Critical exponents of words over 3 letters

Elise Vaslet

Institut de Mathématiques de Luminy, Université Aix-Marseille II
Marseille, France

vaslet@iml.univ-mrs.fr

Submitted: Oct 5, 2010; Accepted: May 25, 2011; Published: Jun 6, 2011
Mathematics Subject Classification: 68R15

Abstract

For all $\alpha \geq RT(3)$ (where $RT(3) = 7/4$ is the repetition threshold for the 3-letter alphabet), there exists an infinite word over 3 letters whose critical exponent is α.

1 Introduction

Let A be a finite alphabet. Any finite word v over A, $v \neq \epsilon$, can be factorized as $v = p^k e$, where:

- $k \geq 1$
- e is a prefix of p
- $|p|$ is minimal

We then say that v has period p, excess e, and exponent $E(v) = |v|/|p|$. For example, the English word *church* has period *chur*, excess *ch*, and exponent 3/2, while the French word *entente* has period *ent*, excess *e*, and exponent 7/3. A (finite or infinite) word is said to be α-free (resp. α^+-free) if none of its subwords has exponent β, with $\beta \geq \alpha$ (resp. $\beta > \alpha$).

The critical exponent of an infinite word w over A is defined as

$$E_c(w) = \sup\{E(v) \in \mathbb{Q},\ v \text{ subword of } w\}.$$

For example, the binary word $abab^2 \cdots ab^n \cdots$ has critical exponent $+\infty$. The Thue-Morse word, fixed point of the morphism $0 \mapsto 01, 1 \mapsto 10$, has critical exponent 2 ([10] and [1]). The Fibonacci word, fixed point of the morphism $0 \mapsto 01, 1 \mapsto 0$, has critical exponent $2 + \phi$, where ϕ is the golden number [8].

The problem to determine if, for a given real number $\alpha > 1$, there is an infinite word w_α with critical exponent α, has been solved by Krieger and Shallit [7]. The number of letters they use to construct the word w_α grows very fast as α tends to 1, and they left open the construction of w_α over an alphabet with a fixed size.

Let k be a natural number, and let A_k be the k-letter alphabet. The repetition threshold on k letters is the real number (see [6] and [3] for more details)

$$RT(k) = \inf\{E_c(w), w \in A_k^\omega\}.$$

Dejean [6] conjectured that

$$RT(k) = \begin{cases} 2 & \text{if } k = 2 \\ 7/4 & \text{if } k = 3 \\ 7/5 & \text{if } k = 4 \\ k/(k-1) & \text{if } k > 4 \end{cases}$$

and this conjecture is now proved (see [9] and [5]).

It is clear that if $\alpha < RT(k)$, no word over A_k has critical exponent α. Currie and Rampersad [4] proved the following result for a binary alphabet:

> For each $\alpha \geq 2 = RT(2)$, there is an infinite binary word with critical exponent α.

And they conjectured:

> Let $k \geq 2$. For each $\alpha \geq RT(k)$, there is an infinite word over k letters with critical exponent α.

We will prove that this is true for $k = 3$. Let $A_3 = \{a, b, c\}$ be the 3-letter alphabet. Dejean [6] proved that $RT(3) = 7/4$. We immediately remark that if $\alpha \geq 2$, the result of Currie and Rampersad gives us a binary word, and then also a word over A_3, with critical exponent α. That's why we only have to consider the case $7/4 \leq \alpha < 2$. To demonstrate her result, Dejean considered the morphism μ (which we will call Dejean's morphism) defined as follows:

$$\mu : \begin{cases} a \mapsto abc\ acb\ cab\ c\ bac\ bca\ cba \\ b \mapsto \pi(\mu(a)) = bca\ bac\ abc\ a\ cba\ cab\ acb \\ c \mapsto \pi^2(\mu(a)) = cab\ cba\ bca\ b\ acb\ abc\ bac \end{cases}$$

where π is the permutation $(a\ b\ c)$, and proved that its fixed point $\mu^\infty(a)$ has critical exponent $7/4$.

2 Exponents and Dejean's morphism

Dejean [6] noticed the existence of specific subwords in $\mu(a)$, $\mu(b)$, and $\mu(c)$, which can be used to desubstitute μ. She called them characteristic factors:

Proposition 1. *The words* $f_1 = abcacbc$, $f_2 = cabcbac$ *and* $f_3 = cbcacba$ *only appear in* $\mu(A_3^*)$ *respectively as prefix, central factor, and suffix, of* $\mu(a)$.
Similarly, $\pi(f_1)$, $\pi(f_2)$, $\pi(f_3)$ *only appear as prefix, central factor, and suffix of* $\mu(b)$, *and* $\pi^2(f_1)$, $\pi^2(f_2)$, $\pi^2(f_3)$ *as prefix, central factor, and suffix of* $\mu(c)$.

Definition 1. *The words* f_1, f_2, *and* f_3 *(resp.* $\pi(f_1)$, $\pi(f_2)$, $\pi(f_3)$ *; resp.* $\pi^2(f_1)$, $\pi^2(f_2)$, $\pi^2(f_3)$*) are called characteristic factors of* $\mu(a)$ *(resp.* $\mu(b)$*; resp.* $\mu(c)$*).*

We use these characteristic factors to prove the following desubstitution results for μ.

Proposition 2. *Let* w *be a word over* A_3, *and* u *a subword of* $\mu(w)$. *If* $|u| \geq 18$, *then, there exists a unique* $x \in A_3$, *a unique* $y \in A_3$, *and there exist some unique* $s, v, p \in A_3^*$, *such that* $u = s\mu(v)p$, *where* $s \neq \epsilon$ *is a suffix of* $\mu(x)$, *and* $p \neq \epsilon$ *is a prefix of* $\mu(y)$.

Proof. As $|u| \geq 18$, u has a characteristic factor as a subword. The result is then clear since μ is a 19-uniform morphism, and since $\forall x \in A_3$, $\mu(x)$ begins and ends with x. □

Theorem 1. *Let* $w \in A_3^*$, *and let* u *be a subword of* $\mu(w)$. *Assume* u *has period* p, *excess* e *and exponent* $7/4 < |u|/|p| < 2$. *Then,* w *has a subword* v *of length* $|v| \geq \lceil |u|/19 \rceil$, *with period* q *such that* $|q| = \lceil |p|/19 \rceil$, *and with exponent* $E(v) \geq |u|/|p|$.

Proof. There are two cases: either $|e| \geq 18$, or $|e| < 18$.

Suppose first that $|e| \geq 18$. Then, we also have $|u| \geq 18$. Then, by Proposition 2,

- There exist some unique $x_u, y_u \in A_3$, and some unique $m_u, s_u, p_u \in A_3^*$, $s_u \neq \epsilon$ suffix of $\mu(x_u)$, p_u prefix of $\mu(y_u)$, such that $u = s_u\mu(m_u)p_u$,
- There exist some unique $x_e, y_e \in A_3$, and some unique $m_e, s_e, p_e \in A_3^*$, $s_e \neq \epsilon$ suffix of $\mu(x_e)$, p_e prefix of $\mu(y_e)$, such that $e = s_e\mu(m_e)p_e$.

Let:

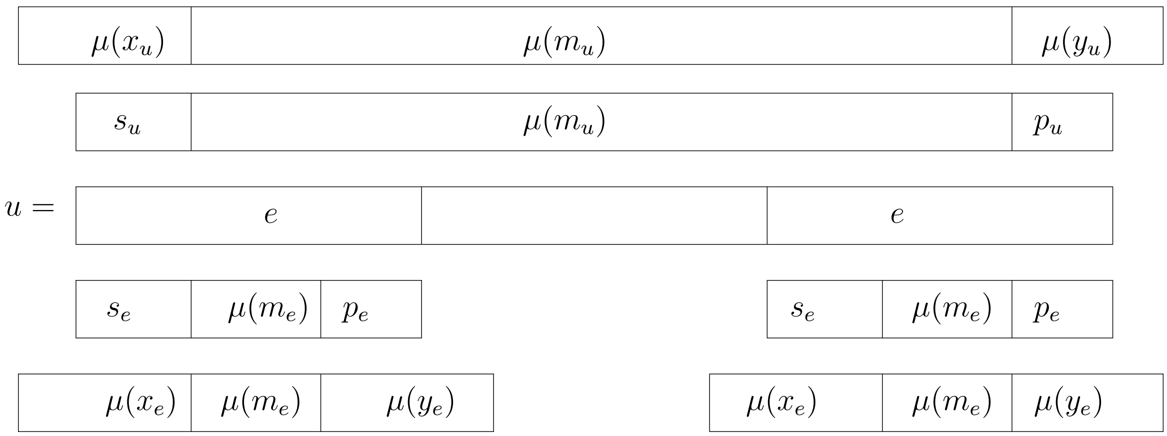
$$f = x_e m_e y_e$$

$$v = x_u m_u y_u$$

As e is a suffix of u, we have: $p_u = p_e$, $y_u = y_e$, and $\mu(x_e)\mu(m_e)$ is a suffix of $\mu(x_u)\mu(m_u)$, thus $x_e m_e$ is a suffix of $x_u m_u$. Moreover, e is a prefix of u, so $s_u = s_e$, $x_e = x_u$, $\mu(m_e)\mu(y_e)$ is a prefix of $\mu(m_u)\mu(y_u)$, and $m_e y_e$ is a prefix of $m_u y_u$ (see the following figure).

Therefore, f is a prefix and a suffix of v, and v has excess f ($|f|$ is maximal, otherwise $|e|$ would not be maximal in u). Let us denote its period by q, i.e. $v = qf$.
It is clear that

$$|v| \geq \lceil \frac{|u|}{19} \rceil.$$



Moreover, q has length

$$\begin{aligned}
|q| &= |v| - |f| = |m_u| - |m_e| \\
&= \frac{|\mu(m_u)| - |\mu(m_e)|}{19} \\
&= \frac{(|u| - |s_u| - |p_u|) - (|e| - |s_e| - |p_e|)}{19} \\
&= \frac{|u| - |e|}{19} \text{ (because } s_u = s_e \text{ and } p_u = p_e) \\
&= \frac{|p|}{19}
\end{aligned}$$

Finally, v has exponent

$$E(v) = \frac{|v|}{|q|} \geq \lceil \frac{|u|}{19} \rceil / \frac{|p|}{19} \geq \frac{|u|}{19} . \frac{19}{|p|} = E(u)$$

Thus, v is the word we were looking for.

Suppose now that $|e| < 18$. Then, as $E(u) > 7/4$, $|u| \leq 24 + 18 = 42$. As μ is 19-uniform, u is a subword of a word $\mu(x)$, where $x \in A_3^*$, $|x| \leq 4$. Moreover, $E(u) > 7/4$, so x is not a subword of $\mu^\infty(a)$. Then, by looking at the $\mu(x)$ obtained if x is not a subword of $\mu^\infty(a)$, we can again reduce the set of possible x:

$$x \in \{aa, bb, cc, aaa, bbb, ccc, aaaa, bbbb, cccc, abab, acac, baba, bcbc, caca, cbcb\}$$

since in the other cases, $\mu(x)$ has no subword u such that $|u| \leq 42$ and $E(u) > 7/4$. Then, we have:

- if u is a subword of $\mu(x)$, with $x \in \{aa, bb, cc\}$, consider $v = x$. v has exponent 2, and period q of length $|q| = 1$. We can also remark that u necessarily has a period of length $19 = 19.|q|$, and has exponent $E(u) \leq 2 = E(v)$. Therefore, v is the word we were looking for.

- if u is a subword of $\mu(x)$, with $x \in \{aaa, bbb, ccc\}$, consider $v = x$. v has exponent 3, and period q of length $|q| = 1$. As u necessarily has a period of length $19 = 19.|q|$, and has exponent $E(u) \leq 3 = E(v)$, v is the word we were looking for.

- if u is a subword of $\mu(x)$, with $x \in \{aaaa, bbbb, cccc\}$, consider $v = x$. v has exponent 4, and period q of length $|q| = 1$. As u necessarily has a period of length $19 = 19.|q|$, and has exponent $E(u) \leq 4 = E(v)$, v is the word we were looking for.

- finally, if u is a subword of $\mu(x)$, with $x \in \{abab, acac, baba, bcbc, caca, cbcb\}$, consider $v = x$. v has exponent 2, and period q of length $q = 2$. As u necessarily has a period of length $2.19 = 19.|q|$, and has exponent $E(u) \leq 2 = E(v)$, v is the word we were looking for.

□

3 Construction of an infinite α-free word over A_3

In the following, we will use the operator, denoted by δ, that removes the first letter of a word: for example, $\delta(0110) = 110$.

Lemma 1. *Let $\mathfrak{L}$ be the set $Fact(\mu(A_3^*))$ of all subwords of the words in $\mu(A_3^*)$. Let $\alpha \in]7/4, 2[$ and $v \in A_3^*$, such that:*

- *$abcbabcv \in \mathfrak{L}$*

- *$abcbabcv$ is α-free.*

Suppose that $babcbabcv = xuy$, where u has exponent $E(u) \geq \alpha$. Then, $x = \epsilon$, and $u = babcbabc$.

Proof. By hypothesis, $abcbabcv$ is α-free. Since $E(u) \geq \alpha$, u is necessarily a prefix of $babcbabcv$. Moreover, $babcbab$ has exponent $7/4 < \alpha \leq E(u)$. Therefore, $u = babcbabcv'$, where v' is a prefix of v. Suppose that $v' \neq \epsilon$.
By hypothesis, $abcbabcv' \in \mathfrak{L}$. Moreover, the subword $abcbabc$ only appears in $\mathfrak{L}$ as a subword of $\mu(c)$. So $v' = abacbabcbac....$
The excess of u is at most $babcbab$. Indeed, otherwise, the word $babcbabc$, whose exponent is 2, is a subword of $abcbabcv$, which is impossible since $abcbabcv$ is α-free. Then, u has excess e, with $|e| \leq 7$, and so has period p with $|p| \leq 9$. So u is a subword of $babcbabcabacbabc$. By looking at the factors of this word, we deduce that the only possibility is $u = babcbabc$. □

Lemma 2. *Let $\alpha \in]7/4, 2[$ be given. Let s, t be natural numbers such that $\mu^s(b) = xabcbabcy$, with $|x| = t$. Let $\beta = 2 - \frac{t}{4.19^s}$. Suppose that $7/4 < \beta < \alpha$, and that $abcbabcv \in \mathfrak{L}$ is α-free. Consider the word $w = \delta^t\mu^s(babcbabcv)$. Then, we have:*

1. *w has a prefix with exponent β.*

2. *If abcbabcv has a subword with exponent γ and period p, then, w has a subword with exponent γ and a period of length $19^s|p|$.*

3. *w is α-free.*

Proof. 1. $\mu^s(babcbabc)$ has exponent 2 and period $\mu^s(babc)$. We have $|\mu^s(babc)| = 4.19^s$, and $|\mu^s(babcbabc)| = 8.19^s$, so the prefix $\delta^t\mu^s(babcbabc)$ of w has exponent

$$\frac{|w|}{|\mu^s(babc)|} = \frac{|\mu^s(babcbabc)| - t}{|\mu^s(babc)|} = \beta$$

2. Let u be a subword of $abcbabcv$, with exponent γ and period p. Then $\mu^s(u)$ is a subword of $\mu^s(abcbabcv)$, with exponent γ and period $19^s|p|$. Moreover, $\mu^s(abcbabcv)$ is a suffix of $\delta^t\mu^s(babcbabcv)$, since $t = |x| \leq |\mu^s(b)|$. So $\mu^s(u)$ is a subword of w, with the required properties.

3. Suppose that w has a subword u, with exponent $\kappa \geq \alpha$ and period p. Then, by iteration of Theorem 2, $babcbabcv$ has a subword u' with exponent $\kappa' \geq \kappa$ and q such that $|q| = \frac{|p|}{19^s}$. By Lemma 1, as $\kappa' \geq \alpha$, we deduce that $\kappa' = 2$ and that $u' = babcbabc$. Then $q = babc$, and $|p| = |q|.19^s = 4.19^s$.
Moreover, u is not a subword of $\mu^s(abcbabcv)$, otherwise $abcbabcv$ has a subword with exponent $\geq \alpha$. u is not a subword of $\delta^t\mu^s(babcbabc)$ either, otherwise, we would have $|u| \leq |\delta^t\mu^s(babcbabc)| = 8.19^s - t$, and so:

$$E(u) = \frac{|u|}{|p|} \leq \frac{8.19^s - t}{4.19^s} = \beta < \alpha,$$

which is absurd since $E(u) \geq \alpha$.
Therefore, u has a prefix z such that $z = z_1\mu^s(abcbabc)z_2$, where $z_1 \neq \epsilon$ is a suffix of $\mu^s(b)$, and $z_2 \neq \epsilon$ is a prefix of $\mu^s(a)$ (since we remarked that the first letter of v is a a). z being a prefix of u, z has a period of length 4.19^s. Then, z has period $z_1\mu^s(abc)z_1'$, with z_1' such that $\mu^s(b) = z_1'z_1$. So we have:

$$z = z_1\mu^s(abc)z_1'z_1\mu^s(abc)z_2.$$

We deduce that either z_2 is a prefix of z_1', or z_1' is a prefix of z_2. Yet neither is possible. Indeed, z_2 begins with the letter a, and z_1' begins with the letter b.
Finally, w is α-free.

□

4 A word over A_3 with critical exponent $\alpha \geq RT(3)$

Definition 2. *A real number $\beta < \alpha$ is said to be obtainable if β can be written as $\beta = 2 - \frac{t}{4.19^s}$, where the natural numbers s and t verify:*

- $s \geq 3$

- *the word* $\delta^t(\mu^s(b))$ *begins with abcbabc.*

We note that for any given $s \geq 3$, it is possible to choose t such that

- $7/4 < \beta = 2 - \frac{t}{4.19^s} < \alpha$
- $|\alpha - \beta| \leq \frac{19^2}{4.19^s}$

Indeed, $\mu^2(a)$, $\mu^2(b)$, and $\mu^2(c)$ have length 19^2, and each have *abcbabc* as a subword. Therefore, choosing a large enough s, we can always find some obtainable real numbers β, arbitrarily close to α.

Theorem 2. *Let* $\alpha \geq RT(3) = 7/4$. *Then, there is an infinite word over* A_3 *with critical exponent* α.

Proof. If $\alpha = 7/4$, we already know that $\mu^\infty(a)$ has critical exponent $7/4$. If $\alpha \geq 2$, by the theorem for $k = 2$, we can find a word over A_3^* with critical exponent α. Now, let $\alpha \in]7/4, 2[$.
Let $(\beta_i)_{i \in \mathbb{N}}$ be an increasing sequence of obtainable numbers, converging to α. For each i, we write β_i as:

$$\beta_i = 2 - \frac{t_i}{4.19^{s_i}}$$

where s_i and t_i are such that:

- $s_i \geq 3$
- $\delta^{t_i}\mu^{s_i}(b)$ begins with *abcbabc*.

For all words $v \in \mathfrak{L}$, let $\Phi_i(v)$ denote the word $\delta^{t_i}\mu^{s_i}(bv)$, and consider the following sequence:

$$\begin{array}{rcl}
v_1 & = & \Phi_1(abcbabc) = \delta^{t_1}\mu^{s_1}(babcbabc) \\
v_2 & = & \Phi_1\Phi_2(abcbabc) = \delta^{t_1}\mu^{s_1}(b\delta^{t_2}\mu^{s_2}(babcbabc)) \\
v_3 & = & \Phi_1\Phi_2\Phi_3(abcbabc) \\
\vdots & & \\
v_n & = & \Phi_1\Phi_2\Phi_3...\Phi_n(abcbabc) \\
\vdots & &
\end{array}$$

By iteration of Lemma 2, as $abcbabc \in \mathfrak{L}$ is α-free, we deduce that each v_i is α-free. Moreover, once again by Lemma 2, each v_i has a subword with exponent β_j, $j = 1, 2, ..., i$. Finally, consider the word $w = \lim_{n\to\infty} \in A_3^\omega$ (it is possible to take this limit since each v_i is a prefix of v_{i+1}). w then has critical exponent α: it is α-free, yet has subwords with exponents β_i converging to α. □

The conjecture proposed by Currie and Rampersad in [4] is then true for alphabets of size 2 and 3. It still have to be proved for alphabets of size ≥ 4. For that, another method must be found, because of Brandenburg's result in [2] : if $k \geq 4$, there is no $RT(k)$-free morphism, i.e., no morphism which maps, as Thue-Morse morphism for $k = 2$ or Dejean's morphism for $k = 3$, every $RT(k)$-free word to an $RT(k)$-free word.

References

[1] J. Berstel. Axel Thue's papers on repetitions in words: a translation. *Publications du LaCIM*, 20, 1995.

[2] F. J. Brandenburg. Uniformly growing k-th powerfree homomorphisms. *Theoret. Comput. Sci.*, 23:69–82, 1983.

[3] A. Carpi. On Dejean's conjecture over large alphabets. *Theoret. Comput. Sci.*, 385:137–151, 2007.

[4] J. D. Currie and N. Rampersad. For each $\alpha > 2$ there is an infinite binary word with critical exponent α. *Electron. J. Combin.*, 15:N34, 2008.

[5] J. D. Currie and N. Rampersad. A proof of Dejean's conjecture. *Preprint*, http://arxiv.org/abs/0905.1129, 2009.

[6] F. Dejean. Sur un théorème de Thue. *J. Comb. Theory A*, 13:90–99, 1972.

[7] D. Krieger and J. Shallit. Every real number greater than 1 is a critical exponent. *Theoret. Comput. Sci.*, 381:177–182, 2007.

[8] F. Mignosi and G. Pirillo. Repetitions in the Fibonacci infinite word. *RAIRO Inform. Theor. Appl.*, 26:199–204, 1992.

[9] M. Rao. Last cases of Dejean's conjecture. *WORDS 2009*, http://www.labri.fr/perso/rao/publi/dejean.ps, 2009.

[10] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske Vid. Selsk. Skr. Mat. Nat. Kl.*, 10:1–67, 1912.