $|I|$. The usual notation for a finite word of length n is $x = x_1 \cdots x_n$ where all $x_i \in \{a, b\}$. We write $\{a, b\}^*$ for the collection of finite words and $\{a, b\}^+$ for the non-empty finite words. (The empty word will be denoted throughout by ϵ). The concatenation xy of two words x, y is defined by writing x in front of y, which is only defined under the obvious restrictions. A finite word x is called a factor of w, notation $x \subset w$, if $w = yxz$ for some words y, z. It is called a left-factor (prefix) of w if $w = xy$ for some y and a right-factor (suffix) if $w = yx$ for some y. The factor-set of w will be denoted by $F(w)$.

An n-factor of w is a factor of w of length n. The collection of n-factors is defined by $\mathcal{B}(w, n)$ and we write $P(w, n) := |\mathcal{B}(w, n)|$. The mapping $P(w, n) : \mathbf{N} \to \mathbf{N}$ is known as the complexity function of w. A factor $x \subset w$ has multiple right extension (m.r.e.) in w if $xa, xb \subset w$. We also say that x

___

is right-special in w and the collection of right-special factors is denoted by $\mathrm{MRE}(w)$. We write $\mathrm{MRE}_n(w)$ for the collection of right-special n-factors and similarly we define $\mathrm{MLE}(w)$ and $\mathrm{MLE}_n(w)$. A factor $x \in \mathrm{MLE}(w) \cap \mathrm{MRE}(w)$ is called a bispecial factor. The bispecials are usually divided into three classes. We call a factor x weakly/normally/strictly bispecial if the number of symbols σ for which $\sigma x \in \mathrm{MRE}(w)$ equals $0/1/2$, respectively. We follow the notation in [Ca] and write $\mathrm{BF}(w)/\mathrm{BO}(w)/\mathrm{BS}(w)$ for the collection of weakly/normally/ strictly bispecials in w, respectively.

A language K is just a collection of words. Most of the notions above can be generalised without problems to languages, e.g. $F(K) := \cup_{w \in K} F(w)$ and $\mathrm{MRE}(K) = \{x | xa, xb \in F(K)\}$. All languages we consider in this paper are closed under factors and extendible. This means $x \subset y, y \in K \Rightarrow x \in K$ and that $\{ax, bx\} \cap K, \{xa, xb\} \cap K$ are non-empty for any finite $x \in K$. We call this a CE-language. The following fundamental proposition can also be found in [Ca].

Proposition 1. *Let K be a CE-language with $a, b \in K$. Then $P(K, n) = n + 1 + \sum_{i=0}^{n-1} (n - 1 - i)(|\mathrm{BS}_i(K)| - |\mathrm{BF}_i(K)|)$.*

Proof. We have $P(K, n+1) - P(K, n) = |\mathrm{MRE}_n(K)|$ and $|\mathrm{MRE}_{n+1}(K)| - |\mathrm{MRE}_n(K)| = |\mathrm{BS}_n(K)| - |\mathrm{BF}_n(K)|$ for all $n \geq 0$. Now perform a double summation. $\square$

Hence: to know $P(K, n)$ we need only consider the weakly and strongly bispecial factors of K. A substitution is a mapping $T : \{a, b\}^* \to \{a, b\}^*$ satisfying $T(xy) = T(x)T(y)$ for all finite x, y. Obviously T is determined by $X := Ta, Y := Tb$ and we write $T = (X, Y)$. Substitutions extend in a natural way to infinite words and we will not distinguish between T and this extension.

If $T = (AB, AC), T' = (BA, CA)$ and σ is a **Z**-word then $T\sigma = T'(\sigma)$. It follows that $F(TK) = F(T'K)$ for any CE-language K. Now consider a substitution $T = (X, Y)$. Let $x = X^{+\infty} := XXX \cdots, y := Y^{+\infty}$ (right-infinite words) and write $|X| =: m, |Y| =: n$. If $x = y$ then one can show, for instance with the Defect Theorem [L, Thm 1.2.5], that X, Y are powers of one word π, i.e. $X = \pi^k, Y = \pi^l$, and then we have $T\sigma = \pi^\infty$ for any σ. In this case we call T trivial. If T is not trivial there exists a smallest i with $x_i \neq y_i$ and then $T\sigma = T'(\sigma)$ where $T' := (x_i \cdots x_{i+m-1}, y_i \cdots y_{i+n-1})$. For non-trivial T and K a CE-language we have therefore shown that $F(TK) = F(T'K)$ for some T' where $T'a, T'b$ have different initial symbols. From now on we assume that $T = (X, Y)$ is a substitution with $X_1 = a, Y_1 = b$. We call such T an a/b-substitution or ABS.

1. A general theorem

In the next theorem we write $\{X, Y\}^{-\infty}$ for the image under T of all left-infinite words, hence $\{X, Y\}^{-\infty} := T(\{a, b\}^{-\infty})$.

Theorem 1. *Let $T = (X, Y)$ be an ABS, let K be a CE-language and $L := F(TK)$. Let $\Sigma := \{X, Y\}^{-\infty}$ and let S be the greatest common suffix of elements of Σ. Then there exists a constant M such that for $x \in L$ with $|x| \geq M$ we have*

$$x \in \mathrm{BS}(L) \iff x = ST(\xi), \xi \in \mathrm{BS}(K) \text{ and}$$
$$x \in \mathrm{BF}(L) \iff x = ST(\xi), \xi \in \mathrm{BF}(K).$$

Proof. Since $X^{-\infty} \neq Y^{-\infty}$ as above we find that S is finite and well-defined. From T we construct a graph $G(T)$ as follows. We have a point $O \in G(T)$, the origin, such that $G(T)$ consists of two directed cycles α, β from O to itself such that they only intersect in O. Every edge has a label in $\{a, b\}$ in such a way that the labels from α, β (starting at O) read X, Y, respectively. As an example we have drawn $G(T)$ when $T = (abb, bba)$.

G(T)
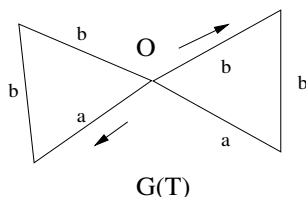
We call $G(T)$ the representing graph for T. We let Δ be the collection of finite paths in $G(T)$ and define $\chi : \{a, b\}^* \to \Delta$ by $\chi(a) = \alpha, \chi(b) = \beta$ such that χ respects concatenation. We define the label-map $\lambda : \Delta \to \{a, b\}^*$ in the obvious way and set Γ equal to the collection of subpaths of $\chi(FK)$. Then $L = \lambda(\Gamma)$. We let σ be the symbol such that all words from ΣX have suffix σS and then all words from ΣY have suffix $\overline{\sigma} S$. (In this paper we write $\overline{a} = b, \overline{b} = a$). If x is a finite word we write $\Gamma(x) := \{\gamma \in \Gamma | \lambda(\gamma) = x\}$ and its elements are called representing paths for x (in Γ). For ($\Rightarrow$) we assume that $x \in L$ is bispecial and we consider two cases.

α) All $\gamma \in \Gamma(x)$ have the same final vertex. This vertex is then O since $x \in \mathrm{MRE}(L)$. Since $x \in \mathrm{MLE}(L)$ we have $|x| \geq |S|$ and x has suffix S. If $|x| > |S|$ and x has suffix $\sigma S / \overline{\sigma} S$, then all $\gamma \in \Gamma(x)$ end with an edge in α/β. Since $x \in \mathrm{MLE}(L)$ we even have that every $\gamma \in \Gamma(x)$ ends with α/β. Continuing in this way one finds that every $\gamma \in \Gamma(x)$ is of the form $\gamma = \phi\chi(\xi)$ where ϕ is a path of length $\leq |S|$ ending in O and where ξ depends only on x (not on γ). Using $x \in \mathrm{MLE}(L)$ we find $|\phi| = |S|, \lambda(\phi) = S$ and $x = ST(\xi)$ where $\xi \in K$. (The last fact follows

from $\chi(\xi) \in \Gamma$). Now assume that $\gamma' \in \Gamma$ represents $\sigma x = \sigma ST(\xi)$. Then $\gamma' = \gamma'' \chi(\xi)$ where γ'' ends in O. Since $\lambda(\gamma'') = \sigma S$ we find that every γ' ends in $\omega_a \chi(\xi)$ where ω_a denotes the final edge of α. This implies

$$\sigma x \in \mathrm{MRE}(L) \iff a\xi \in \mathrm{MRE}(K).$$

Similarly we have $\overline{\sigma} x \in \mathrm{MRE}(L) \iff b\xi \in \mathrm{MRE}(K)$. We deduce that $x \in \mathrm{BF}(L) \iff \xi \in \mathrm{BF}(K)$ and likewise for BO and BS. This proves $(\Rightarrow)$ in case α.

β) Not all $\gamma \in \Gamma(x)$ have the same final vertex.

Definition. *A finite word x is called traceable from P if a $\gamma \in \Gamma(x)$ exists with initial vertex P. We write $P \in \mathrm{Tr}(x)$. An infinite word $\tau : \mathbf{N} \to \{a, b\}$ is called traceable from P if every prefix of τ is traceable from P. We write $P \in \mathrm{Tr}(\tau)$.*

Definition. *A finite word x is called distinctly traceable (dt) if $\gamma, \delta \in \Gamma(x)$ exist with different endpoints. An infinite word $\tau : \mathbf{N} \to \{a, b\}$ is distinctly traceable if every prefix of τ is distinctly traceable.*

Suppose that x is dt of length n and choose $\gamma = P_0 \cdots P_n, \delta = Q_0 \cdots Q_n$ in $\Gamma(x)$ with $P_n \neq Q_n$. Then induction shows that $P_i \neq Q_i$ for all $0 \leq i \leq n$. In particular $P_i \neq O \vee Q_i \neq O$ for all $0 \leq i \leq n$. It follows that x is completely determined by the triple (P_0, Q_0, n) and we call (P_0, Q_0) a starting pair for x. If x, x' are dt with a common starting pair then it is easily seen that x, x' are prefix-related.

Corollary. *There exists a finite collection $\{\tau_i\}$ of words (finite or right-infinite) such that every dt word x is prefix of some τ_i.*

Corollary. *There exists a constant M and a finite collection $\mathcal{T}$ of dt words $\tau_i : \mathbf{N} \to \{a, b\}$ such that every finite dt word x of length at least M is prefix of exactly one τ_i.*

If w is a word we will write $[w]_n$ for its prefix of length n if it exists and $[w]^n$ for its suffix of length n (if it exists). Now let τ be any right-infinite word. Then there exists a constant c_τ such that $\mathrm{Tr}(\tau) = \mathrm{Tr}([\tau]_i)$ for all $i \geq c_\tau$. Indeed, $(\mathrm{Tr}([\tau]_i))_{i=0}^{\infty}$ is a decreasing sequence of finite sets. Such a sequence is always ultimately constant and its ultimate value is $\mathrm{Tr}(\tau)$. Enlarging the previous constant M if necessary we may assume that $M \geq c_\tau, c_{a\tau}, c_{b\tau}$ for all $\tau \in \mathcal{T}$.

To sum up, we assume that $x \in L$ is bispecial and dt with $|x| \geq M$. We let $\tau \in \mathcal{T}$ be the unique element from which x is a prefix. We let μ be the

symbol for which $x\mu$ is also a prefix of τ and we denote by δ_a, δ_b the initial edges of α, β.

By extendibility a symbol ν exists such that $\nu x\overline{\mu} \in L$. Choose $\gamma \in \Gamma(\nu x\overline{\mu})$ with initial vertex P. Then $P \in \text{Tr}(\nu x) = \text{Tr}(\nu\tau) = \text{Tr}(\nu x\mu)$ hence $\nu x \in \text{MRE}(L)$ and $x \notin \text{BF}(L)$. Hence to prove $(\Rightarrow)$ we need only consider $x \in \text{BS}(L)$, which we do. Choose $\gamma \in \Gamma(x\overline{\mu})$ with initial point P. Then $P \in \text{Tr}(x) = \text{Tr}(\tau) = \text{Tr}(x\mu)$, hence γ has final edge $\delta_{\overline{\mu}}$. Since $x\overline{\mu} \in \text{MLE}(L)$ we find, as in α, that $\gamma = \phi\chi(\xi)\delta_{\overline{\mu}}$ where $\lambda(\phi) = S$ and where ξ depends only on x, not on γ. Hence $x = ST(\xi)$, any $\gamma \in \Gamma(\sigma x\overline{\mu})$ ends in $\omega_a\chi(\xi)\delta_{\overline{\mu}}$ and any $\gamma \in \Gamma(\overline{\sigma}x\overline{\mu})$ ends in $\omega_b\chi(\xi)\delta_{\overline{\mu}}$. Considering a $P \in \text{Tr}(\sigma x\overline{\mu})$ yields $a\xi \in \text{MRE}(K)$ and similarly $b\xi \in \text{MRE}(K)$. Therefore $\xi \in \text{BS}(K)$ and this concludes the proof of $(\Rightarrow)$.

Now assume that $x = ST(\xi)$ with $|x| \geq M$ and ξ bispecial in K. Then it is easily seen that x is bispecial in L and the previous arguments apply. We may assume that x is dt since otherwise we are done. We define τ, μ as before. Any $\gamma \in \Gamma(\sigma x\overline{\mu}) = \Gamma(\sigma ST(\xi)\overline{\mu})$ ends in $\omega_a\chi(\xi)\delta_{\overline{\mu}}$ and any $\gamma \in \Gamma(\overline{\sigma}x\overline{\mu})$ ends in $\omega_b\chi(\xi)\delta_{\overline{\mu}}$. We deduce

$$\sigma x \in \text{MRE}(L) \Rightarrow a\xi \in \text{MRE}(K), \ \overline{\sigma}x \in \text{MRE}(L) \Rightarrow b\xi \in \text{MRE}(K).$$

The reverse implications in this line are clear when $x = ST(\xi)$. Hence

$$\sigma x \in \text{MRE}(L) \iff a\xi \in \text{MRE}(K), \ \overline{\sigma}x \in \text{MRE}(L) \iff b\xi \in \text{MRE}(K).$$

Hence $x \in \text{BS/BO/BF}(L) \iff \xi \in \text{BS/BO/BF}(K)$ and we are done. $\square$

Remark.   The given proof shows $x \in \text{BO}(L) \Leftarrow x = ST(\xi), \xi \in \text{BO}(K)$ for $|x| \geq M$. We wondered if the reversed implication is also true. This is not the case, as pointed out by the referee of this article. He kindly supplied us with the next example. Let $K = \{a, b\}^*$ and $T = (aba, b)$. Then $\text{BO}(K) = \emptyset$ and $\text{BO}(L) = a(baba)^* \cup ab(abab)^* \cup ba(baba)^* \cup bab(abab)^*$.

2.1. An application to balanced words. Again we start off with some necessary definitions. The content $c(x)$ of a finite word x is the number of a's that it contains.

Definition.   *A word w is balanced if $|c(x) - c(y)| \leq 1$ for all $x, y \subset w$ with $|x| = |y|$.*

We denote the language of balanced words by Bal. One can show, see [H, Thm 2.3], that $K := \text{Bal}$ is a CE-language. The balanced **Z**-words can be classified, see [H, Thm 2.5], [MH, Thms 5.1,5.2,5.3], [T, Thm 2], and the bispecial factors in K are also known. More precisely, $\text{BF}(K) = \emptyset$

(there are no weakly bispecials) and $BS(K) = \mathcal{H}$, the class of finite Hedlund words. Finite Hedlund words can be introduced in various ways. In [H, Section 2.3] we define them by means of infinite Hedlund words, in [H, Section 2.4] we show $BF(K) = \emptyset, BS(K) = \mathcal{H}$ and we give an arithmetical description of finite Hedlund words. To give this description we need some more notation. Let $A \subset I$ where I is a subinterval of $\mathbf{Z}$. Then A induces a word $w : I \rightarrow \{a, b\}$ by setting $w_i = a \iff i \in A$. By abuse of notation we denote this induced word by $A \subset I$. The finite Hedlund words are then given by

$$\{\lfloor \frac{i(r+s)}{k+l} \rfloor\}_{i=1}^{k+l-1} \subset [1, r+s-2]$$

where k, l, r, s are integers with $0 \le k \le r, 0 \le l \le s$ and $lr - ks = 1$. See [H, p. 27] where this word is denoted by $\mathrm{per}(s, r, 1)$. Note that $(s, r) = 1$.

Remark. A word $w : I \rightarrow \{a, b\}$ is called constant if $|w(I)| = 1$ and w is said to have period p if $w_i = w_{i+p}$ whenever $i, i + p \in I$. A famous theorem by Fine and Wilf [FW, Thm 1] implies that a non-constant word x with coprime periods s, r has length at most $s + r - 2$ and one can show that such an x is unique up to exchange of a and b. See also [T, Thm 3]. De Luca and Mignosi define PER as the collection of such x and show that $BS(K) = \mathrm{PER}$. Hence the class PER in [dL/Mi] equals our class $\mathcal{H}$.

The word $x := \mathrm{per}(s, r, 1)$ satisfies $c(x) = k + l - 1, |x| = r + s - 2$ and from this one easily deduces that the pair (s, r) is unique. It also follows from the above that a word $x \in \mathcal{H}$ with $c(x) = \lambda, |x| = n$ exists if and only if $0 \le \lambda \le n, (\lambda + 1, n + 2) = 1$ and that such an x is unique. We leave these details to the reader. As a direct consequence we have $|BS_n(K)| = \phi(n + 2), |BF_n(K)| = 0$ and using Proposition 1 we find the following formula for $\mathrm{bal}(n) := P(K, n) :$

$$\mathrm{bal}(n) = 1 + \sum_{k=1}^{n} (n + 1 - k)\phi(k)$$

This formula is well-known and has a number of alternative proofs, see [Be/Po][dL/Mi, Thm 7][Mi]. From this one can deduce the asymptotics $\mathrm{bal}(n) = \frac{n^3}{\pi^2} + O(n^2 \ln n)$, see [Mi]. We now investigate what happens when K is subjected to a substitution T.

Theorem 2. *Let $T = (X, Y)$ be an ABS, let $K = \mathrm{Bal}$ and $L = F(TK)$. Then*

$$\lim_{n \to \infty} \frac{P(L, n)}{P(K, n)} = \frac{1}{|X| \cdot |Y|}.$$

Proof. We define S as in Theorem 1 and $\mathrm{BS}^*(L) = ST(\mathcal{H}), \mathrm{BF}^*(L) = \emptyset$. Then Theorem 1 implies that the symmetric differences $\mathrm{BS}^*(L) \triangle \mathrm{BS}(L)$ and $\mathrm{BF}^*(L) \triangle \mathrm{BF}(L)$ are finite. Combining Proposition 1 with this fact and writing $f(n) \sim g(n)$ for $f(n) - g(n) = O(n)$ we find

$$P(L, n) \sim \sum_{i=0}^{n-1} (n - 1 - i)|\mathrm{BS}_i^*(L)| \qquad (1)$$

If $x \in \mathrm{BS}^*(L)$ then by definition we have $x = ST(\xi)$ where $\xi \in \mathcal{H}$. Writing $|x| - |S| = p, c(x) - c(S) = q, |X| = \alpha, |Y| = \beta, c(X) = \gamma, c(Y) = \delta, c(\xi) = \lambda, |\xi| - c(\xi) = \mu$ we have

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \qquad (2)$$

As we have seen above in a slightly different form, there exists a $\xi \in \mathcal{H}$ with parameters (λ, μ) iff $\lambda, \mu \in \mathbf{N}, (\lambda + 1, \mu + 1) = 1$ and ξ is unique in this case. We can now translate this into conditions for (p, q). Put $d := (\alpha, \beta)$. Then obviously $p \in d\mathbf{Z}$ and we write $\alpha = d\tilde{\alpha}, \beta = d\tilde{\beta}, p = d\mathcal{P}$. Then

$$\begin{pmatrix} \mathcal{P} \\ q \end{pmatrix} = \begin{pmatrix} \tilde{\alpha} & \tilde{\beta} \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \qquad (3)$$

Choose integers $k, l \in \mathbf{Z}$ with $0 < l \leq \tilde{\alpha}$ and $k\tilde{\alpha} + l\tilde{\beta} = 1$. Since $\lambda\tilde{\alpha} + \mu\tilde{\beta} = \mathcal{P}$ we have

$$\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \in \begin{pmatrix} \mathcal{P}k \\ \mathcal{P}l \end{pmatrix} + \mathbf{Z}\begin{pmatrix} \tilde{\beta} \\ -\tilde{\alpha} \end{pmatrix} \qquad (4)$$

It follows that $|\mathrm{BS}^*_{|S|+d\mathcal{P}}(L)|$ equals the number of integers t that satisfy $\mathcal{P}k + t\tilde{\beta}, \mathcal{P}l - t\tilde{\alpha} \geq 0, (\mathcal{P}k + t\tilde{\beta} + 1, \mathcal{P}l - t\tilde{\alpha} + 1) = 1$. Some calculation then shows that $|\mathrm{BS}^*_{|S|+d\mathcal{P}}(L)|$ equals the number of integers t with

$$\frac{-\mathcal{P}k}{\tilde{\beta}} \leq t \leq \frac{\mathcal{P}l}{\tilde{\alpha}}, \ (t + l - k, \mathcal{P} + \tilde{\alpha} + \tilde{\beta}) = 1$$

We now need a small lemma on the Euler ϕ-function.

Lemma 2.1. *Let $0 \leq \alpha < \beta \leq 1$ be given and define $\phi_{\alpha,\beta}(n) := |\{k : \alpha n \leq k \leq \beta n, \ (k, n) = 1\}|$. Then $\lim_{n \to \infty} \frac{\phi_{\alpha,\beta}(n)}{\phi(n)} = \beta - \alpha$.*

Proof. We will generalize the classical inclusion/exclusion proof for the formula $\phi(n) = n\Pi_{p|n}(1 - \frac{1}{p})$. All statements in the proof will have fixed n, α, β. We write $n = p_1^{k_1} \cdots p_s^{k_s}$ for the prime decomposition of n and we denote by $N(i)$ the number of multiples of i in $[\alpha n, \beta n]$. By the principle of inclusion/exclusion we have

$$\phi_{\alpha,\beta}(n) = N(1) - N(p_1) - \cdots - N(p_s) \cdots + (-1)^s N(p_1 \cdots p_s)$$

This implies

$$\phi_{\alpha,\beta}(n) = (\beta - \alpha)n - \frac{(\beta - \alpha)n}{p_1} \cdots - \frac{(\beta - \alpha)n}{p_s} \cdots + (-1)^s \frac{(\beta - \alpha)n}{p_1 \cdots p_s} + O(2^s)$$

since the termwise difference is at most 1. Dividing by $\phi(n)$ we find $\frac{\phi_{\alpha,\beta}(n)}{\phi(n)} = \beta - \alpha + \frac{O(2^s)}{\phi(n)}$ and we will now show that the final term has limit 0 as $n \to \infty$. Choose $\epsilon > 0$, choose $\sigma \in \mathbf{N}^+$ such that $s \geq \sigma \Rightarrow \frac{2^s}{s!} < \epsilon$ and finally choose $N \in \mathbf{N}^+$ such that $n \geq N \Rightarrow \frac{2^\sigma}{\phi(n)} < \epsilon$. Let $n \geq N$ be arbitrary. If $s = s(n) \geq \sigma$ we have $\frac{2^s}{\phi(n)} = \frac{2^s}{n\Pi(1-1/p)} \leq \frac{2^s}{\Pi(p-1)} \leq \frac{2^s}{s!} < \epsilon$. If $s < \sigma$ we have $\frac{2^s}{\phi(n)} < \frac{2^\sigma}{\phi(n)} < \epsilon$. Since ϵ was arbitrary it follows that $\lim_{n\to\infty} \frac{2^s}{\phi(n)} = 0$ and we are done. $\square$

Applying this to the "formula" for $|\mathrm{BS}^*_{|S|+d\mathcal{P}}(L)|$ and using that $\frac{l}{\tilde{\alpha}} + \frac{k}{\tilde{\beta}} = \frac{1}{\tilde{\alpha}\cdot\tilde{\beta}}$ we find that $\lim_{\mathcal{P}\to\infty} \frac{|\mathrm{BS}^*_{|S|+d\mathcal{P}}(L)|}{\phi(\mathcal{P}+\tilde{\alpha}+\tilde{\beta})} = \frac{1}{\tilde{\alpha}\cdot\tilde{\beta}}$. To finish things off we have the following small Tauberian lemma.

Lemma 2.2.   *Suppose that $a_n \in \mathbf{R}, b_n \in \mathbf{R}^+$ for all $n \geq 1$ with $\lim_{n\to\infty} \frac{a_n}{b_n} = \lambda$ and $\sum_1^\infty b_n = \infty$. If $A_n := \sum_{k=1}^n a_k$ and $B_n := \sum_{k=1}^n b_k$, then $\lim_{n\to\infty} \frac{A_n}{B_n} = \lambda$.*

Proof. Replacing a_i by $a_i - \lambda b_i$ we may assume $\lambda = 0$. Fix $\epsilon > 0$ and choose $N \in \mathbf{N}^+$ such that $|a_n| < \epsilon b_n$ for $n \geq N$. For all $i \geq 0$ we have

$$|\frac{A_{N+i}}{B_{N+i}}| \leq \frac{|A_N| + \epsilon(b_{N+1} + \cdots + b_{N+i})}{B_{N+i}} \leq \frac{|A_N|}{B_{N+i}} + \epsilon$$

Choose N' s.t. $B_{N+i} \geq \frac{|A_N|}{\epsilon}$ for $i \geq N'$. Then $|\frac{A_j}{B_j}| < 2\epsilon$ for $j \geq N + N'$. $\square$

If we apply Lemma 2.2 not once but twice and use partial summation we also obtain $\lim_{n\to\infty} \frac{\sum_1^n (n+1-i)a_i}{\sum_1^n (n+1-i)b_i} = \lambda$. In the following we write $\sigma = \lfloor \frac{n-1-|S|}{d} \rfloor$, we write $f(n) \cong g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$ and $f(n) \approx g(n)$ if $f(n) - g(n) = O(n^2)$. Note that $\sum_{k=1}^n \phi(k) = O(n^2)$ since $\phi(k) \leq k$ for all

k. Combining the previous material we find

$$P(L, n) \sim \sum_{i=0}^{n-1} (n - 1 - i)|\mathrm{BS}_i^*(L)|$$

$$= d \sum_{\mathcal{P}=0}^{\sigma} \left(\frac{n - 1 - |S|}{d} - \mathcal{P}\right)|\mathrm{BS}_{|S|+d\mathcal{P}}^*(L)|$$

$$\approx d \sum_{\mathcal{P}=1}^{\sigma} (\sigma + 1 - \mathcal{P})|\mathrm{BS}_{|S|+d\mathcal{P}}^*(L)|$$

$$\cong \frac{d^3}{\alpha\beta} \sum_{\mathcal{P}=1}^{\sigma} (\sigma + 1 - \mathcal{P})\phi(\mathcal{P} + \tilde\alpha + \tilde\beta)$$

$$= \frac{d^3}{\alpha\beta}\{\mathrm{bal}(\sigma + \tilde\alpha + \tilde\beta) - \mathrm{bal}(\tilde\alpha + \tilde\beta) - \sigma \sum_{\mathcal{P}=1}^{\tilde\alpha+\tilde\beta} \phi(\mathcal{P})\}$$

$$\sim \frac{d^3}{\alpha\beta}\mathrm{bal}(\sigma + \tilde\alpha + \tilde\beta)$$

$$\cong \frac{d^3}{\alpha\beta} \cdot \frac{\mathrm{bal}(n)}{d^3}$$

$$= \frac{\mathrm{bal}(n)}{|X| \cdot |Y|}.$$

$\square$

Intermezzo. Theorem 2 relies heavily on the properties of Bal. To illustrate this we consider another example, $K = \{a, b\}^*$. The reader can skip this part and continue with Section 2.2 if he wishes to continue directly with applications to balanced words.

Theorem 3. *Let $T = (X, Y)$ be an ABS and let S be the greatest common suffix of $\{X, Y\}^{-\infty}$ as in Theorem 1. We write $\alpha := |X|, \beta := |Y|, d := (\alpha, \beta), \tilde\alpha := \frac{\alpha}{d}, \tilde\beta := \frac{\beta}{d}, p := \max(\tilde\alpha, \tilde\beta), q := |\tilde\alpha - \tilde\beta|$. Let ϕ be the unique positive root of $x^p - x^q - 1$ (existence and uniqueness will be established below). We write $\mathrm{fr}(x) := x - \lfloor x \rfloor$ for fractional parts, for $n \in \mathbf{N}$ we write $\sigma = \sigma(n) := \lfloor \frac{n-1-|S|}{d} \rfloor$ and $\epsilon_n := 1 - \mathrm{fr}(\frac{n-1-|S|}{d})$. Also let $K = \{a, b\}^*$ and $L = F(TK)$. Then there exists a constant $\gamma \in \mathbf{R}^+$ such that*

$$P(L, n) = \gamma(\phi - (\phi - 1)\epsilon_n)\phi^\sigma + o(\phi^\sigma)$$

Proof. We will use techniques similar to those in the proof of Theorem 2. Note that $\mathrm{BF}(K) = \emptyset, \mathrm{BS}(K) = K$ and let $\mathrm{BF}^*(L) = \emptyset, \mathrm{BS}^*(L) = ST(K)$. Formula (1) remains valid as it stands. It is clear that $\mathrm{BS}_i^*(L) = \emptyset$ when

$i \notin |S| + d\mathbf{N}$. We define $f(\mathcal{P}) := |\mathrm{BS}^*_{|S|+d\mathcal{P}}(L)|$, then $f(\mathcal{P}) = \{x \in K : |Tx| = d\mathcal{P}\}|$. Dividing up with respect to the final symbol of x, we find $f(\mathcal{P}) = f(\mathcal{P} - \tilde{\alpha}) + f(\mathcal{P} - \tilde{\beta})$ for $\mathcal{P} \geq \tilde{\alpha}, \tilde{\beta}$. Replacing $\mathcal{P}$ by $\mathcal{P} + p$ we find $f(\mathcal{P}+p) = f(\mathcal{P}) + f(\mathcal{P}+q)$ for all $\mathcal{P} \in \mathbf{N}$. The minimal polynomial of this recurrence relation equals $g = x^p - x^q - 1$. Note $p > q \geq 0$ and $(p, q) = 1$. We will now prove some relevant properties of g in order to establish the asymptotics of $f(\mathcal{P})$.

Lemma 3.1. *Let $g = x^p - x^q - 1$ with $p > q \geq 0$ and $(p, q) = 1$. Then:*

a) *g has only simple roots;*
b) *g has exactly one positive root ϕ and $\phi \in (1, 2]$;*
c) *g has another real root iff p is even in which case this other root lies in $(-1, 0)$;*
d) *Any circle $C(0, r)$ with $r > 0$ contains at most 2 roots of g. If it contains two distinct roots x, y then x, y are non-real and conjugate;*
e) *ϕ is the unique root of maximal length.*

Proof. We assume $q \geq 1$ since $q = 0$ implies $p = 1, g = x - 2$ and then the statements are clear. We have $g' = x^{q-1}(px^{p-q} - q)$.
a) If g has a double root x then $x^q(x^{p-q}-1) = 1$ and $x^{p-q} = \frac{q}{p}$. Substitution yields $x^q = \frac{-p}{p-q}$ and then multiplication yields $x^p = \frac{-q}{p-q}$. Hence $x^p, x^q \in \mathbf{Q}$ and from $(p, q) = 1$ we deduce $x \in \mathbf{Q}$. N ow x is an algebraic integer, hence $x \in \mathbf{Z}$, and all integer roots of g must divide 1. But $g(-1) \equiv 1(2)$ and $g(1) = -1$.
b) This follows from g' and $g(0) = g(1) = -1, g(2) = 2^p - 2^q - 1 \geq 2^q - 1 > 0$.
c) From g' one deduces that g has at most one root in $\mathbf{R}^-$. If N is the number of real roots then $N \equiv p(2)$ since non-real zeroes appear in conjugate pairs. Hence another real root exists iff p is even. Then q is odd, $g(-1) = 1$ and $g(0) = -1$.
d) Fix r and suppose $g(x) = 0$ with $|x| = r$. Then $x^{p-q} - 1 = \frac{1}{x^q}$. Writing $\tau := x^{p-q}$ we have $\tau \in C(0, r^{p-q}) \cap C(1, \frac{1}{r^q})$. Hence at most two τ are possible. If roots x, y of g with $|x| = |y| = r$ give the same τ then $x^{p-q} = y^{p-q}$ and $x^q = y^q$. Hence $x = y$. It follows that $C(0, r)$ contains at most two distinct zeroes. If it contains a non-real root x, then $\overline{x}$ is also a root and the number of zeroes is 2. If it contains a real root x, then also $\tau \in \mathbf{R}$. Since $C(0, r^{p-q})$ and $C(1, \frac{1}{r^q})$ have a real point of intersection they are tangent. Hence for such r there is only one τ and only one x.
e) $g(x) = 0 \Rightarrow |x|^p = |x^q + 1| \leq |x|^q + 1 \Rightarrow g(|x|) \leq 0 \Rightarrow |x| \leq \phi$. Hence ϕ is indeed a zero of maximal length and its uniqueness follows from d). $\square$

Proof of Theorem 3 (continued). We denote the roots of g by $(\omega_i)_{i=1}^p$

where we suppose that $|\omega_1| \geq \cdots \geq |\omega_p|$. Since g has simple roots we have

$$f(\mathcal{P}) = C_1\omega_1^{\mathcal{P}} + \cdots + C_p\omega_p^{\mathcal{P}}, \; \mathcal{P} \in \mathbf{N}$$

for uniquely determined constants $C_i \in \mathbf{C}$. Note that $\omega_1 = \phi$. We will show that $C_1 > 0$. Suppose that $C_1 = 0$ and let k be the minimal index with $C_k \neq 0$. Then $\omega_k \notin \mathbf{R}$ since $\omega_k \in \mathbf{R}$ would imply $|\omega_k| < 1$ and then $\lim_{\mathcal{P}\to\infty} f(\mathcal{P}) = 0$. Since $f : \mathbf{N} \to \mathbf{N}$ this would imply $f(n) = 0$ for n large and the recursive formula for $f(\mathcal{P})$ then shows $f(\mathbf{N}) = \{0\}$. A contradiction. Hence $\omega_k \notin \mathbf{R}$ and by d) we have $\omega_{k+1} = \overline{\omega_k}$. We can conjugate the above formula for $f(\mathcal{P})$ and by unicity of the C_i we obtain $C_{k+1} = \overline{C_k}$. With $C := C_k, \omega := \omega_k, \rho := |\omega|$ we find $f(\mathcal{P}) = 2\mathrm{Re}(C\omega^{\mathcal{P}}) + o(\rho^{\mathcal{P}})$. Writing $C =: re^{i\psi}, \omega =: \rho e^{i\theta}$ this implies $\frac{f(\mathcal{P})}{\rho^{\mathcal{P}}} = 2r\cos(\psi + \theta\mathcal{P}) + o(1)$. We can assume $0 < \theta < \pi$, interchanging ω_k, ω_{k+1} if necessary. Since $\frac{f(\mathcal{P})}{\rho^{\mathcal{P}}} \geq 0$ for all $\mathcal{P} \in \mathbf{N}$ we deduce $\liminf_{k\to\infty} \cos(\psi + \theta k) \geq 0$. This is clearly impossible for $0 < \theta < \pi$: the set $\{\cos(\psi + \theta k)\}_{k=1}^{\infty}$ lies dense in $[-1, 1]$ if $\frac{\theta}{\pi} \notin \mathbf{Q}$ and in the other case there exists an arithmetic sequence $k_n = k_0 + nK$ such that $\cos(\psi + \theta k_n)$ is constant and negative. This contradiction shows that $C_1 \neq 0$. Hence

$$f(\mathcal{P}) \cong C_1\phi^{\mathcal{P}} \; (5)$$

and $C_1 > 0$. We now write $f \approx g$ if $f - g = o(\phi^{\sigma})$. We now combine (1),(5), Lemma 2.2 and the remark following that lemma. Then

$$P(L, n) \sim \sum_{i=0}^{n-1}(n - 1 - i)|\mathrm{BS}_i^*(L)|$$

$$= d\sum_{\mathcal{P}=0}^{\sigma}(\frac{n - 1 - |S|}{d} - \mathcal{P})f(\mathcal{P})$$

$$= d\{\sum_{\mathcal{P}=0}^{\sigma}(\sigma + 1 - \mathcal{P})f(\mathcal{P}) - \epsilon_n\sum_{\mathcal{P}=0}^{\sigma}f(\mathcal{P})\}$$

$$\approx C_1d\{\sum_{\mathcal{P}=0}^{\sigma}(\sigma + 1 - \mathcal{P})\phi^{\mathcal{P}} - \epsilon_n\sum_{\mathcal{P}=0}^{\sigma}\phi^{\mathcal{P}}\}$$

$$= C_1d\{\frac{\phi^{\sigma+2} - (\phi - 1)(\sigma + 1) - \phi}{(\phi - 1)^2} - \frac{\epsilon_n(\phi^{\sigma+1} - 1)}{\phi - 1}\}$$

$$\sim C_1d\{\frac{\phi^{\sigma+2}}{(\phi - 1)^2} - \frac{\epsilon_n\phi^{\sigma+1}}{\phi - 1}\}.$$

Theorem 3 now follows with $\gamma := \frac{\phi C_1 d}{(\phi-1)^2}$. $\qquad\qquad\square$

2.2. More application to balanced words. We now consider $K = \text{Bal}$ again. One can show that every balanced $\mathbf{Z}$-word has a density in the following sense. Let $x_n \subset w$ be a factor of length n for each n. Then $\lim_{n\to\infty} \frac{c(x_n)}{|x_n|}$ exists and only depends on w, not on (x_n). Its value $\alpha(w)$ is called the density of w. Note that $\alpha(w) \in [0,1]$. A $\mathbf{Z}$-word w is called recurrent if each factor $x \subset w$ appears at least twice in w.

Proposition 2. *Let w be a recurrent balanced $\mathbf{Z}$-word of density α. If $\alpha = 0$ then $w = b^{\infty}$. Otherwise we set $\zeta := \frac{1}{\alpha}$ and we define the set $W \subset \mathbf{Z}$ by $w_i = a \iff i \in W$. Then $W = \{\lfloor \zeta i + \phi \rfloor\}_{i \in \mathbf{Z}}$ for some $\phi \in \mathbf{R}$ or $W = \{\lceil \zeta i + \phi \rceil\}_{i \in \mathbf{Z}}$ for some $\phi \in \mathbf{R}$.*

Proof. This follows from the classification of balanced $\mathbf{Z}$-words in [H, Thm 2.5] by considering which $\mathbf{Z}$-words are recurrent. $\qquad\square$

The $\mathbf{Z}$-words in Proposition 2, including b^{∞}, are called Beatty $\mathbf{Z}$-words (BZW). They can also be obtained as the coding of a rotation on the unit circle, see for instance [H, Section 2.5.2] for details. If $\mathcal{S}$ is the collection of all BZW's, then $\text{Bal} = F(\mathcal{S})$. This is well-known and can be seen, for instance, by combining Theorems 2.3, 2.5, 2.8 in [H]. We now investigate what happens when we put restrictions on the density α. It turns out that balanced words are, in some sense, uniformly distributed w.r.t. density. We would like to thank Julien Cassaigne from IML Marseille for suggesting that this might be the case.

Theorem 4. *Let $I \subset [0,1]$ be given and let $\mathcal{S}_I$ be the collection of BZW's σ with $\alpha(\sigma) \in I$. Also let $\text{bal}_I(n) := P(\mathcal{S}_I, n)$ and denote Lebesgue measure on $[0,1]$ by λ. Then*

$$\lambda(I^{\circ}) \leq \liminf_{n\to\infty} \frac{\text{bal}_{I^{\circ}}(n)}{\text{bal}(n)} \leq \limsup_{n\to\infty} \frac{\text{bal}_{\overline{I}}(n)}{\text{bal}(n)} \leq \lambda(\overline{I}).$$

In particular we have $\lim_{n\to\infty} \frac{\text{bal}_I(n)}{\text{bal}(n)} = \lambda(I)$ (*) when $\lambda(\partial I) = \emptyset$. To prove Theorem 4 we show (*) directly for a particular class of intervals, the so-called Farey intervals.

2.3. Sturmian substitutions and Farey intervals. A BZW σ with irrational density is called a Sturmian $\mathbf{Z}$-word. We write $\mathcal{S}'$ for the class of sturmian $\mathbf{Z}$-words. Let $L = (ba, b), R = (ab, b), C = (b, a)$. Since substitutions form a monoid (halfgroup with unit), we can consider the submonoid $\mathcal{M} := < L, R, C >$ generated by $\{L, R, C\}$. We call it the monoid of Sturmian substitutions. The name is explained by the next well-known

proposition.

Proposition 3. *For a substitution T we have $T \in \mathcal{M} \iff T\sigma \in \mathcal{S}'$ for some $\sigma \in \mathcal{S}' \iff T\sigma \in \mathcal{S}'$ for all $\sigma \in \mathcal{S}'$.*

Proof. See [Mi/S]. It also follows from [H, Thm 3.1]. □

Now let $T = (X, Y) \in \mathcal{M}$ with $c(X) =: l, |X| =: s, c(Y) =: k, |Y| =: r$. An easy induction argument shows that $lr - ks \in \{-1, 1\}$. Hence $\mathrm{CH}(\frac{l}{s}, \frac{k}{r})$, where CH denotes convex hull, is a Farey interval.

Definition. *A Farey interval is an interval $I = [\frac{\alpha}{\beta}, \frac{\gamma}{\delta}]$ with $\alpha, \gamma \in \mathbf{N}, \beta, \delta \in \mathbf{N}^+, 0 \le \frac{\alpha}{\beta} \le \frac{\gamma}{\delta} \le 1$ and $\beta\gamma - \alpha\delta = 1$. Also $\mathcal{F}$ is the collection of all Farey intervals.*

We define the density of a finite non-empty word x as $\alpha(x) := \frac{c(x)}{|x|}$. Then we have shown that there exists a mapping $\kappa : \mathcal{M} \to \mathcal{F}$ sending $T = (X, Y)$ to $\kappa(T) := \mathrm{CH}(\alpha(X), \alpha(Y))$. We define $\mathcal{M}^* \subset \mathcal{M}$ as the submonoid generated by $A = (a, ba) = CRC$ and $B = (ab, b) = R$. We call it the monoid of distinguished sturmian substitutions. In fact we have $\mathcal{M}^* = \mathcal{M} \cap \mathrm{ABS}$, the proof of which we leave as an exercise to the reader. With induction one shows $\alpha(X) > \alpha(Y)$ when $(X, Y) \in \mathcal{M}^*$. It is easy to see that $\mathcal{M}^* :=< A, B >$ is freely generated by A, B. Indeed, suppose that $A\Phi = B\Psi$ where $\Phi, \Psi \in \mathcal{M}^*$. Then $ba \subset \Phi(ba)$ since $[\Phi(a)]_1 = a, [\Phi(b)]_1 = b$. Applying A we find $aa \subset A\Phi(ba) = B\Psi(ba)$, a contradiction. To proceed we introduce some more structure on $\mathcal{F}$. If $I = [\frac{\alpha}{\beta}, \frac{\gamma}{\delta}]$ is any interval with $\alpha, \gamma \in \mathbf{N}, \beta, \delta \in \mathbf{N}^+, (\alpha, \beta) = (\gamma, \delta) = 1$ we define $\mathcal{L}I = I^- = [\frac{\alpha}{\beta}, \frac{\alpha+\gamma}{\beta+\delta}]$ and $\mathcal{R}I = I^+ = [\frac{\alpha+\gamma}{\beta+\delta}, \frac{\beta}{\delta}]$. Direct calculation shows $\mathcal{L}, \mathcal{R} : \mathcal{F} \to \mathcal{F}$ and $\mathcal{L}(\mathcal{F}) \cup \mathcal{R}(\mathcal{F}) \cup \{[0, 1]\} = \mathcal{F}$. Hence $\mathcal{F}$ is the smallest collection of intervals containing $[0, 1]$ which is closed under $\mathcal{L}, \mathcal{R}$. It follows that each $I \in \mathcal{F}$ has a unique notation $I = C_1 \cdots C_s([0, 1])$ where all $C_i \in \{\mathcal{L}, \mathcal{R}\}$.

Lemma 4.1.

a) *Let $\kappa : \mathcal{M} \to \mathcal{F}$ be the mapping described in the preceding paragraph and let $T = C_1 \cdots C_s \in \mathcal{M}^*$ where all $C_i \in \{A, B\}$. Define $D_i = \mathcal{L}$ if $C_i = B$ and $D_i = \mathcal{R}$ if $C_i = A$. Then $\kappa(T) = D_s \cdots D_1([0, 1])$;*

b) *The restriction $\kappa : \mathcal{M}^* \to \mathcal{F}$ is bijective;*

c) *$\mathcal{S}_{[0, \frac{1}{2}]} = B\mathcal{S}$ and $\mathcal{S}_{[\frac{1}{2}, 1]} = A\mathcal{S}$;*

d) *Suppose $\kappa(T) = I$ where $T \in \mathcal{M}^*, I \in \mathcal{F}$. Then $\mathcal{S}_I = T(\mathcal{S})$.*

Proof. a) Let $T = (X, Y) = C_1 \cdots C_s \in \mathcal{M}^*$, let $c(X) = l, |X| = s, c(Y) = k, |Y| = r$ and suppose that the statement is true for T. Hence $[\frac{k}{r}, \frac{l}{s}] = [\alpha(Y), \alpha(X)] = \kappa(T) = D_s \cdots D_1([0, 1])$. For $T' = TA$ we have $\alpha(T'a) = \frac{c(T'a)}{|T'a|} = \frac{l}{s}$ and $\alpha(T'b) = \frac{k+l}{r+s}$. Hence $\kappa(TA) = \mathcal{R}\kappa(T)$ and similarly $\kappa(TB) = \mathcal{L}\kappa(T)$. Part a) now follows since it is true for $T = (a, b)$.

b) Immmediate.

c) See [H, Lemma 2.14.1].

d) Suppose the statement is true for $T = C_1 \cdots C_s, I = D_s \cdots D_1([0, 1])$ and write $T = (X, Y)$. Suppose that σ is a BZW with $\alpha(\sigma) \in I$. Then $\sigma = T(\tau)$ for some BZW τ and with $\alpha(\tau) =: \alpha$ we have $\alpha(\sigma) = \frac{\alpha c(X) + (1-\alpha)c(Y)}{\alpha |X| + (1-\alpha)|Y|} =: g(\alpha)$. Since g is strictly increasing on $[0, 1]$ with $g(\frac{1}{2}) = \frac{c(X)+c(Y)}{|X|+|Y|}$, we have $\alpha(\sigma) \in \mathcal{L}I \iff \alpha(\tau) \leq \frac{1}{2} \iff \tau \in B\mathcal{S}$. Hence $\mathcal{S}_{\mathcal{L}I} = TB(\mathcal{S})$ and similarly $\mathcal{S}_{\mathcal{R}I} = TA(\mathcal{S})$. $\qquad\square$

Proof of Theorem 4. First let $I = [\frac{k}{r}, \frac{l}{s}] \in \mathcal{F}$ with all fractions in reduced form. Let $T = (X, Y) := \kappa^{-1}(I)$, then Lemma 4.1d tells us that $\mathcal{S}_I = T(\mathcal{S})$. Since $\frac{c(X)}{|X|} = \frac{l}{s}, \frac{c(Y)}{|Y|} = \frac{k}{r}$ and all fractions are in reduced form, we have $|X| = s, |Y| = r$. We write $K = F(\mathcal{S}) = \text{Bal}$ and $L = F(\mathcal{S}_I) =: \text{Bal}_I$. Applying Theorem 2 we find

$$\lim_{n \to \infty} \frac{\text{bal}_I(n)}{\text{bal}(n)} = \lim_{n \to \infty} \frac{P(L, n)}{P(K, n)} = \frac{1}{|X| \cdot |Y|} = \frac{1}{rs} = \lambda(I),$$

which is (*). We define the level of an interval $I = D_t \cdots D_1([0, 1]) \in \mathcal{F}$ as the integer t and we denote the class of Farey intervals of level t by $\mathcal{F}_t$. Then it is easy to show with induction that $\lambda(I) \leq \frac{1}{t+1}$ when $I \in \mathcal{F}_t$.

Lemma 4.2.

    a) *Let $A, B \subset [0, 1]$ be disjoint and compact. Then $\text{Bal}_A(n) \cap \text{Bal}_B(n) = \emptyset$ for n large;*

    b) *If $A, B \subset [0, 1]$ are closed intervals with $A \cap B = \{\alpha\}$ then $|\text{Bal}_A(n) \cap \text{Bal}_B(n)| = o(n^3)$.*

Proof. a) If σ is a BZW of density α and $x \subset w$ is a non-empty finite factor, then it is well-known that $|\alpha(x) - \alpha| \leq \frac{1}{|x|}$. See [H, Lemma 2.5.1]. Let $d := d(A, B) > 0$ and suppose that $\sigma \in \mathcal{S}_A, \tau \in \mathcal{S}_B$ have a non-empty factor x in common. Then $d \leq |\alpha(\sigma) - \alpha(\tau)| \leq |\alpha(\sigma) - \alpha(x)| + |\alpha(x) - \alpha(\tau)| \leq \frac{2}{|x|}$, whence $|x| \leq \frac{2}{d}$.

b) We assume A to the left of B. If $\sigma \in \mathcal{S}_A, \tau \in \mathcal{S}_B$ have a non-empty finite factor x in common then, as before, $|\alpha(\sigma) - \alpha(\tau)| \leq \frac{2}{|x|}$. Hence

$|\alpha(\tau) - \alpha| \le \frac{2}{|x|}$ and $\mathrm{Bal}_A(n) \cap \mathrm{Bal}_B(n) \subset \mathrm{Bal}_{[\alpha, \alpha + \frac{2}{n}]}(n)$. Let $t \ge 1$ and let I_t be the right-most element of $\mathcal{F}_t$ containing α. Then we have $\mathrm{Bal}_A(n) \cap \mathrm{Bal}_B(n) \subset \mathrm{Bal}_{I_t}(n)$ for $n \ge N(t)$. Hence $\limsup_{n \to \infty} \frac{P(\mathrm{Bal}_A \cap \mathrm{Bal}_B, n)}{\mathrm{bal}(n)} \le \frac{1}{t+1}$ since (*) holds for I_t. Now let $t \to \infty$. $\qquad\qquad\square$

Using Lemma 4.2 we find (*) for any finite union of Farey intervals and we denote the collection of such sets by Ω'. Now let I be any interval and fix $\delta > 0$. We can choose $I_1, I_2 \subset \Omega'$ with $I_1 \subset I \subset I_2$ and $\lambda(I_2 \setminus I_1) < \delta$. For n large we then have

$$\lambda(I) - 2\delta < \lambda(I_1) - \delta < \frac{\mathrm{bal}_{I_1}(n)}{\mathrm{bal}(n)} \le \frac{\mathrm{bal}_I(n)}{\mathrm{bal}(n)}$$
$$\le \frac{\mathrm{bal}_{I_2}(n)}{\mathrm{bal}(n)} < \lambda(I_2) + \delta < \lambda(I) + 2\delta.$$

This proves (*) for I and using Lemma 4.2 one finds (*) for any finite union of closed intervals. We denote the collection of such sets by Ω. Now let $I \subset [0,1]$ be any subset. Then I° is a countable union of disjoint open intervals. Approximating I° from the inside by elements of Ω one finds $\lambda(I^\circ) \le \liminf \frac{\mathrm{bal}_{I^\circ}(n)}{\mathrm{bal}(n)}$. Approximating $[0,1] \setminus \overline{I}$ from the inside by elements of Ω and applying Lemma 4.2a) we find $\limsup \frac{\mathrm{bal}_{\overline{I}}(n)}{\mathrm{bal}(n)} \le \lambda(\overline{I})$. $\qquad\square$

3. Recurrent Z-words of minimal block growth

If w is any **Z**-word then $P(w, n) = 1 + \sum_{i=0}^{n-1} |\mathrm{MRE}_i(w)|$ as in the proof of Proposition 1. Also it is clear that $\mathrm{MRE}_i(w) = \emptyset$ implies $\mathrm{MRE}_{i+1}(w) = \emptyset$. Hence $P(w, n)$ is strictly increasing in n, in which case $P(w, n) \ge n + 1$ for all n, or $P(w, n)$ is ultimately constant. In the last case it is well-known that w is periodic. This, in a sense, explains the following terminology introduced in [CH].

Definition. *Let w be a **Z**-word. Then w has minimal block growth (MBG) if there exist constants k, N with $P(w, n) = n + k$ for $n \ge N$.*

We have $k \ge 1$ and $k = k(w)$ is called the stiffness of w. A word w is called k-stiff if $P(n) \le n + k$ for all n. Hence a **Z**-word w has MBG if it is not periodic and k-stiff for some k. The minimal such k then equals $k(w)$. In this paper we concentrate on recurrent **Z**-words of minimal block growth. For the non-recurrent case we refer the interested reader to [H, Thms A,B of Section 3.1]. The situation for $k = 1$ is classical.

Proposition 4. *The Sturmian **Z**-words are precisely those recurrent **Z**-words with $P(n) = n + 1$ for all n.*

Proof. This follows from the classification of stiff **Z**-words in [H, Thm 2.5, 2.6] by considering which ones are recurrent and not periodic.   □

In a sense Sturmian words form the basis of general recurrent **Z**-words of MBG. We will see this in Proposition 5. An ABS $T = (X, Y)$ is called reduced if X, Y satisfy no suffix relations, i.e. if none is a suffix of the other. Then $T = (A\sigma C, B\overline{\sigma}C)$ for uniquely determined A, B, C, σ (where σ is a symbol). Conversely, any substitution of this form with $[A\sigma]_1 = a, [B\overline{\sigma}]_1 = b$ is reduced. The stiffness of such a T is defined as $k(T) := |ABC| + 1$. The name is explained by the following proposition.

Proposition 5. *Let w be a recurrent **Z**-word. Then $P(w, n) = n + k$ for large n iff $w = T(\sigma)$ where T is a reduced ABS of stiffness k and σ a Sturmian **Z**-word.*

Proof. This follows from [H, Thm 3.1] after noting that every ABS T is of the form $T = T^{\mathrm{RED}}\Phi$ with T^{RED} reduced and $\Phi \in \mathcal{M}^*$.   □

We denote the collection of reduced substitutions of stiffness k by $\mathcal{T}_k$. Since $|ABC| = k - 1$ one immediately finds that $\mathcal{T}_k$ is finite and in fact $|\mathcal{T}_k| = (k^2 + k + 2)2^{k-3}$. See [H, Lm 3.6.1] for a proof. We write $\mathcal{S}_k$ for the class of **Z**-words which are k-stiff but not $(k-1)$-stiff. Also we write $\mathcal{S}_k^{\mathrm{per}}/\mathcal{S}_k^{\mathrm{rnp}}/\mathcal{S}_k^{\mathrm{nr}}$ for those **Z**-words in $\mathcal{S}_k$ which are periodic/ recurrent but not periodic/ non-recurrent, respectively. Then $\mathcal{S}_k^{\mathrm{rnp}}, \mathcal{S}_k^{\mathrm{nr}}$ represent the **Z**-words of MBG of stiffness k. It has been shown that $F(\mathcal{S}_k^{\mathrm{per}}) = F(\mathcal{S}_k^{\mathrm{rnp}})$. See [H, Prop 3.3]. Hence on the level of finite factors we need not distinguish between $\mathcal{S}_k^{\mathrm{per}}$ and $\mathcal{S}_k^{\mathrm{rnp}}$. This is the reason why we concentrate on $\mathcal{S}_k^{\mathrm{rnp}}$.

For T an ABS and $K := \mathrm{Bal}$ we define $L_T := F(TK)$. Then $F(\mathcal{S}_k^{\mathrm{rnp}}) = \cup_{T \in \mathcal{T}_k} L_T$ by Proposition 5 and $\mathcal{B}(\mathcal{S}_k^{\mathrm{rnp}}, n) = \cup_{t \in \mathcal{T}_k} \mathcal{B}(L_T, n)$. If $T = (X, Y)$ is an ABS then $\lim_{n \to \infty} \frac{P(L_T, n)}{\mathrm{bal}(n)} = \frac{1}{|X| \cdot |Y|}$ by Theorem 2. We will show, in a sense to be made precise, that the languages L_T $(T \in \mathcal{T}_k)$ can be considered pairwise disjoint. This can be used to calculate the asymptotics of $P(\mathcal{S}_k^{\mathrm{rnp}}, n)$ and the result is the following.

Theorem 5. *For $k \in \mathbf{N}$ we have*

$$\lim_{n \to \infty} \frac{P(\mathcal{S}_k^{\mathrm{rnp}}, n)}{\mathrm{bal}(n)} = \frac{2^{k-1}}{k^2} + 2^{k-2} \cdot \sum_{i=1}^{k-1} \frac{1}{i^2}$$

Proof of Theorem 5. Let $T \in \mathcal{T}_k$ and write $T = (A\sigma C, B\overline{\sigma}C)$ as before. Then $\lim_{n\to\infty} \frac{P(L_T, n)}{\text{bal}(n)} = \frac{1}{(|AC|+1)(|BC|+1)} =: l(T)$. We write $V_k = \{(i, j)|0 \leq i, j \leq k - 3, i + j \geq k - 3\}$ and consider a few cases.

- $A = B = \epsilon$. Then $|C| = k - 1, \sigma = a$ and $l(T) = \frac{1}{k^2}$.
- $A = \epsilon, B = bB'$. Then $|B'C| = k - 2, \sigma = a$ and $l(T) = \frac{1}{k(|C|+1)}$. Writing $|C| + 1 =: i$ we find that exactly 2^{k-2} substitutions T of this type have $l(T) = \frac{1}{ki}$. The case $A \neq B = \epsilon$ is similar.
- Both A and B are non-empty. Then $A = aA', B = bB', |A'B'C\sigma| = k - 2$. Writing $|A'| =: k - 3 - i, |B'| =: k - 3 - j$ we have $0 \leq i, j \leq k - 3$ and $i + j = k - 3 + |C| \geq k - 3$. Hence $(i, j) \in V_k$ and $l(T) = \frac{1}{(i+2)(j+2)}$.

Taking "disjointness" for granted for the moment we find

$$\lim_{n\to\infty} \frac{P(\mathcal{S}_k^{\text{rnp}}, n)}{\text{bal}(n)} = \frac{2^{k-1}}{k^2} + \frac{2^{k-1}}{k}\sum_{i=1}^{k-1}\frac{1}{i} + 2^{k-2}\sum_{(i,j)\in V_k}\frac{1}{(i+2)(j+2)}$$

Setting $a_k := \frac{2}{k}\sum_{i=1}^{k-1}\frac{1}{i} + \sum_{V_k}\frac{1}{(i+2)(j+2)}$ we have

$$a_{k+1} - a_k = \frac{2}{k+1}\sum_{i=1}^{k}\frac{1}{i} - \frac{2}{k}\sum_{i=1}^{k-1}\frac{1}{i} + \sum_{V_{k+1}\setminus V_k} - \sum_{V_k\setminus V_{k+1}}$$

We have

$$\sum_{V_{k+1}\setminus V_k} = \sum_{\max(i,j)=k-2}\frac{1}{(i+2)(j+2)} = \frac{2}{k}\sum_{i=2}^{k}\frac{1}{i} - \frac{1}{k^2} \qquad \text{and}$$

$$\sum_{V_k\setminus V_{k+1}} = \sum_{i+j=k-3} = \sum_{i=0}^{k-3}\frac{1}{(i+2)(k-1-i)} = \frac{2}{k+1}\sum_{i=2}^{k-1}\frac{1}{i}$$

Inserting this we find $a_{k+1} - a_k = \frac{1}{k^2}$ and since $a_1 = 0$ we find $a_k = \sum_{i=1}^{k-1}\frac{1}{i^2}$. The theorem follows. We thank the referee of this article who suggested the shorter formula for a_k. We now turn to the problem of disjointness. We outline the remainder of the proof in a few lemmata.

Lemma 5.1. *Let T_1, T_2 be ABS, let K be a CE-language and let $L_i := F(T_iK)$. There exist a finite set Σ of **N**-words and a finite collection V of finite words such that $L_1 \cap L_2 \subset \text{Pref}(\Sigma) \cup V(T_1(K) \cap T_2(K))V$.*

Lemma 5.2. *Let T_1, T_2 be ABS and let $\Omega := \{a, b\}^*$. There exist unique words $C \in \{\epsilon\} \cup a\Omega, D \in \{\epsilon\} \cup b\Omega$ such that $T_1(\Omega) \cap T_2(\Omega) = \{C, D\}^*$.*

In the setting of Lemma 5.2 we can obviously write $C = T_1(A) = T_2(A')$, $D = T_1(B) = T_2(B')$. Writing $\Phi := (A, B), \Psi := (A', B')$ we have $T_1(x) =$

$T_2(y) \iff \exists z : x = \Phi(z), y = \Psi(z)$. Hence the pair (Φ, Ψ) parametrizes the solutions of $T_1(x) = T_2(y)$ and we call it the pair associated to (T_1, T_2). Note that $T_1\Phi = T_2\Psi$. We call T_1, T_2 compatible if the words C, D from Lemma 5.2 are non-empty or, equivalently, if Φ, Ψ are non-erasing. (A substitution T is non-erasing if $|T\sigma| \geq 1$ for each symbol σ). We write $T_1 \& T_2$ if T_1, T_2 are compatible and $T_1\underline{\&}T_2$ if they are not. The following two properties have easy proofs which are left to the reader.

Proposition 6.

  a) *Let T_1, T_2, T_3 be ABS. Then $T_1 \& T_2T_3 \Rightarrow T_1 \& T_2$;*
  b) *Let T_1, T_2, T be ABS and let (T_1, T_2) have associated pair (R, S). Then $T_1 \& T_2T \iff S \& T$.*

The set $\mathcal{M}^*$ of distinguished substitutions can be seen as a rooted binary tree where the successors of T are TA, TB. The root is then $T = (a, b)$. If $T = C_1 \cdots C_s \in \mathcal{M}^*$ with all $C_i \in \{A, B\}$, then s is called the level of T. It is well-defined since A, B generate $\mathcal{M}^*$ freely and we denote the set of elements of $\mathcal{M}^*$ with level s by $\mathcal{M}_s^*$. Now let Φ be an ABS. We write $\mathcal{G}_\Phi := \{T \in \mathcal{M}^* | T \& \Phi\}$. By Proposition 6a this is a rooted subtree of $\mathcal{M}^*$.

Lemma 5.3. *Let $\Phi \neq (a, b)$ be a reduced ABS. Then $\mathcal{G}_\Phi$ has at most one branchpoint and $|\mathcal{G}_\Phi \cap \mathcal{M}_s^*| \leq 2$ for all $s \in \mathbf{N}$.*

Lemma 5.4. *Let T_1, T_2 be distinct reduced ABS, let $K := \mathrm{Bal}$ and $L_i := F(T_iK)$. Then $P(L_1 \cap L_2, n) = o(n^3)$.*

It is clear that the itemized computations, together with Lemma 5.4, imply Theorem 5. Having said this we conclude the proof of Theorem 5 and turn to the lemmata. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Lemma 5.1. We set $G_i := G(T_i)$, the representing graph of T_i. We form the directed graph $G = G_1 \times G_2$ where $(P, Q) \to (R, S)$ is an arrow in G iff $P \to R, Q \to S$ is an arrow in G_1, G_2, respectively. We define Γ_i as in the proof of Proposition 1, hence as the collection of subpaths of $\chi_i(FK)$. We let Γ be the collection of paths $\gamma = (\gamma_1, \gamma_2)$ in G satisfying $\gamma_i \in \Gamma_i$ and $\lambda_1(\gamma_1) = \lambda_2(\gamma_2)$. In that case we call $x := \lambda_1(\gamma_1) = \lambda_2(\gamma_2)$ the label of γ. Conversely we call γ a representing path for x and we write $\Gamma(x)$ for the collection of paths in Γ representing x. Note that $x \in L_1 \cap L_2 \iff \Gamma(x) \neq \emptyset$. Now let $x \in L_1 \cap L_2$ and choose a $\gamma = \gamma(0) \cdots \gamma(n) \in \Gamma(x)$. We consider two cases.

  • $\gamma_i \neq O := (O_1, O_2)$ for $0 \leq i \leq n$. In that case γ_{i+1} is completely determined by γ_i for $0 \leq i < n$, hence γ is completely determined by

$\gamma(0)$ and its length. It follows that can find a collection Σ of $\mathbf{N}$-words, with $|\Sigma| \leq |G|$, such that each $\lambda(\gamma) \in \mathrm{Pref}(\Sigma)$. Of course we consider only those γ that do not hit O.

- $\gamma_i = O$ for some $0 \leq i \leq n$. We can write uniquely $\gamma = \gamma'\gamma''\gamma'''$ where γ'/γ''' hit O only in their last/first vertex, respectively. Since γ' is determined by $\gamma(0)$ we have $|\{\gamma'\}| \leq |G|$ and similarly $|\{\gamma'''\}| \leq |G|$. Now let $V := \{\lambda(\gamma')\} \cup \{\lambda(\gamma''')\}$ and note that $\lambda(\gamma'') \in T_1(K) \cap T_2(K)$.

$\square$

Proof of Lemma 5.2. We use the same notation as in the previous proof. Let $x \in T_1(\Omega) \cap T_2(\Omega)$ and choose a $\gamma = \gamma(0) \cdots \gamma(n) \in \Gamma(x)$ with $\gamma(0) = \gamma(n) = O$. Let $0 \leq s < t \leq n$ be consecutive elements of $\gamma^{-1}(O)$. Then $\gamma(s) \cdots \gamma(t)$ is completely determined by x_{s+1}. It is a path from O to itself that does not hit O inbetween and we call this a primitive cycle. We have at most two primitive cycles. We define $C = \lambda(\gamma)$ if a primitive cycle γ exists with initial label a and $C = \epsilon$ otherwise. Similarly we define D. Then it is clear that $C \in \{\epsilon\} \cup a\Omega, D \in \{\epsilon\} \cup b\Omega$ and that $T_1(\Omega) \cap T_2(\Omega) = \{C, D\}^*$. Now suppose that C, D have these properties. Then C equals the element of $T_1(\Omega) \cap T_2(\Omega) \cap a\Omega$ of minimal length if this set is non-empty and $C = \epsilon$ otherwise. Hence C is unique and so is D. $\square$

Proof of Lemma 5.3. Suppose that $T_1 \neq T_2$ are branchpoints. We can see T_i as words in A, B and we define $T := \mathrm{gcp}(T_1, T_2)$, the greatest common prefix. Without loss of generality, $T_2 \neq T$. Let (Φ, T) have associated pair (R, S). Since $TA, TB \& \Phi$ we have $S \& A, B$ by Proposition 6b. We will show the following two statements about S:

a) $S = ((ab)^m a, (ba)^n b)$ where $m, n \in \mathbf{N}$ with $m + n > 0$.
b) $\mathcal{G}_S = \{(a, b)\} \cup AB^* \cup BA^*$

Let G_S, G_A, G_B be the representing graphs for S, A, B, respectively. We write $\{O_A, 1\}$ for the vertex set of A and $\{O_B, 2\}$ for the vertex set of $G(B)$. Hence
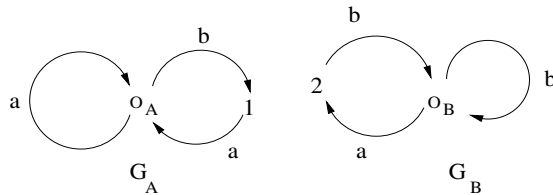
Since $S \& A$ we know that $Sa, Sb \in \mathrm{Pref}(A(\Omega))$. Hence $S = (A(au)v, A(w)x)$ where $v, x \in \{\epsilon, b\}$. Since $S \& A$ we can define $\gamma = \gamma(0) \cdots \gamma(n)$ in $G_S \times G_A$ as

the primitive cycle with initial label a. First suppose that $v = b$. The initial path $\gamma' = \gamma(0) \cdots \gamma(i)$ of γ with label $A(au)v$ then ends in $\gamma(i) = (O_S, 1)$. Let P be the unique point of G_S for which $\lambda(O_S P) = a$. Note that $P \neq O_S$ since $|Sa| > 1$. Then $\gamma(1) = \gamma(i+1) = (P, O_A)$. Hence $\gamma(1), \cdots, \gamma(n)$ has period i and since $O \notin \{\gamma(1), \cdots, \gamma(i)\}$ we find $\gamma(n) \neq O$, a contradiction. Hence $v = \epsilon$ and

$$S = (A(au), A(v)w) \text{ where } w \in \{\epsilon, b\}.$$

Similarly, $S = (B(x)y, B(bz))$ where $y \in \{\epsilon, a\}$. If $y = \epsilon$ there would exist a primitive cycle $\gamma = \gamma(0) \cdots \gamma(n)$ in $G_A \times G_B$ with initial label a. Since $\gamma(0) = (O_A, O_B)$ we obtain $\gamma(1) = (O_A, 2), \gamma(2) = (1, O_B)$ and $\gamma(3) = (O_A, 2)$. Hence $\gamma(1), \cdots, \gamma(n)$ has period 2 and does not hit O, a contradiction. Therefore $y = a$ and tracing γ as above we find $u = b^m$. Similarly $w = b, z = a^n$ and substitution yields $S = ((ab)^m a, (ba)^n b)$. If $m = n = 0$ then $S = (a, b)$ and $\Phi R = TS$ then implies $\Phi R = T$. We can write $R = R^{\mathrm{RED}} R'$ where R^{RED} is reduced and $R' \in \mathcal{M}^*$. Then ΦR^{RED} is also reduced with $k(\Phi R^{\mathrm{RED}}) =: k$. We have $\Phi R^{\mathrm{RED}}(R'(\mathcal{S}')) = T(\mathcal{S}') \subset \mathcal{S}'$ and also $\Phi R^{\mathrm{RED}}(R'(\mathcal{S}')) \subset \Phi R^{\mathrm{RED}}(\mathcal{S}') \subset \mathcal{S}_k^{\mathrm{rnp}}$. Hence $k = 1, \Phi R^{\mathrm{RED}} = (a, b)$ and $\Phi = R^{\mathrm{RED}} = (a, b)$, contradicting the hypotheses of the lemma. Hence $m + n > 0$ and a) is proved.

The pair (S, A) has associated pair $(A, (ab^m, b^{m+n+1}))$ and the pair (S, B) has associated pair $(B, (a^{m+n+1}, ba^n))$, as can be easily verified by tracing the primitive cycles in $G_S \times G_A$ and $G_S \times G_B$, respectively. Let $S' := (ab^m, b^{m+n+1})$, we will calculate $\mathcal{G}_{S'}$. We have $S' \& B^t$ for all $t \in \mathbf{N}$ since $ab^{m+t(m+n+1)}, b^{m+n+1} \in S'(\Omega) \cap B^t(\Omega)$. If (S', B^t) has associated pair (ϕ, ψ) then $\phi(b) = b, \psi(b) = b^{m+n+1}$. Since $m + n + 1 \geq 2$ we have $A \underline{\&} \psi$, hence $B^t A \underline{\&} S'$ by Proposition 6b. Combining all this with Proposition 6a we find $\mathcal{G}_{S'} = B^*$. Hence $S \& A \square \iff \square \& S' \iff \square \in B^*$ when $\square \in \mathcal{M}^*$. Similarly $S \& B \square \iff \square \in A^*$. This proves b).

We can write $T_2 = T T_3$ where $T_3 \in \mathcal{M}^*$. Since $T T_3 A, T T_3 B \& \Phi$ we have $T_3 A, T_3 B \& S$ and by b) we have $T_3 = (a, b)$. Hence $T_2 = T$, contradicting our initial hypothesis. This proves the first part of Lemma 5.3 and the second part is a direct consequence. $\square$

Proof of Lemma 5.4. If $T_1 \underline{\&} T_2$ then an element of $T_1(\Omega) \cap T_2(\Omega)$ is completely determined by its length. Hence $P(T_1(\Omega) \cap T_2(\Omega), n) \in \{0, 1\}$ for all n and $P(L_1 \cap L_2, n) = O(1)$ by Lemma 5.1. Hence we may assume that $T_1 \& T_2$. Let (T_1, T_2) have associated pair (Φ, Ψ). We note that, by Lemma 5.1, it is sufficient to show that $P(T_1(K) \cap T_2(K), n) = o(n^3)$. We will show that also Φ, Ψ are reduced.

Suppose, for instance, that $\Phi = (RS, S)$ for some R, S. Then $T_2(\Psi b) = T_1(S)$ is a suffix of $T_1(RS) = T_2(\Psi a)$. Since T_2 is reduced we find that Ψb

is a suffix of Ψa, hence $\Psi = (UV, V)$ for some U, V. We have $T_1(RS) = T_2(UV)$, $T_1(S) = T_2(V)$, hence $T_1(R) = T_2(U)$. This implies $R = \Phi(\zeta)$, $U = \Psi(\zeta)$ for some ζ, hence $R \in \{RS, S\}^*$. Since $|R| < |RS|$ we find $R \in S^*$, a contradiction since $R_1 = a, S_1 = b$. The other cases are dealt with in a similar way and we conclude that Φ, Ψ are indeed reduced. We assume wlog that $\Phi \neq (a, b)$. Let $t \in \mathbf{N}$. Since $\kappa(\mathcal{M}_t^*) = \mathcal{F}_t$ covers $[0, 1]$ we have $K = \cup_{T \in \mathcal{M}_t^*} F(TK) = \cup_{T \in \mathcal{M}_t^*} L_T$. We note that $T_1(x) = T_2(y), x, y \in \text{Bal}$ implies $x = \Phi(z), y = \Psi(z)$ for some z. Then of course $\Phi(z) \in \text{Bal}$. Hence

$$P(T_1 K \cap T_2 K, n) \leq |\{z : \Phi(z) \in \text{Bal}, |T_1\Phi(z)| = n\}|$$

$$\leq \sum_{T \in \mathcal{M}_t^*} |\{z : \Phi(z) \in L_T, |T_1\Phi(z)| = n\}|$$

$$\leq \sum_{T \in \mathcal{M}_t^*} |\{\zeta \in L_\Phi \cap L_T : |T_1(\zeta)| = n\}|.$$

If $T \underline{\&} \Phi$ then $P(L_\Phi \cap L_T, n) = O(1)$ as above, hence $|\{\zeta \in L_\phi \cap L_T : |T_1(\zeta)| = n\}| \leq |\{\zeta \in L_\Phi \cap L_T : |\zeta| \leq n\}| = \sum_{k=1}^n P(L_\Phi \cap L_T, k) = O(n)$. We now assume $T\&\Phi$: by Lemma 5.3 this can hold for at most two $T \in \mathcal{M}_t^*$. We have

$$|\{\zeta \in L_\Phi \cap L_T : |T_1(\zeta)| = n\}| \leq |\{\zeta \in L_T : |T_1(\zeta)| = n\}|$$

$$= |T_1(\{\zeta \in L_T : |T_1(\zeta)| = n\})|$$

$$\leq P(L_{T_1 T}, n).$$

Hence

$$\limsup_{n \to \infty} \frac{|\{\zeta \in L_\Phi \cap L_T : |T_1(\zeta)| = n\}|}{\text{bal}(n)} \leq \frac{1}{|T_1 T a| \cdot |T_1 T b|}$$

$$\leq \frac{1}{|Ta| \cdot |Tb|}$$

$$\leq \frac{1}{t + 1}.$$

It follows that $\limsup_{n \to \infty} \frac{P(T_1 K \cap T_2 K, n)}{\text{bal}(n)} \leq \frac{2}{t+1}$. Now let $t \to \infty$. $\qquad\square$

Final remark. We would like to finish with the following nice but impractical characterization of $\mathbf{Z}$-words of MBG: A $\mathbf{Z}$-word w has MBG iff $|\text{BS}(w)| = |\text{BF}(w)| < \infty$.

Proof. If w has MBG then $\text{MRE}_n(w) =: \{x_n\}$ for $n \geq s$. Then x_m is a suffix of x_n if $n \geq m \geq s$, hence $\text{BS}_n(w) = \text{BF}_n(w) = \emptyset$ if $n \geq s$. It follows that $P(w, n) = n + 1 + \sum_{i=0}^s (n - 1 - i)(|\text{BS}_i(w)| - |\text{BF}_i(w)|)$ for $n > s$. Since w has MBG we find that the coefficient of n equals $1 = 1 + \sum_{i=0}^s (|\text{BS}_i(w)| - |\text{BF}_i(w)|)$, hence $|\text{BS}(w)| = \sum_0^s |\text{BS}_i(w)| = \sum_0^s |\text{BF}_i(w)|$

$= |\text{BF}(w)|$. It is clear that the proof above also works in the opposite direction. $\square$

References

[Be/Po] J. Berstel, M. Pocchiola, *A geometric proof of the enumeration formula for Sturmian words.* Internat. J. Algebra Comput. **3** (1993), 394–355.

[Ca]   J. Cassaigne, *Complexité et facteurs spéciaux.* Bull. Belg. Math. Soc. **4** (1997), 67–88.

[CH]   E.M. Coven, G.A. Hedlund, *Sequences With Minimal Block Growth.* Math. Systems Th. **7** (1971), 138–153.

[FW]   N.J. Fine, H.S. Wilf, *Uniqueness theorems for periodic functions.* Proc. Amer. Math. Soc. **16** (1965), 109–114.

[H]   A. Heinis, *Arithmetics and combinatorics of words of low complexity.* Doctor's Thesis Rijksuniversiteit Leiden (2001). Available on `http://www.math.leidenuniv.nl/~tijdeman`

[L]   M. Lothaire, *Mots.* Hermès Paris 1990.

[dL/Mi] A. de Luca, F. Mignosi, *Some combinatorial properties of Sturmian words.* Theoret. Comp. Sci. **136** (1994), 361–385.

[Mi]   F. Mignosi, *On the number of factors of Sturmian words.* Theoret. Comp. Sci. **82** (1991), 71–84.

[Mi/S]   F. Mignosi, P. Séébold, *Morphismes sturmiens et règles de Rauzy.* J. Th. Nombres Bordeaux **5** (1993), 211–233.

[MH]   M. Morse, G.A. Hedlund, *Symbolic dynamics II: Sturmian trajectories.* Amer. J. Math. **62** (1940), 1–42.

[T]   R. Tijdeman, *Intertwinings of periodic sequences.* Indag. Math. **9** (1998), 113–122.

Alex Heinis
Rode Kruislaan 1403 D
1111 XD Diemen, Pays-Bas
E-mail : `alexheinis@hotmail.com`