On the number of episturmian palindromes

Michelangelo Bucci[1], Aldo de Luca[1], and Alessandro De Luca[2]

[1]Dipartimento di Matematica e Applicazioni "R. Caccioppoli"
Università degli Studi di Napoli Federico II
Via Cintia, Monte S. Angelo, I-80126 Napoli, Italy
{micbucci,aldo.deluca}@unina.it

[2]LaCIM, Université du Québec à Montréal
Case postale 8888, succursale Centre-ville
Montréal (Québec) H3C 3P8, Canada
de_luca.alessandro@courrier.uqam.ca

May 27, 2010

Abstract

Episturmian words are a suitable generalization to arbitrary alphabets of Sturmian words. In this paper we are interested in the problem of enumerating the palindromes in all episturmian words over a k-letter alphabet A_k. We give a formula for the map g_k giving for any n the number of all palindromes of length n in all episturmian words over A_k. This formula extends to $k > 2$ a similar result obtained for $k = 2$ by the second and third author in 2006. The map g_k is expressed in terms of the map P_k counting for each n the palindromic prefixes of all standard episturmian words (epicentral words). For any $n \geq 0$, $P_2(n) = \varphi(n+2)$ where φ is the totient Euler function. The map P_k plays an essential role also in the enumeration formula for the map λ_k counting for each n the finite episturmian words over A_k. Similarly to Euler's function, the behavior of P_k is quite irregular. The first values of P_k and of the related maps g_k, and λ_k for $3 \leq k \leq 6$ have been calculated and reported in the Appendix. Some properties of P_k are shown. In particular, broad upper and lower bounds for P_k, as well as for $\sum_{m=0}^{n} P_k(m)$ and g_k, are determined. Finally, some conjectures concerning the map P_k are formulated.

Keywords: Episturmian words; Sturmian words; Epicentral words; Palindromes; Palindromization map.

1 Introduction

Since 2001, many papers have been written on the subject of episturmian words a suitable generalization to arbitrary alphabets of Sturmian words which are on a two-letter alphabet (see, for instance, [2, 10]). A generalization of standard Sturmian words was obtained in [9] by extending in a natural way to an arbitrary alphabet a palindromization map, introduced in [4], capable to generate all standard Sturmian words. In this way one obtains the class of standard episturmian words. A word is episturmian if it has the same set of factors as a standard one. We observe that in the passage from Sturmian to episturmian words some properties (such as the balance property) are lost, and others are saved.

As a consequence of the above construction, one has that episturmian words have many palindromic factors. In fact, episturmian words are rich in palindromes, i.e., any factor has the maximum possible number of palindromic factors [9]. A special class of palindromes are the palindromic prefixes of the standard episturmian words, called epicentral words in analogy with the term central used in the case of Sturmian words (see, for instance, [13]). Epicentral words satisfy several interesting combinatorial properties (see, for instance, [7, 8]).

In this paper we shall consider the problem of enumerating the palindromes in all episturmian words over a k-letter alphabet A_k. We denote by g_k the map counting for any n all palindromes of length n which are factors of episturmian words over A_k. In the case of Sturmian words, i.e., $k = 2$, a formula for g_2 was given in [5]; the map g_2 depends on the enumeration map P_2 of central words. As is well known [6] for any $n \geq 0$, $P_2(n) = \varphi(n + 2)$ where φ is the totient Euler function.

For any k, let us denote by P_k the enumeration map of all epicentral words over A_k. In Section 5 we shall give a formula (see Theorem 5.3) expressing $g_k(n)$ in terms of the values $P_k(n - 2i)$ with $1 \leq i \leq \lceil n/2 \rceil - 1$. This formula extends to arbitrary alphabets the result obtained in [5] in the case $k = 2$.

We stress that the preceding proof requires some remarkable structural properties of episturmian words and epicentral words which are proved in Section 4. In particular, it is proved (cf. Theorem 4.3) that the set PER_k of all epicentral words over A_k equals the set SBS_k of all strictly bispecial factors of the set $\mathrm{Fact}\, EP_k$ of all factors of the episturmian words over A_k. (A word w is strictly bispecial in $\mathrm{Fact}\, EP_k$ if for all $x, y \in A_k$, $xwy \in \mathrm{Fact}\, EP_k$.) The elements of $\mathrm{Fact}\, EP_k$ are usually called finite episturmian words over A_k.

The map P_k plays an essential role also in the enumeration formula for the map λ_k counting for each n the finite episturmian words over A_k of length n. In Section 6 we provide a correct version of the formula given in [14]; for the sake of completeness, we essentially repeat the arguments considered there with a more accurate proof which uses the preceding structural properties of epicentral words.

In Section 7 canonical epicentral words are introduced. If the alphabet $A_k = \{1, \ldots, k\}$ is totally ordered by setting $1 < 2 < \cdots < k$, an epicentral word w is canonical if the order of the first occurrences of the letters in w is the same as the order of A_k. Any epicentral word can be obtained from a canonical one by a word isomorphism (i.e., a renaming of its letters). Some properties of the canonical epicentral words which are of interest for counting the epicentral words are shown.

The enumeration map P_k is studied in Section 8. The first values of P_k and of the related maps g_k, and λ_k for $3 \leq k \leq 6$ are respectively reported in Tables 1, 2, and 4 of the Appendix. In Tab. 3 are reported for $3 \leq k \leq 6$ the first values of the map U_k counting for each n the number of canonical epicentral words of length n having exactly k distinct letters. A quite efficient algorithm to compute P_k is briefly described in Section 8.3.

Similarly to Euler's function the behavior of P_k is quite irregular and oscillating. Some simple properties of P_k are shown. Moreover, upper and lower bounds for P_k, as well as for $\sum_{m=0}^{n} P_k(m)$, are determined. In this way one obtains also bounds for g_k and λ_k. However, these bounds are far to be tight.

In conclusion, we formulate some conjectures concerning the map P_k based on the numerical values obtained by implementing the algorithm described in Section 8.3.

2 Notations and preliminaries

In the following A_k, or simply A, will denote a finite alphabet of cardinality $k > 0$ and A^* the *free monoid* generated by A. The elements $a_1, a_2, \ldots, a_k$ of A are usually called *letters* and those of A^* *words*. We suppose that A is totally ordered by setting $a_1 < a_2 < \cdots < a_k$ and we shall denote often the letters simply by digits $1, 2, \ldots, k$, when there is no ambiguity. The identity element of A^*, called the *empty word*, is denoted by ε. We shall set $A^+ = A^* \setminus \{\varepsilon\}$. For any word $w \in A^*$ the *length* of w is denoted by $|w|$. The length of ε is conventionally 0.

For any letter $x \in A$, $|w|_x$ denotes the number of occurrences of x in w. The *Parikh vector* of w is a vector whose components are $|w|_x$ with $x \in A$. For any $w \in A^*$, $\operatorname{alph} w$ denotes the set of distinct letters of A occurring in w.

A word u is a *factor* of $w \in A^*$ if $w = rus$ for some words r and s. In the special case $r = \varepsilon$ (resp., $s = \varepsilon$), u is called a *prefix* (resp., *suffix*) of w. A factor u of w is *proper* if $u \neq w$. A factor u of w is *median* if $w = rus$ with $|r| = |s|$. If u is a factor of w, w is also called an *extension* of u.

A right infinite word, or simply *infinite word*, x is just an infinite sequence of letters:

$$x = x_1 x_2 \cdots x_n \cdots \text{ where } x_i \in A, \text{ for all } i \geq 1 \ .$$

For any integer $n \geq 0$, $x_{[n]}$ will denote the prefix $x_1 x_2 \cdots x_n$ of x of length

n. A factor of x is either the empty word or any sequence $x_i \cdots x_j$ with $i \leq j$. The set of all infinite words over A is denoted by A^ω. We also set $A^\infty = A^* \cup A^\omega$. For any $w \in A^\infty$ we denote respectively by $\operatorname{Fact} w$ and $\operatorname{Pref} w$ the sets of all factors and prefixes of the word w. For $Y \subseteq A^\infty$, $\operatorname{Fact} Y$ and $\operatorname{Pref} Y$ will denote respectively the sets of the factors and of the prefixes of all the words of Y.

A factor u of a word $w \in A^\infty$ is called *right special* if there exist $a, b \in A$, $a \neq b$, such that ua and ub are both factors of w. Symmetrically, u is said *left special* if $au, bu \in \operatorname{Fact} w$ for some distinct letters a and b. A factor u of w is called *bispecial* if it is both a right and a left special factor of w. A word u is called a right (resp., left) special factor of *a set* $Y \subseteq A^\infty$ if there exists letters $a, b \in A$ such that $a \neq b$ and $ua, ub \in \operatorname{Fact} Y$ (resp., $au, bu \in \operatorname{Fact} Y$). If u is both a right and left special factor of Y, then it is called a *bispecial factor* of Y. A word u is called a *strictly bispecial factor* of Y if

$$xuy \in \operatorname{Fact} Y, \quad \text{for all} \quad x, y \in A.$$

Note that if $\operatorname{card}(A) = k > 1$, then a strictly bispecial factor of Y is certainly bispecial, whereas the converse is not in general true.

Let $w = w_1 w_2 \cdots w_n$ with all $w_i \in A$, be a word. The *reversal* $w^\sim$ is the word $w_n \cdots w_1$. If $w = \varepsilon$, one sets $\varepsilon^\sim = \varepsilon$. A word $w \in A^*$ is a *palindrome* if $w = w^\sim$. The set of all palindromes of A^* is denoted by PAL_A, or simply PAL.

One can introduce in A^* the map ${}^{(+)} : A^* \to PAL$ which associates to any word $w \in A^*$ the palindrome $w^{(+)}$ defined as the shortest palindrome having the prefix w (cf. [4]). We call $w^{(+)}$ the *right palindrome closure of* w. If Q is the longest palindromic suffix of $w = uQ$, then one has

$$w^{(+)} = uQu^\sim .$$

For instance, if $w = abacbca$, then $w^{(+)} = abacbcaba$.

Let us define the map

$$\psi : A^* \to PAL,$$

as follows: $\psi(\varepsilon) = \varepsilon$ and for all $u \in A^*$, $x \in A$,

$$\psi(ux) = (\psi(u)x)^{(+)} .$$

For instance, if $v = abc$, one has $\psi(abac) = abaabacabaaba$. It is readily verified that for all words $v \in A^+$ if u is a prefix of v, then $\psi(u)$ is a palindromic prefix (and suffix) of $\psi(v)$ and, conversely, every palindromic prefix of $\psi(v)$ is of the form $\psi(u)$ for some prefix u of v.

The map ψ, called *(right) iterated palindrome closure* or simply *palindromization map*, is injective. For any $w \in \psi(A^*)$ the unique word u such that $\psi(u) = w$ is called the *directive word* of w. The directive word u of

w can be read from w just by taking the subsequence of w formed by all letters immediately following all proper palindromic prefixes of w.

For any $x \in A$, μ_x denotes the injective endomorphism of A^*

$$\mu_x : A^* \to A^*$$

defined as:

$$\mu_x(x) = x, \quad \mu_x(y) = xy, \text{ for } y \in A \setminus \{x\}.$$

If $v = x_1 x_2 \cdots x_n$, with $x_i \in A$, $i = 1, \ldots, n$, then we set:

$$\mu_v = \mu_{x_1} \circ \cdots \circ \mu_{x_n};$$

moreover, if $v = \varepsilon$, $\mu_\varepsilon = \text{id}$, where id is the identity map on A^*.

For any fixed letter $a \in A$, we set $Z = \mu_a(A)$ and denote by Z^* (resp., Z^ω) the set of all finite (resp., infinite) words generated by Z. Moreover, we set $Z^\infty = Z^* \cup Z^\omega$. The set Z is a code (cf. [1]) with a finite deciphering delay equal to 1. A consequence of this property is that any element of Z^∞ can be uniquely factorized by the elements of Z. Moreover, any pair (z, a) with $z \in A$ is synchronizing for Z, i.e., for all $\lambda \in A^*$ and $\zeta \in A^\infty$

$$\lambda z a \zeta \in Z^\infty \Longrightarrow \lambda z, \ a\zeta \in Z^\infty.$$

A letter $a \in A$ is said to be *separating* for $w \in A^\infty$ if it occurs in each factor of w of length 2. It is readily verified the following:

Proposition 2.1. *Let a be the first letter of $w \in A^\infty$. Then a is separating for w if and only if there exists, and is unique, a word $u \in A^\infty$ such that $w = \mu_a(u)$.*

The following formula, due to Justin [11] relates the palindromization map with the endomorphisms μ_v, $v \in A^*$:

$$\psi(vu) = \mu_v(\psi(u))\psi(v), \ v, u \in A^*.$$

In the particular case $v = x \in A$, one has:

$$\psi(xu) = \mu_x(\psi(u))x, \ x \in A, \ u \in A^*. \tag{1}$$

For definitions and notations on words not explicitly given in the text, the reader is referred to the books of Lothaire [12, 13].

3 Episturmian words

An infinite word $t \in A^\omega$ is *standard episturmian* if it is *closed under reversal* (that is, if $w \in \text{Fact}\, t$, then $\tilde{w} \in \text{Fact}\, t$) and each of its left special factors is a prefix of t. We denote by $SEP(A)$, or by SEP_k, the set of all standard episturmian words over the alphabet A with $k = \mathrm{card}(A)$.

An infinite word $s \in A^\omega$ is called *episturmian* if there exists a standard episturmian word $t \in A^\omega$ such that $\mathrm{Fact}\, s = \mathrm{Fact}\, t$. We shall denote by EP_k the class of all episturmian words over A.

The map ψ can then be naturally extended to A^ω by setting, for any infinite word x,

$$\psi(x) = \lim_{n\to\infty} \psi(x_{[n]}) \ .$$

The following important result was proved in [9]:

Theorem 3.1. *An infinite word t is standard episturmian over A if and only if there exists $\Delta \in A^\omega$ such that $t = \psi(\Delta)$.*

For any $t \in SEP_k$, there exists a *unique* Δ such that $t = \psi(\Delta)$. This Δ is called the *directive word* of t. If every letter of A occurs infinitely often in Δ, the word t is called a standard *Arnoux-Rauzy word*. In the case of a binary alphabet, a standard Arnoux-Rauzy word is usually called a *standard Sturmian word*. An infinite word s is called Arnoux-Rauzy if there exists a standard Arnoux-Rauzy word t such that $\mathrm{Fact}\, s = \mathrm{Fact}\, t$. We shall denote by SAR_k (resp., AR_k) the set of standard Arnoux-Rauzy (resp., Arnoux-Rauzy) words over A_k.

Example 3.2. *Let $A = \{a, b\}$ and $\Delta = (ab)^\omega$. The word $\psi(\Delta)$ is the famous* Fibonacci word

$$f = abaababaabaababaababa \cdots \ .$$

If $A = \{a, b, c\}$ and $\Delta = (abc)^\omega$, then $\psi(\Delta)$ is the so-called Tribonacci word

$$\tau = abacabaabacababacabaabacabaca \cdots \ .$$

The following proposition proved in [9] is a further characterization of standard episturmian words:

Proposition 3.3. *An infinite word t is standard episturmian if and only if any leftmost occurrence of a palindrome in t is a median factor of a palindromic prefix of t.*

The following lemmas are well known. The proof of the first is an immediate consequence of the definition. We report a proof of the second for the sake of completeness (see, for instance, [3]).

Lemma 3.4. *If $s \in AR_k$ then for each $n \geq 0$ there exists a unique right special factor u of length n such that $ux \in \mathrm{Fact}\, s$ for all $x \in A_k$.*

Lemma 3.5. *For all $k > 0$, $\mathrm{Pref}\, SEP_k = \mathrm{Pref}\, SAR_k$ and $\mathrm{Fact}\, EP_k = \mathrm{Fact}\, AR_k$.*

Proof. Let u be a prefix of a standard episturmian word $s \in SEP_k$. One has $s = \psi(\Delta)$ where $\Delta = t_1 t_2 \cdots t_n \cdots$ with $t_i \in A$, $i > 0$. Therefore there exists a palindromic prefix p of s such that $u \in \operatorname{Pref} p$. Now $p = \psi(t_1 \cdots t_i)$ for some i. We can consider $\Delta' = t_1 \cdots t_i t$ with $t \in A^\omega$ such that any letter of A occurs infinitely many times in t. Hence, $s' = \psi(\Delta') \in SAR_k$ and has the prefix p, so that $u \in \operatorname{Pref} s'$. Hence, $\operatorname{Pref} SEP_k \subseteq \operatorname{Pref} SAR_k$. Since the inverse inclusion is trivial, one obtains $\operatorname{Pref} SEP_k = \operatorname{Pref} SAR_k$.

Let $v \in \operatorname{Fact} EP_k = \operatorname{Fact} SEP_k$. Therefore, there exists a standard episturmian word s such that $v \in \operatorname{Fact} s$. There exists a prefix u of s such that $v \in \operatorname{Fact} u$. Since, as we have proved above, $u \in \operatorname{Pref} s'$ with $s' \in SAR_k$, it follows that $\operatorname{Fact} EP_k \subseteq \operatorname{Fact} AR_k$. Since the inverse inclusion is trivial the result follows. □

The proofs of the two following lemmas are in [9].

Lemma 3.6. *Let t be a standard episturmian word and a be its first letter. Then a is separating for t.*

Lemma 3.7. *Let $t \in A^\omega$ and $a \in A$. Then $\mu_a(t)$ is a standard episturmian word if and only if so is t.*

In the next sections we shall use the following interesting proposition ([14], Lemma 1) on Arnoux-Rauzy words:

Proposition 3.8. *Let u be a right special factor of* $\operatorname{Fact} AR_k$. *Then there exists $s \in AR_k$ such that u is a right special factor of s.*

4 Epicentral words

In the following, PER_k will denote the set $\psi(A_k^*)$ of the palindromic prefixes of all standard episturmian words over the k-letter alphabet A_k. Since by Lemma 3.5, $\operatorname{Pref} SEP_k = \operatorname{Pref} SAR_k$, one has that PER_k equals the set of palindromic prefixes of all standard Arnoux-Rauzy words over A_k. One trivially has that $\operatorname{Fact} SEP_k = \operatorname{Fact} PER_k$.

The words of PER_k are also called *k-central*, or without making reference to k, simply *epicentral words*. The class PER_2 is usually denoted by PER and the 2-central words are called *central words* [2, 13].

We shall set for any $n \geq 0$,

$$P_k(n) = \operatorname{card}(PER_k \cap A_k^n),$$

i.e., $P_k(n)$ counts the number of epicentral words over A_k.

If $k = 1$, then $P_1(n) = 1$ for all $n \geq 0$. In the case $k = 2$, it was proved [6] that $P_2(n) = \varphi(n+2)$, where φ is the Euler function.

As we shall see (cf. Theorem 4.3) the set of epicentral words PER_k is equal to the set SBS_k of strictly bispecial factors of $\operatorname{Fact} SEP_k$. This was proved in the case $k = 2$ in [6]. However, the proof in [6] cannot be extended to the case $k > 2$ since it uses the *balance property* of Sturmian words which, in general, does not hold for episturmian words.

Let us remark that a bispecial factor of $\mathrm{Fact}\, SEP_k$ is not in general strictly bispecial. For instance, it is readily verified that ab is a bispecial factor of $\mathrm{Fact}\, SEP_k$ but it is not strictly bispecial since by Lemma 3.6, $aabb$ is not a factor of any episturmian word.

The following lemma, needed to prove Theorem 4.3, is an immediate consequence of the local balance property introduced by Richomme in [15] as a characterization of recurrent words which are episturmian. We report a proof of the lemma for the sake of completeness and remark that the two conditions are not equivalent[1].

Lemma 4.1. *Let s be any episturmian word. For all $w \in A^*$ and $a, b \in A$*

$$awa \in \mathrm{Fact}\, s \text{ and } b \neq a \Longrightarrow bwb \notin \mathrm{Fact}\, s. \qquad (2)$$

Proof. We can assume without loss of generality that s is a standard episturmian word. The proof is by induction on the length of the word w. The result is trivially true when $|w| = 0$. Indeed, from the uniqueness of the separating letter of s one cannot have $aa \in \mathrm{Fact}\, s$ and also $bb \in \mathrm{Fact}\, s$ with $b \neq a$. Let us then assume the statement true for any $w' \in A^*$ with $|w'| < |w|$ and $s' \in SEP_k$ and suppose by contradiction that $awa, bwb \in \mathrm{Fact}\, s$ for $s \in SEP_k$. By Lemma 3.6 the first letter x of s is separating for s. By Proposition 2.1, $s = \mu_x(t)$ for a unique word $t \in A^\omega$. By Lemma 3.7, $t \in SEP_k$.

Since $a \neq b$, one has $x \neq a$ or $x \neq b$; we can assume without loss of generality that $x \neq a$. Since $awa \in \mathrm{Fact}\, s$ and the separating letter of s is x, one derives that the first and the last letter of w have to be x. Hence, we can write by Proposition 2.1

$$w = \mu_x(w')x$$

for a suitable word $w' \in A^*$. Moreover, as $xawax \in \mathrm{Fact}\, s$, one has also that $\mu_x(aw'a) \in \mathrm{Fact}\, s$. From this and the synchronization property of the code $\mu_x(A)$, it follows that $aw'a \in \mathrm{Fact}\, t$. If $b \neq x$ by the same argument one derives $bw'b \in \mathrm{Fact}\, t$; if $b = x$, then one has $bwb = xwx = x\mu_x(w')xx = \mu_x(xw'x)x \in \mathrm{Fact}\, s$, so that $xw'x = bw'b \in \mathrm{Fact}\, t$. In all cases one obtains $aw'a, bw'b \in \mathrm{Fact}\, t$. Since $|w'| < |w|$ we reach a contradiction which concludes our proof. □

Lemma 4.2. *If w is a palindromic right special factor of* $\mathrm{Fact}\, AR_k$*, then* $w \in PER_k$.

[1] In fact, the word $w = (acabcd)^\omega$ is a recurrent word which satisfies condition (2) but does not satisfy the local balance property.

Proof. Let w be a right special factor of $\operatorname{Fact} AR_k$. By Proposition 3.8 there exists an Arnoux-Rauzy word s, that we can always take standard, such that w is a right special factor of s. Since w is a palindrome, it follows that w is also a left special factor of s and then a palindromic prefix of s. Thus $w \in PER_k$. □

Theorem 4.3. *For all $k \geq 1$, $PER_k = SBS_k$.*

Proof. If $k = 1$, i.e., $A = \{a\}$ the result is trivial since $PER_1 = \psi(a^*) = a^*$. Moreover, $SEP_1 = a^\omega$, so that $\operatorname{Fact} SEP_1 = a^*$ and $SBS_1 = a^*$. Let us then suppose $k > 1$. We first prove that $PER_k \subseteq SBS_k$. Let $w \in PER_k$. Then there exists a directive word $v \in A_k^*$ such that $w = \psi(v)$. For all $x, y \in A_k$, one has from Justin's formula $\psi(vx) = \mu_v(x)\psi(v)$ and

$$\psi(vxy) = (\psi(vx)y)^{(+)} = \mu_v(x)\psi(v)y\zeta$$

for a suitable $\zeta \in A_k^*$. Since $\mu_v(x)$ terminates with the letter x one obtains that $x\psi(v)y = xwy \in \operatorname{Fact} AR_k$, i.e., $w \in SBS_k$.

Let us now prove that $SBS_k \subseteq PER_k$. Let $w \in SBS_k$. Since w is right special in $\operatorname{Fact} AR_k$, in view of Lemma 4.2 it is sufficient to prove that $w \in PAL$. Indeed, suppose by contradiction that w is not a palindrome. We can write $w = a_1 \cdots a_n$ with $a_i \in A$, $1 \leq i \leq n$ and set $m = \min\{h \mid 1 \leq h \leq n \text{ and } a_h \neq a_{n-h+1}\}$. Let $a_m = x \in A$ and set $a_{n-m+1} = y \neq x$. We can then write:

$$w = uxtyu^{\sim},$$

with $t \in A^*$ and $u = a_1 \cdots a_{m-1}$. By the hypothesis $xwy \in \operatorname{Fact} AR_k$, so that there exists a standard episturmian word s such that $xwy \in \operatorname{Fact} s$. This implies that $xux, yu^{\sim}y \in \operatorname{Fact} s$. Since s is closed under reversal, we have also $yuy \in \operatorname{Fact} s$, a contradiction in view of Lemma 4.1. □

5 Counting palindromes in episturmian words

In this section we are interested in counting for each length n the palindromic factors of length n in the set of all episturmian words over A_k. For any fixed $k > 0$ we set:

$$G_k = PAL \cap \operatorname{Fact} EP_k = PAL \cap \operatorname{Fact} AR_k,$$

where the last equality is a consequence of Lemma 3.5. We introduce the map $g_k : \mathbb{N} \to \mathbb{N}$ defined, for all $n \geq 0$, as:

$$g_k(n) = \operatorname{card}(G_k \cap A^n).$$

As we shall see (cf. Theorem 5.2) the computation of g_k will be reduced to the computation of P_k.

Proposition 5.1. *If $v \in PAL \cap \mathrm{Fact}\, AR_k$ but v is not k-central, then there exists and is unique a letter $x \in A_k$ such that*

$$xvx \in \mathrm{Fact}\, AR_k.$$

Proof. If $v \in PAL \cap \mathrm{Fact}\, AR_k$, then there exists a standard Arnoux-Rauzy word s such that $v \in \mathrm{Fact}\, s$. Since $v \notin PER_k$, by Proposition 3.3 the leftmost occurrence of v is s is a proper median factor of a palindromic prefix of s. Therefore, certainly there exists a letter $x \in A_k$ such that $xvx \in \mathrm{Fact}\, s \subseteq \mathrm{Fact}\, AR_k$.

Now let us suppose that there exists a letter $y \neq x$ such that $yvy \in \mathrm{Fact}\, AR_k$. Since v is a right special factor of $\mathrm{Fact}\, AR_k$ and a palindrome, from Lemma 4.2 we obtain $v \in PER_k$, which is a contradiction. □

Theorem 5.2. *Let $k > 0$. For all $n \geq 0$*

$$g_k(n+2) = g_k(n) + (k-1)P_k(n). \tag{3}$$

Proof. Let v be a palindrome of length n in $\mathrm{Fact}\, AR_k$. If v is not epicentral, then by Proposition 5.1 there exists a unique letter x such that $xvx \in PAL \cap \mathrm{Fact}\, AR_k$.

If, on the contrary, v is epicentral, then by Theorem 4.3 for all $x \in A_k$ one has:

$$xvx \in PAL \cap \mathrm{Fact}\, AR_k.$$

Hence, all $g_k(n)$ palindromes of length n in $\mathrm{Fact}\, AR_k$ can be extended to palindromes of $\mathrm{Fact}\, AR_k$ of length $n+2$ in number of

$$kP_k(n) + g_k(n) - P_k(n) = g_k(n) + (k-1)P_k(n).$$

Since any palindrome in $\mathrm{Fact}\, AR_k$ of length $n+2$ has a median factor of length n in $\mathrm{Fact}\, AR_k$, our assertion follows. □

From the preceding theorem by iteration of (3) one immediately derives:

Theorem 5.3. *Let $k > 0$. One has $g_k(0) = 1$, $g_k(1) = k$ and for any $n > 0$:*

$$g_k(n) = k + (k-1) \sum_{i=1}^{\lceil n/2 \rceil - 1} P_k(n-2i). \tag{4}$$

We can rewrite equation (4), separating the cases of even and odd integers, as:

$$g_k(2n) = 1 + (k-1) \sum_{m=0}^{n-1} P_k(2m), \;\; n \geq 0, \tag{5}$$

$$g_k(2n+1) = k + (k-1) \sum_{m=0}^{n-1} P_k(2m+1), \;\; n \geq 0. \tag{6}$$

In the case $k = 2$ one has that $P_2(n) = \varphi(n+2)$ and, as it is readily verified, equation (4) becomes (see [5]) :

$$g_2(n) = 1 + \sum_{i=0}^{\lceil n/2 \rceil - 1} \varphi(n - 2i), \text{ for all } n \geq 0.$$

6 Counting finite episturmian words

Let us introduce, for any $k > 0$, the map $\lambda_k : \mathbb{N} \to \mathbb{N}$ defined for any $n \geq 0$ as:

$$\lambda_k(n) = \operatorname{card}(A_k^n \cap \operatorname{Fact} EP_k),$$

i.e., $\lambda_k(n)$ counts the number of finite episturmian words of length n over the alphabet A_k. The following holds:

Theorem 6.1. *Let $k > 0$. For any $n > 0$,*

$$\lambda_k(n) = k + k(k-1)(n-1) + (k-1)^2 \sum_{i=1}^{n-2} (n-i-1) P_k(i). \tag{7}$$

Let us observe that in [14] a similar formula was proved with P_k replaced by the map b_k counting for each n the bispecial factors of $\operatorname{Fact} AR_k$ of length n. Actually, this latter formula, as well as all further results in [14], holds [2] if by b_k one means the map counting the bispecial factors in all Arnoux-Rauzy words (i.e., P_k) or, equivalently in view of Theorem 4.3, if one replaces the map b_k with the map sb_k counting for each length n the strictly bispecial factors of $\operatorname{Fact} AR_k$ of length n.

We report here a proof of Theorem 6.1 for the sake of clarity.

Proof of Theorem 6.1. Let $k > 0$. We denote by r_k the map counting for each n the right special factors of $\operatorname{Fact} AR_k$ of length n. The following formula, proved in [6] in the case $k = 2$, holds:

$$r_k(n+1) = r_k(n) + (k-1) sb_k(n). \tag{8}$$

Indeed, we first observe that by Proposition 3.8 one easily derives that a left special factor of $\operatorname{Fact} AR_k$ can be extended on the left to a word of $\operatorname{Fact} AR_k$ by the k letters of A_k. Moreover, any right special factor u of $\operatorname{Fact} AR_k$ of length n is the suffix of length n of at least one right special factor of $\operatorname{Fact} AR_k$ of length $n+1$. Indeed, by Proposition 3.8 there exists $s \in AR_k$ such that u is a right special factor of s. The Arnoux-Rauzy word s has a unique right special factor of length n, namely u which is suffix of the unique right special factor of s of length $n+1$.

[2] Lemma 2 and Corollary 2 of [14] are not correct. Indeed, for instance, ab is trivially a bispecial factor of $\operatorname{Fact} AR_k$ but there does not exist a standard Arnoux-Rauzy word ω such that ab is a bispecial factor of ω

If u is the suffix of more that one right special factor of $\mathrm{Fact}\, AR_k$ of length $n+1$, then it is also left special, so that it is the suffix of k right special factors of $\mathrm{Fact}\, AR_k$ of length $n+1$, namely xu for any $x \in A_k$. Since each xu is right special it follows that for all $x, y \in A_k$, one has that $xuy \in \mathrm{Fact}\, AR_k$, so that u is a strictly bispecial factor of $\mathrm{Fact}\, AR_k$ of length n.

Therefore, the number $r_k(n+1)$ of right special factors of $\mathrm{Fact}\, AR_k$ of length $n+1$ is obtained by adding to the number $r_k(n)$ of right special factors of $\mathrm{Fact}\, AR_k$ of length n the further number $(k-1)sb_k(n)$, where $sb_k(n)$ is the number of strictly bispecial factors of length n.

The remaining part of the proof can be easily derived from equation (8) by using the relation:

$$\lambda_k(n) = \lambda_k(n-1) + (k-1)r_k(n-1),$$

and Theorem 4.3. $\square$

In the case $k = 2$, taking into account that $P_2(n) = \varphi(n+2)$, equation (7) simply becomes

$$\lambda_2(n) = 1 + \sum_{i=1}^{n} (n-i+1)\varphi(i).$$

This latter formula was proved by several authors using different techniques (see, for instance, [2] and references therein).

7 Canonical epicentral words

Let $v = x_1 \cdots x_n$, $x_i \in A$, $i = 1, \ldots, n$, be a word over A. For any $x \in \mathrm{alph}\, v$ we denote by j_x the first occurrence of x in v, i.e., the least integer i such that $x = x_i$. We order $\mathrm{alph}\, v$ by setting for $x, y \in \mathrm{alph}\, v$:

$$x \prec y \;\text{ if }\; j_x < j_y.$$

We define *word order* of v, the sequence $\mathrm{ord}\, v = a_1 \cdots a_r$, where $r = \mathrm{card}(\mathrm{alph}\, v)$, $a_i \in \mathrm{alph}\, v$, $i = 1, \ldots, r$, and $a_i \prec a_j$ for $i < j$. If $v = \varepsilon$, we set $\mathrm{ord}\, v = \varepsilon$. For instance, if $v = baacba$, then $\mathrm{ord}\, v = bac$. Note that $|\mathrm{ord}\, v| = \mathrm{card}(\mathrm{alph}\, v)$. If $w = \psi(v)$ is an epicentral word, then it is readily verified, by the definition of $\psi(v)$, that $\mathrm{ord}\, w = \mathrm{ord}\, v$.

An epicentral word $w = \psi(v)$ over A_k will be called *canonical* if $\mathrm{ord}\, v = \mathrm{ord}\, w \in \mathrm{Pref}\, 12 \cdots k$. In order to explain the use of this term we give the following definition: we say that two words $w, w' \in A_k^*$ are *similar* if each can be obtained from the other by a word isomorphism (i.e., a renaming of the letters). This is trivially an equivalence relation. A canonical epicentral word is a suitable choice of a representative in a similarity class of an epicentral word, i.e., it is minimal with respect to the lexicographic order. For instance, on a three-letter alphabet $A = \{a, b, c\}$ with $a < b < c$, the epicentral word $w = bcbabcb$ is not canonical since $\operatorname{ord} w = bca \notin \operatorname{Pref} abc$. The canonical epicentral in the similarity class of w is the word $w' = abacaba$.

In general, if $w = \psi(v)$ and $\operatorname{ord} v = x_1 \cdots x_i$, then the canonical epicentral word similar to w is $w' = \psi(v')$ with $v' = \sigma(v)$ where σ is the isomorphism of A_k^* defined by $\sigma(x_j) = j$ for $1 \leq j \leq i$. It is readily verified that $|w'| = |w|$.

Let $w = \psi(v)$ be an epicentral word over A_k. We shall denote by ω the map $\omega : A_k^* \to \mathbb{N}^k$ where for any $v \in A_k^*$, $\omega(v)$ is Parikh vector of $w = \psi(v)$. For $i = 1, \ldots, k$ one has

$$\omega_i(v) = |\psi(v)|_i \quad \text{and} \quad |\omega(v)| = |\psi(v)|,$$

having set:

$$|\omega(v)| = \sum_{i=1}^{k} \omega_i(v).$$

The importance of the Parikh vector of an epicentral word is due to the following (cf. [7, 8])

Proposition 7.1. *The map ω is injective.*

We remark that in fact there exists an effective procedure which allows to construct the epicentral word $\psi(v)$ from the knowledge of $\omega(v)$ (cf. [7, 8]).

An interesting characterization of epicentral words which are canonical in terms of their Parikh vectors, is given by Proposition 7.4. We need two lemmas:

Lemma 7.2. *Let $w = \psi(v)$ be an epicentral word over A_k having Parikh vector $\omega(v)$. Then for any letter $j \in A_k$ one has:*

$$\omega_j(jv) = |\omega(v)| + 1 \quad \textit{and} \quad \omega_i(jv) = \omega_i(v) \quad \textit{for} \quad i \neq j.$$

Proof. By equation (1), for $i \neq j$ one has:

$$\omega_i(jv) = |\psi(jv)|_i = |\mu_j(\psi(v))j|_i = |\mu_j(\psi(v))|_i = |\psi(v)|_i = \omega_i(v);$$

indeed, since $j \neq i$, the application of μ_j to $\psi(v)$ does not change the number of occurrences of the letter i in $\psi(jv)$. For $i = j$ one has:

$$\omega_j(jv) = |\mu_j(\psi(v))j|_j = 1 + |\psi(v)| = 1 + |\omega(v)|;$$

indeed, if one applies μ_j to $\psi(v)$ any occurrence of the letter j in $\psi(v)$ remain unchanged, whereas any occurrence of a letter $\neq j$ produces a new occurrence of j in $\mu_j(\psi(v))$. Hence, $|\mu_j(\psi(v))|_j = |\psi(v)|$. □

Lemma 7.3. *Let* $w = \psi(v)$ *and let* $\operatorname{ord} w = \operatorname{ord} v = x_1 x_2 \cdots x_m$ *with* $m > 1$ *and* $x_h \in A_k$*,* $1 \leq h \leq m \leq k$*. Then, for all* $1 \leq i < j \leq m$*,*

$$\omega_{x_i}(v) > \omega_{x_j}(v). \tag{9}$$

Proof. We proceed by induction on $|v|$. The statement is trivially true for $|v| = 2$, so let us assume that $|v| > 2$ and that the statement holds for all directive words whose length is ≥ 2 and smaller than $|v|$. By Lemma 7.2 the first letter of the directive word of an epicentral word is the most frequent letter; thus equation (9) holds for $i = 1$ and the statement of the lemma is proved for $m = 2$. Let us then assume $i \geq 2$ and $m > 2$. Let $v = x_1\bar{v}$. For all $2 \leq i < j \leq m$, the letter x_i occurs in $\operatorname{ord}\bar{v}$ before x_j. Thus, since $|\bar{v}| < |v|$, by induction hypothesis we have

$$\omega_{x_i}(\bar{v}) > \omega_{x_j}(\bar{v})$$

for all $2 \leq i < j \leq m$. For all $i \geq 2$, $x_i \neq x_1$, so that by Lemma 7.2

$$\omega_{x_i}(v) = \omega_{x_i}(x_1\bar{v}) = \omega_{x_i}(\bar{v}),$$

from which the result follows. □

From the preceding lemma, the following proposition easily follows:

Proposition 7.4. *Let* $w = \psi(v)$ *be an epicentral word over the alphabet* A_k*. Then* w *is canonical if and only if*

$$\omega_1(v) \geq \omega_2(v) \geq \cdots \geq \omega_k(v). \tag{10}$$

Proof. If w is canonical, then $\operatorname{ord} v = 12 \cdots m$ with $0 \leq m \leq k$. If $m = 0$ then $v = \varepsilon$ so that $\omega_i(\varepsilon) = 0$ for $1 \leq i \leq k$ and the result follows. If $m = 1$, then $v = 1$ and $\omega_1(1) = 1$ and $\omega_i(1) = 0$ for $1 < i \leq k$. Let us suppose $m > 1$. By Lemma 7.3 and the fact that $\omega_j(v) = 0$ for $j > m$, equation (10) follows. Conversely, suppose that

$$\omega_1(v) \geq \omega_2(v) \geq \cdots \geq \omega_k(v);$$

we want to show that $w = \psi(v)$ is canonical. Indeed, if there exist $1 \leq i < j \leq k$ such that the letter j occurs before the letter i in $\operatorname{ord} v$, then by the preceding lemma, one would derive that $\omega_j(v) > \omega_i(v)$, which contradicts our hypothesis. □

By the previous proposition if $w = \psi(v)$ is an epicentral word over A_k having the Parikh vector $\omega(v)$, then the Parikh vector of the canonical epicentral word similar to w is obtained from $\omega(v)$ by arranging its components in a decreasing order.

Let $w = \psi(v)$ be a canonical epicentral word over A_k. The following proposition allows to construct for any letter $j \in A_k$ the Parikh vector of the canonical epicentral word similar to $\psi(jv)$.

Proposition 7.5. *Let $w = \psi(v)$ be a canonical epicentral word over A_k having Parikh vector $\omega(v) = (\omega_1(v), \ldots, \omega_k(v))$. Then for any $j \in A_k$ the canonical epicentral word in the similarity class of $\psi(jv)$ has Parikh vector*

$$(|\omega(v)| + 1, \omega_1(v), \ldots, \omega_{j-1}(v), \omega_{j+1}(v), \ldots, \omega_k(v)).$$

Proof. By Proposition 7.4 one has $\omega_i(v) \geq \omega_{i+1}$ for all $i = 1, \ldots, k-1$. From Lemma 7.2 the Parikh vector of $\psi(jv)$ is

$$\omega(jv) = (\omega_1(v), \ldots, \omega_{j-1}(v), |\omega(v)| + 1, \omega_{j+1}(v), \ldots, \omega_k(v)).$$

If we arrange the components of $\omega(jv)$ in a decreasing order we obtain the result in view of Proposition 7.4. □

Example 7.6. Let $w = \psi(abc) = abacaba$ be the canonical epicentral word over $A_3 = \{a, b, c\}$ with $a < b < c$ and $\omega(abc) = (4, 2, 1)$. The word $w_1 = \psi(aabc) = aabaacaabaa$, $w_2 = \psi(babc) = babbabcbabbab$, and $w_3 = \psi(cabc) = cacbcaccacbcac$ having respectively the Parikh vectors $\omega(aabc) = (8, 2, 1)$, $\omega(babc) = (4, 8, 1)$, and $\omega(cabc) = (4, 2, 8)$. One has that w_1 is canonical; the canonical epicentral words similar to w_2 and w_3 are respectively $w_2' = \psi(abac) = abaabacabaaba$ having Parikh vector $(8, 4, 1)$ and $w_3' = \psi(abca) = abacabaabacaba$ whose Parikh vector is $(8, 4, 2)$.

Proposition 7.7. *For any $1 \leq i \leq k$, a canonical epicentral word $w = \psi(v)$ over A_k such that* $\operatorname{ord} v = 1 \cdots i$ *has a minimal length equal to* $2^i - 1$. *This minimal value is reached when $v = 1 \cdots i$.*

Proof. Let us first prove by induction on the integer i that $|\psi(1 \cdots i)| = 2^i - 1$. This is trivial for $i = 1$, so that assume that the statement is true for all positive integers less than or equal to $i < k$ and prove it for $i + 1$. Indeed, $\psi(1 \cdots (i+1)) = \psi(1 \cdots i)(i+1)\psi(1 \cdots i)$ so that $|\psi(1 \cdots (i+1))| = 2|\psi(1 \cdots i)| + 1 = 2^{i+1} - 1$.

Let us now prove that any canonical epicentral word $w = \psi(v)$ with $\operatorname{ord} v = 1 \cdots i$ is such that $|\psi(v)| \geq 2^i - 1$. The proof is obtained by induction on the integer i. The statement is trivially true in the case $i = 1$. Indeed, $v = 1^{|v|}$ and $\psi(v) = |v| \geq 1$. Let us then assume the statement true up to $i < k$ and prove it for $i + 1$. We can write v as $v = u(i+1)\zeta$ where $\zeta \in A_k^*$ and the letter $(i+1)$ does not occur in u. Now the word $\psi(v)$ has the prefix $\psi(u(i+1))$ of length $2|\psi(u)| + 1$. Since $\operatorname{ord} u = 1 \cdots i$, by induction one has $|\psi(u)| \geq 2^i - 1$, so that $|\psi(v)| \geq |\psi(u(i+1))| \geq 2^{i+1} - 1$, which proves the assertion. □

8 Counting epicentral words

As we have seen in the previous sections, for any $k > 0$ the enumeration map g_k of the palindromes in Fact EP_k, as well as the map λ_k counting the finite episturmian words of any length, depends on the map P_k counting for each n the epicentral words of length n. Therefore, the map P_k plays a key role in both counting problems. Another important function is the map $\hat{P}_k$, which counts for any n the number of epicentral words over a k-letter alphabet of length *at most* n, i.e.,

$$\hat{P}_k(n) = \sum_{m=0}^{n} P_k(m).$$

Since for all $n \geq 0$, $P_2(n) = \varphi(n+2)$, one has that $\hat{P}_2(n) = \Phi(n+2) - 1$, where for all $n \geq 1$,

$$\Phi(n) = \sum_{m=1}^{n} \varphi(m).$$

The first values of $P_k(n)$, $g_k(n)$, and $\lambda_k(n)$ for $3 \leq k \leq 6$ are respectively reported in Tables 1, 2, and 4 of the Appendix.

In this section we shall analyze some combinatorial properties of the map P_k that will allow us to give some upper and lower bounds for P_k and $\hat{P}_k$; from this one can derive also bounds for maps g_k and λ_k. Finally, a quite efficient procedure for counting epicentral words is briefly described.

For each $0 \leq i \leq k$ and $n \geq 0$, we shall denote by $U_i(n)$ the number of canonical epicentral words u of length n such that $\operatorname{ord} u = 12 \cdots i$; in the case $i = 0$, $\operatorname{ord} u = \varepsilon$. The first values of $U_i(n)$ for $3 \leq i \leq 6$ are reported in Table 3 of the Appendix.

From the definition one has that $U_0(0) = 1$ and $U_i(0) = 0$ for $i > 0$. For all $n > 0$, one has $U_1(n) = 1$. Moreover, since for all $n \geq 0$, $P_2(n) = \varphi(n+2)$, one derives that for $n > 0$, $U_2(n) = (1/2)\varphi(n+2) - 1$. Tab. 3 shows that for any i several initial values of $U_i(n)$ are 0. Indeed, by Proposition 7.7 one has that $U_i(n) = 0$ for $n < 2^i - 1$.

Proposition 8.1. *For $n \geq 0$ the following holds*

$$P_k(n) = \sum_{i=0}^{k} \frac{k!}{(k-i)!} U_i(n). \tag{11}$$

Proof. We can write $P_k(n) = \sum_{i=0}^{k} P_k^{(i)}(n)$, where $P_k^{(i)}(n)$ denotes the number of epicentral words w over A_k of length n such that $\operatorname{card}(\operatorname{alph} w) = i$. These words can be generated from the canonical epicentral words by replacing the letters $1, 2, \ldots, i$ with any i distinct letters taken from A_k (and this can be done in $\binom{k}{i}$ ways) and then permuting the i letters in all possible ways (and this can be done in $i!$ ways), so that $P_k^{(i)}(n) = \frac{k!}{(k-i)!} U_i(n)$. □

A consequence of the preceding proposition is that for $n > 0$, $P_k(n)$ has the same parity as k.

Corollary 8.2. *For $n > 0$ one has:*

$$P_k(n) \equiv k \pmod 2.$$

Proof. The result is trivial for $k = 1$. Let us then suppose $k > 1$. Since for all $n > 0$, $U_1(n) = 1$, from Proposition 8.1, we can write:

$$P_k(n) = k + \sum_{i=2}^{k} k(k-1)\cdots(k-i+1)U_i(n).$$

For $2 \leq i \leq k$, the product $k(k-1)\cdots(k-i+1)$ is always an even integer, so that the result follows. □

Corollary 8.3. *For $n > 0$ one has:*

$$g_k(n) \equiv k \pmod 2$$

and

$$\lambda_k(n) \equiv k \pmod 2.$$

Proof. The proof is immediately derived from the preceding corollary and equations (5), (6), and (7). □

Proposition 8.4. *For all $k > 1$ and $n > 0$*

$$P_k(n) \geq \frac{k}{k-1} P_{k-1}(n). \tag{12}$$

Proof. From (11) one easily derives:

$$P_k(n) - P_{k-1}(n) = \sum_{i=1}^{k} i \; \frac{(k-1)!}{(k-i)!} U_i(n) \geq \sum_{i=1}^{k-1} \frac{i}{(k-i)} \frac{(k-1)!}{(k-1-i)!} U_i(n).$$

Since for i varying in the interval $[1, k-1]$, $i/(k-i) \geq 1/(k-1)$, it follows that the r.h.s. of the above formula is greater than or equal to $1/(k-1)P_{k-1}(n)$. Hence,

$$P_k(n) - P_{k-1}(n) \geq \frac{1}{k-1} P_{k-1}(n),$$

from which the assertion follows.

□

Corollary 8.5. *For all $k > 1$ and $n > 0$ the following holds:*

$$P_k(n) \geq \frac{k}{2} \; \varphi(n+2).$$

Proof. The proof is immediately obtained by iterating (12) and taking into account the fact that $P_2(n) = \varphi(n+2)$. □

8.1 Upper bounds

In this section we shall give some upper bounds for the functions P_k, $\hat{P}_k$, g_k, and λ_k.

Let $w = \psi(v)$ be an epicentral word of length n having the directive word v and Parikh vector $\omega(v)$. From Proposition 7.1 the map ω is injective so that for any $n \geq 0$

$$P_k(n) = \text{card}\{\omega(v) \mid v \in A_k^* \text{ and } \ |\omega(v)| = |\psi(v)| = n\}.$$

The r.h.s. of preceding equation is upper bounded by the number of all k-tuples $(i_1, \ldots, i_k)$ of integers such that $i_1 + \cdots + i_k = n$ and $i_h \geq 0$, $h = 1, \ldots, k$. This latter number is well known [16] and it is called the number of *compositions* of the integer n into k parts. It is given by the binomial coefficient:

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}.$$

Moreover, it is readily verified that $\binom{n+k-1}{n} \leq (n+1)^{k-1}$, so that for all $n \geq 0$

$$P_k(n) \leq \binom{n+k-1}{n} \leq (n+1)^{k-1}. \tag{13}$$

From this one also obtains:

$$\hat{P}_k(n) \leq \sum_{i=0}^{n} \binom{i+k-1}{k-1} = \binom{n+k}{n} \leq (n+1)^k. \tag{14}$$

Note that in the case $k = 2$ from (13) one obtains $P_2(n) \leq n+1$ that implies for $n > 1$, $\varphi(n) \leq n - 1$.

From equations (5) and (6) we can find easily an upper bound for the map g_k. Indeed, from (13) one has:

$$\sum_{m=0}^{n-1} P_k(2m) \leq \sum_{m=0}^{n-1} (2m+1)^{k-1} \leq \frac{1}{2k}(2n+1)^k,$$

$$\sum_{m=0}^{n-1} P_k(2m+1) \leq \sum_{m=0}^{n-1} (2m+2)^{k-1} \leq \frac{1}{2k}(2n+2)^k.$$

Hence,

$$g_k(2n) \leq 1 + \frac{k-1}{2k}(2n+1)^k, \quad g_k(2n+1) \leq k + \frac{k-1}{2k}(2n+2)^k.$$

From this one derives:

Proposition 8.6. *For any $k > 0$ one has*

$$g_k(n) = O(n^k).$$

Let us observe that the preceding proposition can also be immediately derived from equations (5) and (6) by using the obvious relations:

$$g_k(2n) \leq 1 + (k-1)\hat{P}_k(2n-2), \quad g_k(2n+1) \leq k + (k-1)\hat{P}_k(2n-1)$$

and (14).

Proposition 8.7. *For any $k > 0$ one has*

$$\lambda_k(n) = O(n^{k+1}).$$

Proof. The result easily follows from equations (13) and (7). □

8.2 Lower bounds

In this section we will examine some lower bounds for the functions P_k, $\hat{P}_k$, and g_k.

Lemma 8.8. *For any $v \in A^*$,*

$$|\psi(v)| \leq 2^{|v|} - 1.$$

Proof. Let, by induction on the length of v, the statement be true for all the lengths up to n (the base case for $|v| = 0$ is trivial). From the definition of ψ, it is clear that, for any $v \in A^*$ and $a \in A$,

$$|\psi(va)| = \left|(\psi(v)a)^{(+)}\right| \leq 2|\psi(v)| + 1.$$

If $|v| = n$, we have, by induction hypothesis, that

$$|\psi(v)| \leq 2^{|v|} - 1,$$

so

$$|\psi(va)| \leq 2(2^{|v|} - 1) + 1 = 2^{|v|+1} - 1 = 2^{|va|} - 1$$

which proves the thesis. □

We are now able to prove the following lower bound for $\hat{P}_k$:

Theorem 8.9. *For any $k > 1$ and $n \geq 0$ the following inequality holds:*

$$\hat{P}_k(n) \geq \frac{k^{\lfloor \log_2(n+1) \rfloor + 1} - 1}{k-1} + k(n - \lfloor \log_2(n+1) \rfloor).$$

Proof. From Lemma 8.8, we have that if $|v| \leq \log_2(n+1)$, then

$$|\psi(v)| \leq 2^{\log_2(n+1)} - 1 = n.$$

Thus

$$v \in \bigcup_{i=0}^{\lfloor \log_2(n+1) \rfloor} A_k^i \quad \Rightarrow \quad |\psi(v)| \leq n.$$

Moreover, for any $x \in A_k$ and $\lfloor \log_2(n+1) \rfloor + 1 \leq p \leq n$ the directive word $v = x^p$ generates the epicentral word x^p. Therefore, there exist at least $k(n - \lfloor \log_2(n+1) \rfloor)$ additional epicentral words of length $\leq n$. So, since ψ is injective and

$$\mathrm{card}\left(\bigcup_{i=0}^{\lfloor \log_2(n+1) \rfloor} A_k^i \right) = \sum_{i=0}^{\lfloor \log_2(n+1) \rfloor} k^i,$$

the assertion is proved. □

Corollary 8.10. *For any $k \geq 2$ the following holds:*

$$\hat{P}_k(n) = \Omega((n+1)^{\log_2 k}).$$

Proof. The result follows from the preceding theorem. It is sufficient to observe that

$$k^{\log_2(n+1)} = (n+1)^{\log_2 k}.$$

□

Proposition 8.11. *For any $k > 1$ and $n \geq 0$ one has:*

$$\hat{P}_k(2n+1) \geq k\hat{P}_k(n) + kn + 1.$$

Proof. Let $w = \psi(v)$ be an epicentral word of length $0 \leq |w| \leq n$ over the alphabet A_k. Then for any $x \in A_k$ the word $(wx)^{(+)} = \psi(vx)$ is an epicentral word of length $1 \leq |(wx)^{(+)}| \leq 2n+1$. Since the map ψ is injective, one has that from the set of all epicentral words of length $\leq n$ one can generate $k\hat{P}_k(n)$ epicentral words of length ≥ 1 and less than or equal to $2n+1$. Moreover, the words x^p with $x \in A_k$ and $n+2 \leq p \leq 2n+1$ are kn epicentral words which cannot be produced by the preceding procedure starting from epicentral words of length $\leq n$. Therefore, one has:

$$\sum_{j=1}^{2n+1} P_k(j) = \hat{P}_k(2n+1) - 1 \geq k\hat{P}_k(n) + kn,$$

which proves the assertion. □

Theorem 8.12. *Let $k \geq 2$. Then, for every $n \geq 1$,*

$$P_k(n) \geq k \sum_{\substack{d|(n+1) \\ d \neq n+1}} P_{k-1}(d-1).$$

Proof. Let d be a proper divisor of $n+1$, i.e., $d\,|(n+1)$ and $d \neq n+1$. We can write $n+1 = (m+1)d$ with $m > 0$. Consider any letter $a \in A_k$ and $v \in (A_k \setminus \{a\})^*$ such that $|\psi(v)| = d-1$. Since $a \notin \operatorname{alph} v$, we have

$$|(\psi(va^m)| = |\psi(v)a\psi(v)a \cdots a\psi(v)a\psi(v)| = (m+1)(d-1) + m = n.$$

Thus, for every proper divisor d of $n+1$ and for every word $v \in (A_k \setminus \{a\})^*$ such that $|\psi(v)| = d-1$, there exists for each possible choice of $a \in A_k$ a unique word of the type va^m such that $|\psi(va^m)| = n$. Since, by definition $m > 0$ (which makes all the above words distinct) and ψ is injective, we obtain

$$P_k(n) \geq \sum_{\substack{d|(n+1) \\ d \neq n+1}} kP_{k-1}(d-1)$$

and the assertion follows. □

Corollary 8.13. *Let $k \geq 3$. Then, for every $n \geq 1$,*

$$P_k(n) \geq \ k + \frac{k(k-1)}{2} \sum_{\substack{d|(n+1) \\ d \notin \{1,n+1\}}} \varphi(d+1).$$

Proof. This comes from Theorem 8.12 and from the fact that if $k \geq 3$, then from Corollary 8.5, for $n > 0$, $P_{k-1}(n) \geq \frac{k-1}{2}\varphi(n+2)$. Hence,

$$P_k(n) \geq kP_{k-1}(0) + k \sum_{\substack{d|(n+1) \\ d \notin \{1,n+1\}}} P_{k-1}(d-1) \geq k + \frac{k(k-1)}{2} \sum_{\substack{d|(n+1) \\ d \notin \{1,n+1\}}} \varphi(d+1).$$

□

Proposition 8.14. *Let $k \geq 2$. For any $n \geq 0$ the following holds:*

$$P_k(2n+1) \geq kP_{k-1}(n).$$

If $n \geq 2$, then

$$P_k(2n+1) \geq kP_{k-1}(n) + k^2.$$

Proof. From Theorem 8.12 we have:

$$P_k(2n+1) \geq k \sum_{\substack{d|(2n+2) \\ d \neq 2n+2}} P_{k-1}(d-1).$$

If $n \geq 2$, $2n+2$ has the three proper distinct divisors 1, 2, and $n+1$. Therefore, one has:

$$P_k(2n+1) \geq k(P_{k-1}(0) + P_{k-1}(1) + P_{k-1}(n)) = k + k(k-1) + kP_{k-1}(n)$$
$$= kP_{k-1}(n) + k^2.$$

In the case $n = 0$, $n+2$ has the only proper divisor 1 and in the case $n = 1$ the only two distinct proper divisors 1 and 2. Hence, one derives that for all $n \geq 0$, $P_k(2n+1) \geq kP_{k-1}(n)$. □

Corollary 8.15. *Let $k > 2$. The following holds:*

$$\sum_{m=0}^{n-1} P_k(2m+1) \geq k\hat{P}_{k-1}(n-1) = \Omega(n^{\log_2(k-1)}).$$

Proof. From the preceding proposition one has:

$$\sum_{m=0}^{n-1} P_k(2m+1) \geq k \sum_{m=0}^{n-1} P_{k-1}(m),$$

so that the result follows from Corollary 8.10. □

Lemma 8.16. *Let n be an odd integer and $k \geq 2$. The following holds:*

$$P_k(2n) \geq kP_{k-1}(\lfloor n/2 \rfloor).$$

Proof. Let x be any letter in A_k. Let $w = \psi(v)$ be any epicentral word having a unique occurrence of the letter x. Since w is a palindrome, we can write $w = uxu^{\sim}$, with $u \in (A_k \setminus \{x\})^*$, so that $|w| = n$ is an odd integer. Moreover, as $x \in \operatorname{alph} v$ and x has a unique occurrence in w, it follows that u is a palindromic prefix of w and, therefore, an epicentral word over A_{k-1}. From Justin's formula one has that $\psi(xv) = \mu_x(\psi(v))x$. Since $|\mu_x(u)| = 2|u| = n-1$, one has

$$|\psi(xv)| = |\mu_x(uxu)| + 1 = 2|\mu_x(u)| + 2 = 2(n-1) + 2 = 2n.$$

Since for any fixed letter in A_k the number of epicentral words with a unique occurrence of the letter x having a length equal to the odd integer n is $P_{k-1}(\lfloor n/2 \rfloor)$, it follows from the injectivity of ψ that $P_k(2n)$ has the lower bound $kP_{k-1}(\lfloor n/2 \rfloor)$. □

Proposition 8.17. *Let $k > 2$. For $n \geq 2$ the following holds:*

$$\sum_{m=0}^{n-1} P_k(2m) \geq k\hat{P}_{k-1}(\lfloor (n-2)/2 \rfloor) = \Omega(n^{\log_2(k-1)}).$$

Proof. By the preceding lemma one has:

$$\sum_{m=0}^{n-1} P_k(2m) \geq \sum_{\substack{m=1\\ m \text{ odd}}}^{n-1} P_k(2m) \geq k \sum_{\substack{m=1\\ m \text{ odd}}}^{n-1} P_{k-1}(\lfloor m/2 \rfloor).$$

If n is even, then $n-1$ is odd and the set

$$\{\lfloor k/2 \rfloor \mid 1 \leq k \leq n-1,\ k \text{ odd}\} = \{0, 1, 2, \ldots, \lfloor (n-1)/2 \rfloor\}.$$

Hence, one has if n is even:

$$\sum_{\substack{m=1\\ m \text{ odd}}}^{n-1} P_{k-1}(\lfloor m/2 \rfloor) = \sum_{m=0}^{\lfloor (n-1)/2 \rfloor} P_{k-1}(m) = \hat{P}_{k-1}(\lfloor (n-1)/2 \rfloor).$$

If n is odd, then $n-1$ is even and $n-2$ is odd so that, by the previous argument, one has

$$\sum_{m=0}^{n-1} P_k(2m) \geq \sum_{m=0}^{n-2} P_k(2m) \geq k\hat{P}_{k-1}(\lfloor (n-2)/2 \rfloor).$$

For any n, $\hat{P}_{k-1}(\lfloor (n-1)/2 \rfloor) \geq \hat{P}_{k-1}(\lfloor (n-2)/2 \rfloor)$, so that in view of Corollary 8.10 the result follows. □

Corollary 8.18. *For any $k > 2$, the following holds:*

$$g_k(n) = \Omega(n^{\log_2(k-1)}).$$

Proof. From Theorem 5.3 one has that

$$g_k(2n) = \Omega(\sum_{m=0}^{n-1} P_k(2m)) \quad \text{and} \quad g_k(2n+1) = \Omega(\sum_{m=0}^{n-1} P_k(2m+1)).$$

Hence, the result follows from Corollary 8.15 and Proposition 8.17. □

Remark 8.19. The lower bound of preceding corollary is meaningful only if k is quite large. We recall that it was proved in [5] that $g_2(n) = \Omega(n^{1+\alpha})$, with $\alpha = \log_3 2 = 0.6309\cdots$. Hence, since $g_k(n) \geq g_2(n)$ for all $k \geq 2$ and $n > 0$, one trivially has $g_k(n) = \Omega(n^{1+\alpha})$. However, for all $k > 4$ one has that $n^{\log_2(k-1)} > n^{1+\alpha}$.

8.3 Counting procedure

As observed in Section 7, the map ω associating with any $v \in A^*$ the Parikh vector $\omega(v)$ of $\psi(v)$ is injective. Therefore, counting epicentral words of a given length n over A_k is equivalent to counting the vectors $\omega(v)$ such that $v \in A_k^*$ and

$$|\omega(v)| = \sum_{i=1}^{k} \omega_i(v) = n.$$

In view of Proposition 8.1, the counting map P_k can be obtained from the number of *canonical* epicentral words, which have by Proposition 7.4 decreasing Parikh vectors (i.e., such that $\omega_i(v) \geq \omega_{i+1}(v)$ for $i = 1, \ldots, k-1$); a vector corresponding to a canonical epicentral word $\psi(v)$ with $\mathrm{card}(\mathrm{alph}\, v) = j < k$ will have $k - j$ trailing zero components.

A convenient approach for calculating many values of the functions P_k (or g_k etc.) by a computer is then to generate all such vectors for a suitable k ($k = 8$ was used to compute the data in the Appendix) and all lengths $n \leq 2^k - 1$. By Proposition 7.7 at most k distinct letters can occur in an epicentral word of such a length n; this allows us to calculate the values of $U_i(n)$, $P_i(n)$, $g_i(n)$, and $\lambda_i(n)$ for all $i \geq 1$ and $n \leq 2^k - 1$ by straightforward calculations, using the definition of the maps U, P, g, and λ, as well as equations (4), (7), and Proposition 8.1. Thus, in particular, for $n \leq 2^k - 1$ one has for all $j > k$, $U_j(n) = 0$ and by Proposition 8.1

$$P_j(n) = \sum_{i=1}^{k} \frac{j!}{(j-i)!} U_i(n).$$

The main iteration for computing the Parikh vectors of canonical epicentral words is a consequence of Proposition 7.5. For any letter j, every vector $\omega(v) = (\omega_1(v), \ldots, \omega_k(v))$ gives rise to new ones of the form

$$(|\omega(v)| + 1, \omega_1(v), \ldots, \omega_{j-1}(v), \omega_{j+1}(v), \ldots, \omega_k(v)).$$

By Proposition 7.5 these are indeed the Parikh vectors of the canonical epicentral words similar to $\psi(a_j v)$. At each step, new vectors whose corresponding lengths are greater than $2^k - 1$ are discarded. In this way, starting with the zero vector $\omega(\varepsilon)$, one obtains all desired vectors, up to length $2^k - 1$, after which the procedure stops since no new vector can be kept.

9 Conclusions

For any k, the map P_k which counts for any n the epicentral words of length n on the k-letter alphabet, is a suitable extension to the case $k > 2$ of Euler's totient function. Indeed, for $k = 2$, $P_2(n) = \varphi(n+2)$; for $k > 2$ a general arithmetic interpretation for $P_k(n)$ in terms of a multidimensional generalization of the Euclidean algorithm is in [14] (see also [7]).

As it appears from Tab. 1, the behavior of P_k for any fixed k is quite irregular as n increases. However, some remarkable properties seem to be satisfied by $P_k(n)$ even though we are not able to prove them. Therefore, we set the following conjectures.

Conjecture 9.1. *Let $k \geq 3$ and $n > 0$. Then $P_k(n) \geq n - 1$.*

We observe that, since for all k and n, $P_{k+1}(n) \geq P_k(n)$ if the conjecture is true for $k = 3$, then it is true for all $k > 3$. Moreover, the lower bound $n - 1$ is tight since, for example, $P_3(10) = 9$. Note that for $k = 2$, the statement of conjecture is false. For instance, $P_2(6) = \varphi(8) = 4 < 5$.

One can prove that the conjecture is true for all $n = p^m - 2$ where p is an odd prime and $m \geq 1$. This is an immediate consequence of the fact that from Corollary 8.5, $P_3(n) \geq (3/2)\varphi(n + 2)$ and of the relation $\varphi(p^m) = p^{m-1}(p - 1)$. Moreover, the conjecture has been verified by using a computer for $3 \leq k \leq 8$ and $1 \leq n \leq 510$.

The next conjecture is based on the observation that for $k > 2$ the number of epicentral words of odd length seems to be quite larger than that of even length. For instance, $P_6(63) = 18936$, whereas $P_6(62) = 5256$ and $P_6(64) = 1716$. More precisely, we formulate the following:

Conjecture 9.2. *For $k \geq 3$ and $n > 0$ one has:*

$$P_k(2n \pm 1) \geq P_k(2n).$$

Let us first remark that the statement of the preceding conjecture does not hold for $k = 2$, Indeed, one has $P_2(313) = 144 < 156 = P_2(312) = P_2(314)$. By a computer the conjecture was verified for $3 \leq k \leq 8$ and $1 \leq n \leq 254$.

We observe that if Conjecture 9.2 is true, then one would derive from (5) and (6) that for all $k \geq 3$ and $n > 0$, $g_k(2n) \leq g_k(2n \pm 1)$. This relation is satisfied for the values of n and k reported in Tab. 2.

A further conjecture concerns the density $P_k(n)/g_k(n)$ of the epicentral words of length n over A_k with respect to all episturmian palindromes of length n. From the tables it seems that the preceding ratio vanishes when n diverges.

Conjecture 9.3. *For all $k \geq 3$*

$$\lim_{n\to\infty} \frac{P_k(n)}{g_k(n)} = 0.$$

In the case $k = 2$ the result is true as proved in [5].

References

[1] J. Berstel, D. Perrin, *Theory of Codes*, Academic Press, New York, 1985

[2] J. Berstel, Sturmian and Episturmian words (A survey of some recent results), Lecture Notes in Computer Science, vol. 4728, Springer-Verlag (Berlin 2007) pp. 23–47

[3] M. Bucci, A. de Luca, A. De Luca, L.Q. Zamboni, On some problems related to palindrome closure, *Theoretical Informatics and Applications* **42** (2008) 679–700

[4] A. de Luca, Sturmian words: Structure, Combinatorics, and their Arithmetics, *Theoretical Computer Science* **183** (1997) 45–82

[5] A. de Luca, A. De Luca, Combinatorial properties of Sturmian palindromes, *International Journal of Foundations of Computer Science* **17** (2006) 557–573

[6] A. de Luca, F. Mignosi, On some combinatorial properties of Sturmian words, *Theoretical Computer Science* **136** (1994) 361–385

[7] A. de Luca, L. Q. Zamboni, On the graphs of central episturmian words, *Theoretical Computer Science* **411** (2010) 70–90, doi: 10.1016/j.tcs.2009.09.003

[8] A. de Luca, L.Q. Zamboni, Involutions of epicentral words, *European Journal of Combinatorics* **31** (2010) 867–886, doi:10.1016/j.ejc.2009.07.002

[9] X. Droubay, J. Justin, and G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, *Theoretical Computer Science* **255** (2001) 539–553

[10] A. Glen, J. Justin, Episturmian words: A survey, *Theoretical Informatics and Applications* **43** (2009) 403–442

[11] J. Justin, Episturmian morphisms and a Galois theorem on continued fractions, *Theoretical Informatics and Applications* **39** (2005) 207–215

[12] M. Lothaire, *Combinatorics on Words*, Addison-Wesley (Reading, MA, 1983)

[13] M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press (Cambridge, 2002)

[14] F. Mignosi, L.Q. Zamboni, On the number of Arnoux-Rauzy words, *Acta Arithmetica* **101** (2002) 121–129

[15] G. Richomme, A local balance property of episturmian words, in Lecture Notes in Computer Science, vol. 4588, Springer, Berlin, 2007, pp. 371–381

[16] J. Riordan, *An Introduction to Combinatorial Analysis*, Princeton University Press, 1978

Appendix

Table 1: First values of $P_k(n)$, for $3 \leq k \leq 6$.

n	$P_3(n)$	$P_4(n)$	$P_5(n)$	$P_6(n)$
0	1	1	1	1
1	3	4	5	6
2	3	4	5	6
3	9	16	25	36
4	3	4	5	6
5	15	28	45	66
6	9	16	25	36
7	21	52	105	186
8	9	16	25	36
9	27	52	85	126
10	9	16	25	36
11	51	136	285	516
12	15	28	45	66
13	33	88	185	336
14	27	64	125	216
15	63	184	445	936
16	15	28	45	66
17	69	172	345	606
18	21	40	65	96
19	75	232	525	996
20	39	100	205	366
21	87	220	445	786
22	33	88	185	336
23	129	496	1385	3156
24	33	64	105	156
25	87	244	525	966
26	45	112	225	396
27	129	424	1105	2436
28	33	88	185	336
29	159	508	1245	2586
30	57	160	385	816
31	147	568	1685	4236
32	63	160	325	576
33	129	376	825	1536
34	51	136	285	516
35	249	928	2505	5556
36	63	148	285	486
37	141	424	945	1776
38	93	280	625	1176

n	$P_3(n)$	$P_4(n)$	$P_5(n)$	$P_6(n)$
39	243	976	2845	6696
40	57	160	345	636
41	237	844	2265	5046
42	75	184	365	636
43	225	928	2625	5976
44	123	412	1045	2226
45	219	748	2005	4506
46	75	280	805	1896
47	357	1684	5865	16446
48	69	184	425	876
49	243	784	1805	3456
50	111	304	645	1176
51	309	1216	3505	8196
52	87	244	525	966
53	327	1264	3445	7656
54	117	376	945	2016
55	351	1576	5205	14076
56	141	400	865	1596
57	327	1108	2845	6126
58	105	352	865	1776
59	495	2344	7725	20196
60	135	412	1005	2106
61	285	1024	2865	6996
62	189	712	2105	5256
63	405	1912	6585	18936
64	129	400	905	1716
65	495	1876	5085	11286
66	141	376	785	1416
67	399	1792	5405	12876
68	159	592	1605	3576
69	465	1876	5425	12666
70	153	544	1425	3096
71	663	3472	12645	35976
72	141	400	945	1956
73	369	1384	3625	7776
74	225	736	1785	3636
75	513	2296	6945	16596
100	213	712	1745	3576
500	3021	18748	78745	259446

Table 2: First values of $g_k(n)$, for $3 \leq k \leq 6$.

n	$g_3(n)$	$g_4(n)$	$g_5(n)$	$g_6(n)$	n	$g_3(n)$	$g_4(n)$	$g_5(n)$	$g_6(n)$
0	1	1	1	1	42	1359	5140	14085	31806
1	3	4	5	6	43	4209	20920	71305	192936
2	3	4	5	6	44	1509	5692	15545	34986
3	9	16	25	36	45	4659	23704	81805	222816
4	9	16	25	36	46	1755	6928	19725	46116
5	27	64	125	216	47	5097	25948	89825	245346
6	15	28	45	66	48	1905	7768	22945	55596
7	57	148	305	546	49	5811	31000	113285	327576
8	33	76	145	246	50	2043	8320	24645	59976
9	99	304	725	1476	51	6297	33352	120505	344856
10	51	124	245	426	52	2265	9232	27225	65856
11	153	460	1065	2106	53	6915	37000	134525	385836
12	69	172	345	606	54	2439	9964	29325	70686
13	255	868	2205	4686	55	7569	40792	148305	424116
14	99	256	525	936	56	2673	11092	33105	80766
15	321	1132	2945	6366	57	8271	45520	169125	494496
16	153	448	1025	2016	58	2955	12292	36565	88746
17	447	1684	4725	11046	59	8925	48844	180505	525126
18	183	532	1205	2346	60	3165	13348	40025	97626
19	585	2200	6105	14076	61	9915	55876	211405	626106
20	225	652	1465	2826	62	3435	14584	44045	108156
21	735	2896	8205	19056	63	10485	58948	222865	661086
22	303	952	2285	4656	64	3813	16720	52465	134436
23	909	3556	9985	22986	65	11295	64684	249205	755766
24	369	1216	3025	6336	66	4071	17920	56085	143016
25	1167	5044	15525	38766	67	12285	70312	269545	812196
26	435	1408	3445	7116	68	4353	19048	59225	150096
27	1341	5776	17625	43596	69	13083	75688	291165	876576
28	525	1744	4345	9096	70	4671	20824	65645	167976
29	1599	7048	22045	55776	71	14013	81316	312865	939906
30	591	2008	5085	10776	72	4977	22456	71345	183456
31	1917	8572	27025	68706	73	15339	91732	363445	1119786
32	705	2488	6625	14856	74	5259	23656	75125	193236
33	2211	10276	33765	89886	75	16077	95884	377945	1158666
34	831	2968	7925	17736	76	5709	25864	82265	211416
35	2469	11404	37065	97566	77	17103	102772	405725	1241646
36	933	3376	9065	20316	78	6027	27640	88685	229296
37	2967	14188	47085	125346	79	18285	110344	435665	1330176
38	1059	3820	10205	22746	80	6369	29632	96385	251976
39	3249	15460	50865	134226	100	11073	55564	192825	534426
40	1245	4660	12705	28626	201	178311	1457440	7380325	27922956
41	3735	18388	62245	167706	500	550941	4515952	22922985	86857176

Table 3: First values of $U_k(n)$, for $3 \leq k \leq 6$.

n	$U_3(n)$	$U_4(n)$	$U_5(n)$	$U_6(n)$
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	1	0	0	0
8	0	0	0	0
9	0	0	0	0
10	0	0	0	0
11	3	0	0	0
12	0	0	0	0
13	2	0	0	0
14	1	0	0	0
15	3	1	0	0
16	0	0	0	0
17	3	0	0	0
18	0	0	0	0
19	7	0	0	0
20	2	0	0	0
21	4	0	0	0
22	2	0	0	0
23	12	4	0	0
24	0	0	0	0
25	6	0	0	0
26	2	0	0	0
27	8	3	0	0
28	2	0	0	0
29	12	2	0	0
30	2	1	0	0
31	15	4	1	0
32	3	0	0	0
33	10	0	0	0
34	3	0	0	0
35	24	6	0	0
36	2	0	0	0
37	12	0	0	0
38	8	0	0	0

n	$U_3(n)$	$U_4(n)$	$U_5(n)$	$U_6(n)$
39	21	10	0	0
40	4	0	0	0
41	19	6	0	0
42	3	0	0	0
43	26	7	0	0
44	10	2	0	0
45	14	6	0	0
46	5	3	0	0
47	39	21	5	0
48	2	1	0	0
49	25	0	0	0
50	7	0	0	0
51	26	12	0	0
52	6	0	0	0
53	35	8	0	0
54	8	2	0	0
55	41	16	4	0
56	10	0	0	0
57	26	6	0	0
58	10	1	0	0
59	53	30	3	0
60	8	2	0	0
61	30	4	2	0
62	16	6	1	0
63	44	24	5	1
64	12	0	0	0
65	50	12	0	0
66	8	0	0	0
67	45	19	0	0
68	15	4	0	0
69	43	18	0	0
70	14	3	0	0
71	75	52	10	0
72	6	2	0	0
73	42	6	0	0
74	20	2	0	0
75	56	25	0	0
100	20	2	0	0
500	379	340	106	16

Table 4: First values of $\lambda_k(n)$, for $3 \leq k \leq 6$.

n	$\lambda_3(n)$	$\lambda_4(n)$	$\lambda_5(n)$	$\lambda_6(n)$	n	$\lambda_3(n)$	$\lambda_4(n)$	$\lambda_5(n)$	$\lambda_6(n)$
0	1	1	1	1	39	101715	633484	2551405	7923846
1	3	4	5	6	40	110697	693832	2805665	8738076
2	9	16	25	36	41	120651	762964	3105445	9719706
3	27	64	125	216	42	130833	833536	3410745	10717236
4	57	148	305	546	43	141963	911704	3752285	11840916
5	123	376	885	1776	44	153393	991528	4099665	12980496
6	201	640	1545	3156	45	165723	1079704	4489045	14269476
7	339	1156	2925	6186	46	178545	1171588	4895145	15614106
8	513	1816	4705	10116	47	192243	1270204	5333325	17071386
9	771	2944	8165	18696	48	206241	1371340	5784385	18576066
10	1065	4216	12025	28176	49	221667	1487632	6329285	20491896
11	1467	5956	17245	40806	50	237369	1605580	6880985	22429626
12	1905	7840	22865	54336	51	254043	1730584	7461565	24453756
13	2547	10948	33045	80766	52	271161	1858324	8052465	26507286
14	3249	14308	43945	108846	53	289515	1997008	8699445	28765716
15	4083	18460	57805	145326	54	308217	2137888	9354825	31048296
16	5025	23188	73665	187206	55	328227	2290144	10065325	33522276
17	6219	29572	96645	252486	56	348705	2445784	10790945	36046656
18	7473	36208	120345	319416	57	370587	2615608	11599845	38922936
19	9003	44392	149565	401496	58	393033	2789032	12422585	41839116
20	10617	52936	179825	485976	59	416787	2972428	13290845	44908446
21	12531	63568	218485	595356	60	440961	3158992	14172945	48022176
22	14601	75100	260425	713886	61	467115	3366652	15178645	51640806
23	17019	88612	309485	852066	62	493809	3578020	16200425	55312086
24	19569	102916	361505	998646	63	521643	3798604	17268045	59158266
25	22635	121684	435685	1224126	64	550233	4025596	18369345	63135846
26	25833	141028	511545	1453506	65	580443	4269796	19576005	67586826
27	29379	162568	595805	1707036	66	611169	4517596	20797145	72080706
28	33105	185116	683665	1970466	67	643875	4782280	22099645	76856736
29	37347	211480	789205	2294796	68	677145	5050348	23414705	81668166
30	41721	238636	897705	2627526	69	712011	5334544	24816245	86801496
31	46731	270364	1026125	3024906	70	747513	5624068	26243465	92024226
32	51969	303532	1160705	3442686	71	784875	5930476	27757485	97563606
33	57795	341812	1322245	3966366	72	822849	6241780	29294305	103180386
34	63873	381532	1488985	4504446	73	863475	6584332	31033445	109696566
35	70467	424636	1668925	5080926	74	904665	6930484	32787705	116261646
36	77265	468964	1853425	5670306	75	947331	7289092	34599965	123021126
37	85059	521644	2078005	6398586	100	2534937	21417592	109353905	412367376
38	93105	575656	2307145	7139016	200	27341865	287631016	1744815265	7594787376