Theoretical Computer Science 410 (2009) 2885–2888

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

# Dejean's conjecture holds for $n \geq 30$

James Currie *, Narad Rampersad

*Department of Mathematics and Statistics, University of Winnipeg, 515 Portage Ave., Winnipeg, MB R3B 2E9, Canada*

ARTICLE INFO

To Juhani Karhumäki on the occasion of his 60th birthday

*Keywords:*
Repetitive threshold
Fractional power
Dejean's conjecture

ABSTRACT

We extend Carpi's results by showing that Dejean's conjecture holds for $n \geq 30$.


## 1. Introduction

Repetitions in words have been studied starting with Thue [10,11] at the beginning of the previous century. Much study has also been conducted on repetitions with fractional exponent [1–6]. If $n > 1$ is an integer, then an ***n*-power** is a non-empty word $x^n$, i.e., word $x$ repeated $n$ times in a row. For rational $r > 1$, a **fractional *r*-power** is a non-empty word $w = x^{\lfloor r \rfloor}x'$ such that $x'$ is the prefix of $x$ of length $(r - \lfloor r \rfloor)|x|$. For example, 01010 is a 5/2-power. A basic problem is that of identifying the repetitive threshold for each alphabet size $n > 1$:

> What is the infimum of $r$ such that an infinite sequence on $n$ letters exists, not containing any factor of exponent greater than $r$?

We call this infimum the **repetitive threshold** of an $n$-letter alphabet, denoted by $RT(n)$. Dejean's conjecture [3] is that

$$RT(n) = \begin{cases} 7/4, & n = 3 \\ 7/5, & n = 4 \\ n/(n-1) & n \neq 3, 4 \end{cases}$$

The values $RT(2)$, $RT(3)$, $RT(4)$ were established by Thue, Dejean and Pansiot, respectively [11,3,9]. Moulin-Ollagnier [8] verified Dejean's conjecture for $5 \leq n \leq 11$, while Mohammad-Noori and Currie [7] proved the conjecture for $12 \leq n \leq 14$.

An exciting new development has recently occurred with the work of Carpi [2], who showed that Dejean's conjecture holds for $n \geq 33$. Verification of the conjecture is now only lacking for a finite number of values. In the present paper, we sharpen Carpi's methods to show that Dejean's conjecture holds for $n \geq 30$.

## 2. Preliminaries

The following definitions are from Sections 8 and 9 of [2]: Fix $n \geq 30$. Let $m = \lfloor (n-3)/6 \rfloor$. Let $A_m = \{1, 2, \ldots, m\}$. Let $\ker \psi = \{v \in A_m^* \mid \forall a \in A_m, 4 \textit{ divides } |v|_a\}$. (In fact, this is not Carpi's *definition* of $\ker \psi$, but rather the assertion of

* Corresponding author.
*E-mail addresses:* j.currie@uwinnipeg.ca (J. Currie), n.rampersad@uwinnipeg.ca (N. Rampersad).

0304-3975/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2009.01.026

his Lemma 9.1.) A word $v \in A_m^+$ is a $\psi$-**kernel repetition** if it has period $q$ and a prefix $v'$ of length $q$ such that $v' \in \ker \psi$, $(n-1)(|v|+1) \geq nq-3$.

It will be convenient to have the following new definition: If $v$ has period $q$ and its prefix $v'$ of length $q$ is in $\ker \psi$, we say that $q$ is a **kernel period** of $v$.

As Carpi states at the beginning of section 9 of [2]:

> By the results of the previous sections, at least in the case $n \geq 30$, in order to construct an infinite word on $n$ letters avoiding factors of any exponent larger than $n/(n-1)$, it is sufficient to find an infinite word on the alphabet $A_m$ avoiding $\psi$-kernel repetitions.

For $m = 5$, Carpi produces such an infinite word, based on a paper-folding construction. He thus establishes Dejean's conjecture for $n \geq 33$. In the present paper, we give an infinite word on the alphabet $A_4$ avoiding $\psi$-kernel repetitions. We thus establish Dejean's conjecture for $n \geq 30$.

**Definition 1.** Let $f : A_4^* \to A_4^*$ be defined by $f(1) = 121, f(2) = 123, f(3) = 141, f(4) = 142$. Let $w$ be the fixed point of $f$.

It is useful to note that the frequency matrix of $f$, i.e.,

$$[|f(i)|_j]_{4\times 4} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

has an inverse modulo 4.

**Remark 1.** Let $q$ be a non-negative integer, $q \leq 1966$. Fix $n = 32$.

R1: Word $w$ contains no $\psi$-kernel repetition $v$ with kernel period $q$.
R2: Word $w$ contains no factor $v$ with kernel period $q$ such that $|v|/q \geq 35/34$.

The defining condition on $\psi$-kernel repetitions, namely that

$$(n-1)(|v|+1) \geq nq - 3,$$

can be rewritten as

$$\frac{|v|}{q} \geq \frac{n}{n-1} - \frac{n+2}{q(n-1)}. \tag{1}$$

Note that $\frac{32}{31} - \frac{34}{31q} = \frac{35}{34}$ when $q = \frac{34^2}{3} = 385\frac{1}{3}$, so neither piece of the remark implies the other. One verifies that

$$\frac{35}{34} + \frac{9}{2(1967)} \leq \frac{32}{31} - \frac{34}{31q}$$

for $q \geq 1967$. Since the right-hand side of (1) decreases with $n$, once the remark is established, it will remain true if $n$ is replaced by 30 or 31. To show that $w$ contains no $\psi$-kernel repetitions for $n = 30, 31, 32$, it thus suffices to verify R1 and to show that word $w$ contains no factor $v$ with kernel period $q \geq 1967$ such that

$$|v|/q \geq 35/34 + 9/2(1967). \tag{2}$$

The remarks R1 and R2 are verified by computer search, so we will consider the second part of this attack. Fix $q \geq 1967$, and suppose that $v$ is a factor of $w$ with kernel period $q$, and $|v|/q \geq 35/34$. Since $w$ is not ultimately periodic, without loss of generality, suppose that no extension of $v$ has period $q$. Write $v = sf(u)p$ where $s$ (resp. $p$) is a suffix (resp. prefix) of the image of a letter, and $|s|$ ( resp. $|p|$) $\leq 2$.

If $|v| \leq q+2$, then $35/34 \leq (q+2)/q$ and $1/34 \leq 2/q$, forcing $q \leq 68$. This contradicts R2. We will therefore assume that $|v| \geq q+3$.

Suppose $|s| = 2$. Since $|v| \geq q+3$, write $v = s1zs1v'$, where $|s1z| = q$. Examining $f$, we see that the letter $a_s$ preceding any occurrence of $s1$ in $w$ is uniquely determined if $|s| = 2$. It follows that $a_s v$ is a factor of $w$ with kernel period $q$, contradicting the maximality of $v$. We conclude that $|s| \leq 1$.

Again considering $f$, we see that if $t$ is any factor of $w$ of length 3, and $u_1t, u_2t$ are prefixes of $w$, then $|u_1| \equiv |u_2| \pmod 3$. Since $|v| \geq q+3$, we conclude that 3 divides $q$. Write $q = 3q_0$. Since $|s| \leq 1$, $|p| \leq 2$ and $|v| \geq q+3$, we see that $|f(u)| \geq q$. Thus $f(u)$ has a prefix of length $q = 3q_0$ which is in $\ker \psi$. As the frequency matrix of $f$ is invertible modulo 4, the prefix of $u$ of length $q_0$ is in $\ker \psi$. We see that

$$\frac{|v|}{q} \leq \frac{3|u|+3}{3q_0} = \frac{|u|}{q_0} + \frac{1}{q_0}.$$

**Lemma 2.** *Let s be a non-negative integer. If factor $v$ of $w$ has kernel period q, where $q \leq 1966(3^s)$, then*

$$\frac{|v|}{q} < \frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{s-1} 3^{-j}.$$

**Proof.** If $s = 0$, this is implied by R2. Suppose $t > 0$ and the result holds for $s < t$. Suppose that $1966(3^{t-1}) < q \leq 1966(3^t)$ and there is a factor $v$ of $w$ such that $v$ has kernel period $q$. Suppose that $|v|/q \geq 35/34$. Without loss of generality, suppose that no extension of $v$ has period $q$. We have seen that there is a factor $u$ of $w$ with kernel period $q_0 = q/3$, $1966(3^{t-2}) < q_0 \leq 1966(3^{t-1})$ such that

$$\begin{aligned}
|v|/q &\leq |u|/q_0 + 1/q_0 \\
&< \left(\frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{t-2} 3^{-j}\right) + \frac{1}{q_0} \quad \text{(by the induction hypothesis)} \\
&< \frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{t-2} 3^{-j} + \frac{1}{1966(3^{t-2})} \\
&= \frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{t-2} 3^{-j} + \frac{3}{1966(3^{t-1})} \\
&= \frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{t-1} 3^{-j}. \quad \square
\end{aligned}$$

**Theorem 3.** *Word $w$ contains no factor $v$ with kernel period q such that*

$$|v|/q \geq 35/34 + 9/2(1966).$$

**Proof.** Suppose that factor $v$ of $w$ has kernel period $q$ such that (2) holds. By Remark 1, we have $q \geq 1966$. By the previous lemma, for some non-negative $s$,

$$|v|/q < \frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{s-1} 3^{-j} < \frac{35}{34} + \frac{3}{1966}\sum_{j=0}^{\infty} 3^{-j} = \frac{35}{34} + \frac{9}{2(1966)}. \quad \square$$

We may now build on Carpi's result, here restated as a theorem:

**Theorem 4.** *Fix $n \geq 30$. To show that there is an infinite word on n letters avoiding factors of any exponent larger than $n/(n-1)$, it is sufficient to find an infinite word on the alphabet $A_m$ avoiding $\psi$-kernel repetitions.*

**Corollary 5.** *Dejean's conjecture holds for $n = 30, 31, 32$.*

The restriction $n \geq 30$ in the theorem results from Carpi's approach to avoiding the so-called 'short repetitions'. (See [8].) Therefore, our result in some sense optimizes his construction.

## Acknowledgements

The first author is supported by an NSERC Discovery Grant. The second author is supported by an NSERC Postdoctoral Fellowship.

## Appendix. Computer search

Suppose that some factor $v$ of $w$ has kernel period $q \leq 1966$ and either $31(|v| + 1) \geq 32q - 3$ or $|v|/q \geq 35/34 + 9/2(1967)$. Without loss of generality, taking such a $v$ as short as possible, we may assume that

$$|v| \leq \left\lceil \frac{32(1966) - 3}{31} - 1 \right\rceil = 2029.$$

$$\text{(We also have } \left\lceil (1966)\left(\frac{35}{34} + \frac{9}{2(1967)}\right)\right\rceil = 2029.\text{)}$$

If $|v| > 3$, $v$ is a factor of $f(u)$ for some factor $u$ of $w$ with $|u| \leq (|v|+4)/3$. For a non-negative integer $r$, let $g(r) = \lfloor (r+4)/3 \rfloor$. Since $g^7(2029) = 2 < 3$, (here the exponent denotes iterated function composition) word $v$ must be a factor of $f^7(u)$ for some factor $u$ of $w$, $|u| = 2$.

The word $u_0 = 23141121142$ contains all 8 factors of $w$ which have length 2. To establish R1 and R2, one thus checks that they hold for the single word $f^7(u_0)$ (which is of length 24,057).

## Note added in proof

We have now improved the result to $n \geq 27$, as will appear in a forthcoming paper.

## References

[1] F.J. Brandenburg, Uniformly growing *k*-th powerfree homomorphisms, Theoret. Comput. Sci. 23 (1983) 69–82.
[2] A. Carpi, On Dejean's conjecture over large alphabets, Theoret. Comput. Sci. 385 (2007) 137–151.
[3] F. Dejean, Sur un théorème de Thue, J. Combin. Theory Ser. A 13 (1972) 90–99.
[4] L. Ilie, P. Ochem, J Shallit, A generalization of repetition threshold, Theoret. Comput. Sci. 345 (2005) 359–369.
[5] D. Krieger, On critical exponents in fixed points of non-erasing morphisms, Theoret. Comput. Sci. 376 (2007) 70–88.
[6] F. Mignosi, G. Pirillo, Repetitions in the Fibonacci infinite word, RAIRO Inform. Théor. Appl. 26 (1992) 199–204.
[7] M. Mohammad-Noori, J.D. Currie, Dejean's conjecture and Sturmian words, European J. Combin. 28 (2007) 876–890.
[8] J. Moulin-Ollagnier, Proof of Dejean's conjecture for alphabets with 5, 6, 7, 8, 9, 10 and 11 letters, Theoret. Comput. Sci. 95 (1992) 187–205.
[9] J.-J. Pansiot, A propos d'une conjecture de F. Dejean sur les répétitions dans les mots, Discrete Appl. Math. 7 (1984) 297–311.
[10] A. Thue, Über unendliche Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiana 7 (1906) 1–22.
[11] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiana 1 (1912) 1–67.