FINE AND WILF WORDS FOR ANY PERIODS II

R. Tijdeman and L. Zamboni

Abstract. In 2003 the authors introduced a fast algorithm to determine the word $w = w_1...w_n$ of maximal length n, and with a maximal number of distinct letters for this length, such that w has prescribed periods $p_1, ..., p_r$, but not period $\gcd(p_1, ..., p_r)$. They proved that the constructed word is uniquely determined apart from isomorphism and that it is a palindrome. In the present paper the algorithm is extended in such a way that for any given positive integers $n, p_1, \ldots, p_r$ a word w of length n with periods $p_1, \ldots, p_r$ with a maximal number of distinct letters for this length can be constructed. The constructed words are pseudo-palindromes.

1. Introduction

Let $w = w_1...w_n$ be a word on a finite non-empty alphabet A. We denote the length of w by $|w|$. The empty word, denoted ϵ, is the unique word of length 0. A positive integer p is said to be a *period* of w if $w_{i+p} = w_i$ for each $i = 1, ..., n-p$. In 1965 Fine and Wilf [FW] showed that if w is a word having distinct periods p_1 and p_2 and $|w| \geq p_1 + p_2 - \gcd(p_1, p_2)$, then the $\gcd(p_1, p_2)$ is also a period of w. They further showed that if $\gcd(p_1, p_2) \notin \{p_1, p_2\}$, then there exists a word of length $p_1 + p_2 - \gcd(p_1, p_2) - 1$ with periods p_1 and p_2 but not the $\gcd(p_1, p_2)$. In case p_1 and p_2 are relatively prime, this word is unique up to letter to letter isomorphism and is known to be a palindrome and a bispecial factor of an infinite Sturmian word. In 1999 Castelli, Mignosi and Restivo [CMR] obtained an analogous result in the case of three periods p_1, p_2, p_3 : they showed that there exists a constant L (depending on p_1, p_2, p_3) such that any word w with periods p_1, p_2, p_3 and of length $|w| \geq L$, necessarily also has period $\gcd(p_1, p_2, p_3)$, and moreover if $\gcd(p_1, p_2, p_3) \notin \{p_1, p_2, p_3\}$, then there exists a word of length $L - 1$, having periods p_1, p_2, p_3 but not $\gcd(p_1, p_2, p_3)$. Their result was later generalized to any number of periods by Justin [J]. Let $P = \{p_1, p_2, \ldots, p_r\}$ be a set consisting of r distinct positive integers. We call P a *period set* if the $\gcd(p_1, p_2, \ldots, p_r) \notin \{p_1, p_2, \ldots, p_r\}$. For each period set P we denote by $L(P)$ the least positive integer L such that any word w of length $|w| \geq L(P)$ with periods $p_1, p_2, \ldots, p_r$ also has period $\gcd(p_1, p_2, \ldots, p_r)$.

In 2003 the authors [TZ] introduced a fast algorithm for constructing extremal Fine and Wilf words relative to any period set $P = \{p_1, p_2, \ldots, p_r\}$, i.e., words of length $L(P) - 1$ having periods $p_1, p_2, \ldots, p_r$ and on the most number of distinct symbols. This same construction was later rediscovered by Constantinescu and Ilie [CI] and by Š. Holub [H]. There are several reasons for requiring these words to be on the most number of symbols. First of all, it is shown in [TZ] that such

2000 Mathematics Subject Classification: ??

words are unique up to word isomorphism. Secondly, any other word of the same length having periods $p_1, p_2, \ldots, p_r$ is the morphic image (under a letter to letter morphism) of the extremal Fine and Wilf word. The authors also showed that as in the case of two periods, if the $\gcd(p_1, \ldots, p_r) = 1$, then the non-constant word of maximal length having periods $\{p_1, \ldots, p_r\}$ and on the most number of symbols, is a palindrome.

Let w be a finite word on a non-empty alphabet A. We define $\mathrm{Alph}(w)$ to be the set of distinct symbols of A occurring in w, and $\sharp w$ the number of distinct symbols occurring in w. Let $P = \{p_1, p_2, \ldots, p_r\}$ be a set of positive integers. We say w is a *Fine and Wilf* word (or FW word for short) *relative to* P if w has periods $p_1, \ldots, p_r$ and given any other word v of the same length as w having periods $p_1, \ldots, p_r$, we have $\sharp w \geq \sharp v$. For instance, it is readily verified that $w = abcabcababcabcab$ is a FW word relative to the set $\{8, 11\}$. We note in this example that 14 is also a period of the word, thus w is also a FW word relative to the set $\{8, 11, 14\}$. However the period 14 in this case is a consequence of the other two periods. We call a word w an FW word if w is an FW word relative to some set P.

A FW word w relative to a period set $P = \{p_1, p_2, \ldots, p_r\}$ is said to be *extremal* if $|w| = L(P) - 1$, and *trivial* if $|w| \geq L(p)$. Thus, a trivial FW word also has period $\gcd(p_1, p_2, \ldots, p_r)$, while the extremal FW word is the longest FW word having periods $p_1, p_2, \ldots, p_r$ but not $\gcd(p_1, p_2, \ldots, p_r)$. For instance, returning to the example above, the FW word $w = abcabcababcabcab$ relative to periods $\{8, 11\}$ is not extremal since the longer word (of length 17) $w' = abaabaababaabaaba$ also has periods $\{8, 11\}$ but not period $\gcd(8, 11)$. However, it follows from the Fine and Wilf theorem that w' is an extremal FW word relative to the period set $\{8, 11\}$. Note that the extremal FW word w' is a palindrome, while the non-extremal FW word w is not a palindrome. We observe however that the reverse of w (denoted $\bar{w}$) is equal to w with all a's and b's exchanged. The word w is an example of a so-called pseudo-palindrome [AZZ, LL,BLLZ1, BLLZ2]. More precisely, a finite word w on a finite alphabet A is called a *pseudo-palindrome* if w is a fixed point of an involutory antimorphism θ of the free monoid A^*. We recall that an involutory antimorphism is given by a map $\theta : A^* \to A^*$ such that $\theta \circ \theta = \mathrm{id}$, and satisfying $\theta(uv) = \theta(v)\theta(u)$ for any $u, v \in A^*$. The reversal operator

$$R : w \in A^* \mapsto \bar{w} \in A^*$$

is the most basic example. Any involutory antimorphism is the composition $\theta = \tau \circ R = R \circ \tau$ where τ is an involutory permutation of the alphabet A.

In the present paper we give a fast word combinatorial algorithm for constructing FW words of all lengths, i.e., given a set P and a positive integer n we construct a FW word of length n relative to P. This algorithm is an extension of the one originally given in [TZ].

In Section 2 we describe two naive approaches in the construction of FW words. In Section 3 we present the algorithm which computes the FW word w for given length n relative to a given set P, and state the principal theorem showing that the algorithm produces the required word, that this word is unique up to word isomorphism, and is always a pseudo-palindrome. In Section 4 we prove the main theorem, and in Section 5 we present some examples. The authors plan to publish another paper in the near future in which the properties of FW words are further investigated.

2. Naive approach to constructing FW words

In this section we describe two approaches for constructing FW words of a given length relative to a given set. The first approach is as follows: Suppose we want to construct a FW word of length 16 relative to the period set $\{8, 11\}$. We start with a 'blank' word of length 16 consisting of 16 dashes:

$$- - - - - - - - - - - - - - - -$$

Then we assign the value a to the first position and use the periods 8 and 11 to determine all other locations for the value a:

$$a - -a - -a - a - -a - -a-$$

Next we assign the value b to the first vacant position (in this case position 2) and again use the periods to determine all other locations for the value b:

$$ab - ab - abab - ab - ab$$

Next we assign the value c to the first vacant position (in this case position 3) and use the periods to determine all other locations for the value c:

$$abcabcababcabcab$$

This approach is easily generalized to any number of periods. We simply keep assigning a new symbol to the first vacant slot and use the periods to determine all locations for that symbol. The process ends when there are no more vacant positions. This construction suggests that the resulting word is always unique up to word isomorphism.

A graph-theoretic analogue is the following: Given a period set P and positive integer n (representing the length of the desired word), we construct a graph $G_n(P)$ whose vertices are the integers $\{1, 2, \ldots, n\}$, and for each $p \in P$, we put an undirected edge labeled p between vertices x and y if and only if $x - y = \pm p$. We then assign a distinct symbol to each connected component of $G_n(P)$ and construct a word of length n, whose ith entry is simply the symbol assigned to the connected component of $G_n(P)$ containing vertex i. For instance, in the example above with $P = \{8, 11\}$ and $n = 16$, we find $G_n(P)$ consists of three connected components: one component contains vertices $\{7, 15, 4, 12, 1, 9\}$, another contains vertices $\{8, 16, 5, 13, 2, 10\}$, and the third vertices $\{6, 14, 3, 11\}$. If we assign the value a to the first component, b to the second and c to the third, we obtain the FW word w above. If we were to repeat the same process for $n = 17$, the new vertex 17 would form a link between the component containing 9 and the component containing 6 so that $G_{17}(P)$ would consist of two connected components, whence the resulting FW word would be the binary word $w' = abaabaababaabaaba$. Finally for $n = 18$, we would have a single connected component, whence the FW word would simply be the constant word a^{18}.

3. The algorithm

Let $P = \{p_1, ..., p_r\}$ be a period set, and n a positive integer. We now present an algorithm to construct a FW word w of length n relative to the period set P. We shall see that this word is unique up to word isomorphism.

THE ALGORITHM

Input: positive integers $n, p_1, \dots, p_r$.

Reduction

(R0) (Initialization) for $i = 1, \dots, r$ put $p_i[0] := p_i$; put $k := 0, n[0] := n$.

(R1) Let i be the smallest value with $p_i[k] = \min\{\ p_j[k] \mid p_j[k] > 0; j = 1, \dots, r\}$.

(R2) If $n[k] \geq p_i[k]$, then put $p[k] := p_i[k]$, else goto (E0).

(R3) For $j \in \{1, ..., r\}$, $j \neq i$ do, if $p_j[k] > 0$ then $p_j[k+1] := p_j[k] - p[k]$, else $p_j[k+1] := 0$.

(R4) Put $p_i[k+1] := p_i[k], g[k] := i, n[k+1] := n[k] - p[k]; k := k+1$; goto (R1).

Extension

(E0) (Initialization) put $K := k; N := n[K]; v[0] := v_{01} \dots v_{0N}; w[K] := v[0]$;
for $j = 1, \dots, r$ if $p_j[K] > N$ put $h[j] := p_j[K] - N, v[j] := v_{j1} v_{j2} \dots v_{jh[j]}$, else put $h[j] := 0, v[j] := \epsilon$.

(E1) For $k = K-1, K-2, \dots, 0$,
if $n[k] > 2n[k+1]$ put $w[k] := w[k+1]v[g[k]]w[k+1]$, else put $w[k] := w[k+1]w'[k+1]$ where $w'[k+1]$ is the suffix of $w[k+1]$ of length $p[k]$.

(E2) Output $w := w[0]$.

Observe that only those factors $v[j]$ appear in w for which $j = g[k]$ for some k, i.e. column j contains an underlined number. For practical reasons we write the periods p_j in increasing order. We shall prove

Theorem 1. *Let $P = \{p_1, p_2, \dots, p_r\}$ be a period set, and n a positive integer. The above algorithm constructs a FW word w of length n relative to the period set P. The word w is unique apart from isomorphism. It is pseudo-palindromic with respect to the involutary antimorphism ϕ induced by reversion of the $v[j]$'s.*

We present an example first and give the proof in the next section.

Example 1 We construct the FW word w of length 386 relative to the period set $P = \{189, 288, 336, 362\}$. We label the periods: $p_1 = 189, p_2 = 288, p_3 = 336, p_4 = 362$. First we construct a tower as in [TZ] according to the Euclidean algorithm starting with p_1, p_2, p_3, p_4, and n. At each level j, we identify the first 'period' column i (counting from the left) containing the smallest positive p-value. In passing from level j to level $j+1$, this smallest p-value is subtracted from each of the positive values in the other columns (meaning columns different from i), while the entry p in column i remains unchanged from level j to level $j+1$. Any 0 in a period column at level j stays 0 at level $j+1$. The process terminates when the current n-value becomes negative. In Example 1 we stop at level 6, since $p[6] = 3 > 2 = n[6]$.

k	$p_1[k]$	$p_2[k]$	$p_3[k]$	$p_4[k]$	$n[k]$	$g[k]$
0	$\underline{189}$	288	336	362	386	1
1	189	$\underline{99}$	147	173	197	2
2	90	99	$\underline{48}$	74	98	3
3	42	51	48	$\underline{26}$	50	4
4	$\underline{16}$	25	22	26	24	1
5	16	9	$\underline{6}$	10	8	3
6	10	3	6	4	2	

Table 1: Result of the Reduction procedure

We obtain $n[0] = 386, n[1] = 197, \ldots, n[6] = N = 2, g[0] = 1, g[1] = 2, g[2] = 3, g[3] = 4, g[4] = 1, g[5] = 3, p[0] = 189, p[1] = 99, p[2] = 48, p[3] = 26, p[4] = 16, p[5] = 6$. Next we construct a 'seed' of words $v[0], v[1], \ldots, v[r]$ consisting of $N = 2, h[1] = 8, h[2] = 1, h[3] = 4, h[4] = 2$ distinct letters, respectively, $v[0] := xy, v[1] := abcdefgh, v[2] := i, v[3] := jklm, v[4] := no$, say. Then we construct the FW word in seven steps where we concatenate $w[k+1]$ with another word to obtain $w[k]$ so that we reach the length $n[k]$:

$w[6] := v[6]$ (of length 2)
$w[5] := w[6]v[g[5]]w[6]$ (of length 8)
$w[4] := w[5]v[g[4]]w[5]$ (of length 24)
$w[3] := w[4]v[g[3]]w[4]$ (of length 50)
$w[2] := w[3]v[g[5]]w[6]v[g[4]]w[5]v[g[3]]w[4]$ (of length 98)
$w[1] := w[2]v[g[1]]w[2]$ (of length 197)
$w[0] := w[1]v[g[4]]w[5]v[g[3]]w[4]v[g[5]]w[6]v[1g[4]]w[5]v[g[3]]w[4]v[g[1]]w[2]$ (of the required length 386).

Note that $g[5] = g[2]$ and $g[4] = g[0]$ so that the suffix $w'[k]$ of $w[k]$ is exactly the subword which was concatenated after a previous time that column $g[k]$ was chosen for underlining. We conclude that $\sharp w$ equals $2 + 8 + 1 + 4 + 2 = 17$ and that w begins in $xy|jklmxy|abcdefghxyjklmxy|noxyjklmxyabcdefghxyjklmxy|$ where the bars should be ignored, but are inserted to indicate the end of the prefixes $w[6], w[5], w[4], w[3]$ of w. Observe that w is pseudo-palindromic with respect to the function ϕ defined by $\phi(x) = y, \phi(y) = x, \phi(a) = h, \phi(b) = g, \phi(c) = f, \phi(d) = e, \phi(e) = d, \phi(f) = c, \phi(g) = b, \phi(h) = a, \phi(i) = i, \phi(j) = m, \phi(k) = l, \phi(l) = k, \phi(m) = j, \phi(n) = o, \phi(o) = n$.

4. Proof of Theorem 1

For the proof of Theorem 1 we need some lemmas. Lemma 1 is a variant of a lemma of Castelli, Mignosi and Restivo [CMR].

Lemma 1. *Let $u = u_1...u_m$ be a word with s distinct letters and periods $q_1 < ... < q_r$. Put $u' := u_1...u_{m-q_1}$. If $m \geq 2q_1 - y$, with $0 \leq y < q_1$, then u' is a word with at least $s - y$ distinct letters and periods $q_1, q_2 - q_1, ..., q_r - q_1$.*

Proof. Because of the period q_1, every letter of u occurs in u' except possibly the y letters $u_{q_1-y+1}, u_{q_1-y+2}, \ldots, u_{q_1}$. Hence the number of distinct letters in u' is at least $s - y$. For $t \leq m - q_j$ we have $u_t = u_{t+q_j} = u_{t+q_j-q_1}$. So u' has period $q_j - q_1$ for $j = 2, ..., r$. □

Lemma 2 is a counterpart to Lemma 1 and appears in modified form in [TZ].

Lemma 2. *Suppose $u = u_1...u_m$ has periods $q_1, ..., q_r$. Let $u_{m+1}, \ldots, u_{m+q_1}$ satisfy $u_{m+i} = u_{m+i-q_1}$ for $i = \max(1, q_1 + 1 - m), ..., q_1$. Then the word $u' := u_1...u_{m+q_1}$ has periods $q_1, q_1 + q_1, q_2 + q_1, ..., q_r + q_1$.*

Proof. Note that if $q_1 \leq m$, then $u_{m+1}, \ldots, u_{m+q_1}$ is the suffix of length q_1 of u. If $q_1 > m$, then u is a suffix $u_{m+1}, \ldots, u_{m+q_1}$. Clearly u' has period q_1 and therefore $2q_1$. For $t \leq m - q_j$ we have $u_t = u_{t+q_j} = u_{t+q_j+q_1}$ for $j = 2, ..., r$. □

Lemma 3. *If $p_j[k] > n[k]$ for some j, k, then $p_j[k] - n[k]$ equals the length $h[j]$ of $v[j]$.*

Proof. By definition the claim holds for $k = K$. Suppose the statement is correct for $k+1$. If $p_j[k] > n[k]$, then $p_j[k] > p[k]$ and $p_j[k+1] = p_j[k] - p[k], n[k+1] = n[k] - p[k]$. Thus the claim holds for k. If $p_j[k] \leq n[k]$, then $p_j[k-1] - n[k-1] \leq p_j[k] - n[k] \leq 0$. Hence there is a k_0 (possibly -1 or K) such that $p_j[k] - n[k] \leq 0$ for $k \leq k_0$ and $p_j[k] - n[k] = |v[j]|$ for $j > k_0$. □

Lemma 4. *The word w constructed according to the algorithm has length n and periods $p_1, ..., p_r$.*

Proof. By reverse induction on k. The word $w[K]$ has $N = n[K]$ letters. By (R1) and (R2) we know that $N \leq p_j[K]$ for each j with $p_j[K] > 0$. Hence $w[K]$ has periods $p_1[K], \ldots, p_r[K]$.

Suppose the constructed word $w[k+1]$ has length $n[k+1]$ and periods $p_1[k+1], p_2[k+1], \ldots, p_r[k+1]$. Assume first that $n[k] > 2n[k+1]$. Then $w[k] = w[k+1]v[g[k]]w[k+1]$. Since $n[k] - n[k+1] = p[k] = p_{g[k]}[k] = p_{g[k]}[k+1]$, we have, by Lemma 3, that $n[k] - 2n[k+1] = p_{g[k]}[k+1] - n[k+1]$ equals the length of $v[g[k]]$. Thus $w[k]$ has length $2n[k+1] + v[g[k]] = n[k]$. The word $w[k]$ has period $p[k] = p_{g[k]}[k]$ by construction. By applying Lemma 2 to $w[k+1]$ with $q_1 = p[k], m = n[k+1]$ and $\{q_1, \ldots, q_r\} = \{p_1[k+1], \ldots, p_r[k+1]\}$, we obtain that $w[k]$ has periods $p_j[k+1] + p[k] = p_j[k]$ for $j \neq g[k]$.

Suppose $n[k] \leq 2n[k+1]$. Then $n[k] - n[k+1] = p[k]$ is the length of $w'[k+1]$, hence by a similar application of Lemma 2 as in the previous case, $w[k]$ has periods $p_1[k], \ldots, p_r[k]$. Furthermore $w[k]$ has length $n[k+1] + p[k] = n[k]$.

Thus $w[0] = w$ has length $n[0] = n$ and periods $p_1[0] = p_1, \ldots, p_r[0] = p_r$. □

Lemma 5. *The word w constructed according to the algorithm has $n[K] + \sum_j' h[j]$ distinct letters, where the sum ranges over all j for which $j = g[k]$ for some k.*

Proof. The letters in w originate from $v[0]$ at the start and from the introduction of $v[j]$'s when $n[k] > 2n[k+1]$ and $j = g[k]$, in which case the letters of $v[j]$ are introduced. Hence w has at most $n[K] + \sum_j' h[j]$ distinct letters. Conversely, suppose that column j contains an underlined entry. If $p_j[K] = 0$, then column j does not introduce new letters. If $p_j[K] > 0$, then by the proof of Lemma 3 there exists a k_0 such that $p_j[k] - n[k] = |v[j]|$ for all $k > k_0$ and $p_j[k] - n[k] \leq 0$ for $k \leq k_0$. If $0 \leq k_0 < K$, then $j = g[k_0]$ and

$$n[k_0] - n[k_0+1] = p[k_0] = p_j[k_0] = p_j[k_0+1] > n[k_0+1].$$

Thus $n[k_0] > 2n[k_0+1]$ and the letters of $v[j]$ occur in w. If $k_0 = -1$ or $k_0 = K$, then the letters of $v[j]$ do not appear in w whereas $g[k]$ is not defined for these values of k. □

Lemma 6. *An FW word w can be generated as follows:*
$w[K+1] := \epsilon$; for $k \geq 0$
either $w[k] = w[k+1]v[g[k]]w[k+1]$ where $v[g[k]]$ is either the empty word, or consists of distinct letters none of which occur in $w[k+1]$,
or $w[k] = w[k+1]w'[k+1]$ where $w[k+1] = w[l+1]w'[k+1]$ for some $l > k$.

Proof. By reverse induction on k. By definition $w[K] = v[0]$ which consists of N distinct letters. Let $0 \leq k < K$. If $n[k] > 2n[k+1]$, then $w[k] = w[k+1]v[g[k]]w[k+$

1]. We have $n[k] = n[k+1] + p[k] \geq p[k] = p_{g[k]}[k]$. Hence $n[k] \geq p_{[g[k]]}[k]$ and by induction it follows that $n[k-1] \geq p_{[g[k]]}[k-1], \dots, n[0] \geq p_{[g[k]]}[0]$ so that the letters of $v[g[k]]$ are introduced only once. Thus the $h[g[k]]$ letters introduced at level k do not occur in $w[k+1]$.
If $n[k] \leq 2n[k+1]$, then $w[k] = w[k+1]w'[k+1]$ where $w'[k+1]$ is the suffix of $w[k]$ of length $p[k]$. Hence the prefix of $w[k+1]$ which is deleted has length

$$n[k+1] - p[k] = n[k+1] - p_{g[k]}[k+1] = n[k+2] - p_{g[k]}[k+2] = \cdots = n[l] - p_{g[k]}[l]$$

where l is the first level after k where $g[l] = g[k]$, hence $p_{g[k]}[l] = p[l]$. Therefore $n[k+1] - p[k] = n[l+1]$. It follows that the prefix of $w[k+1]$ which is omitted in $w'[k+1]$ equals $w[l+1]$. Thus $w[k+1] = w[l+1]w'[k+1]$. $\square$

Proof of Theorem 1. Let $P = \{p_1, p_2, \dots, p_r\}$ be a period set, and n a positive integer. By Lemma 4 the algorithm generates a word of length n with periods $p_1, \dots, p_r$. Let u be a word of length n having periods $p_1, \dots, p_r$ with $\sharp u \geq \sharp w$. It remains to prove that u is isomorphic with w. Consider the prefix $u[1]$ of u of length $n[1]$. According to Lemma 1 applied to u with $m = n, q_1 = p[0], \{q_1, \dots, q_r\} = \{p_1, \dots, p_r\}$ we find that if $n = n[0] \leq 2n[1]$, then $n \geq 2(n[0] - n[1]) = 2p[0]$, hence $u[1]$ is composed of the same letters as u, whereas $w[1]$ is composed of the same letters as w. If $n > 2n[1]$, then $p_{g[0]}[1] = p_{g[0]}[0] = p[0] = n[0] - n[1] > n[1]$, hence, by Lemma 3, $p[0] - n[1] = h[g[0]] > 0$ which implies $n = n[1] + p[0] = 2p[0] - h[g[0]]$ with $h[g[0]] > 0$ from which it follows that $\sharp u - \sharp u[1] \leq h[g[0]]$ and $\sharp w - \sharp w[1] = h[g[0]]$. Thus, in both cases, $\sharp u - \sharp u[1] \leq \sharp w - \sharp w[1]$. We proceed by induction on k and find for every k that $\sharp u[k] - \sharp u[k+1] \leq \sharp w[k] - \sharp w[k+1]$. Thus $\sharp u - \sharp u[K] \leq \sharp w - \sharp w[K]$. Furthermore, $\sharp u[K] \leq |u(K)| = N = |w(K)| = \sharp w[K]$. We conclude that $\sharp u \leq \sharp w$, hence by our assumption $\sharp u = \sharp w$. It follows that $\sharp u[k] - \sharp u[k+1] = \sharp w[k] - \sharp w[k+1]$ for $k = 0, 1, \dots, K-1$ and that $\sharp u[K] = \sharp w[K]$. Thus $u[K]$ consists of exactly $\sharp w[K] = N$ distinct letters. Moreover, if $n[k] > 2n[k+1]$, then $\sharp u[k] - \sharp u[k+1] = h[g[k]] = n[k] - 2n[k+1], |u[k]| = n[k], |u[k+1]| = n[k+1]$. Since $u[k]$ has period $p_{g[k]} = n[k+1] + h[g[k]]$, it is of the form $u[k+1]t[k+1]u[k+1]$ where $t[k+1]$ is a word of length $h[g[k]]$. Hence $t[k+1]$ consists of $h[g[k]]$ distinct new letters and is therefore isomorphic with $v[k+1]$. If $n[k] \leq 2n[k+1]$, we put $t[k+1] = \epsilon$. We conclude that $t[1] \dots t[K]$ consist of distinct letters, $\sum'_j h[j]$ in total. Since $u[k] = u[k+1]t[k+1]u[k+1]$ if $n[k] > 2n[k+1]$ and $u[k] = u[k+1]u'[k+1]$ with $|u'[k+1]| = n[k] - n[k+1] \leq n[k+1]$ otherwise, we find that u is isomorphic with w indeed.

Let ϕ be the involutory antimorphism induced by $\phi(v[i]) = \underline{v[i]}$ where overlining denotes reversion. According to Lemma 6 the word w is constructed inductively by extensions of the form $w \mapsto wuw$ and $w \mapsto wv$ where in the former case $\phi(w) = \overline{w}, \phi(u) = \overline{u}$ and in the latter case $W = uv$ with $\phi(w) = \overline{w}, \phi(u) = \overline{u}$. In the former case we have

$$\phi(wuw) = \phi(w)\phi(u)\phi(w) = \overline{w} \cdot \overline{u} \cdot \overline{w} = \overline{wuw},$$

in the latter case

$$\phi(wv) = \phi(w)\phi(v) = \overline{w}\phi(v) = \overline{v} \cdot \overline{u}\phi(v) = \overline{v}\phi(u)\phi(v) = \overline{v}\phi(uv) = \overline{v}\phi(w) = \overline{v} \cdot \overline{w} = \overline{wv}.$$

So it follows by induction that the final FW word is pseudo-palindromic with respect to ϕ. $\square$

5. Some more examples

In this section we consider more examples which reveal variations in the outcome of the algorithm and structure of FW words.

Example 2. We construct the FW word of length $n = 200$ relative to the period set $P = \{99, 147, 174, 188, 198, 207\}$. We set $p_1 = 99, p_2 = 147, p_3 = 174, p_4 = 188, p_5 = 198, p_6 = 207$. This yields the following table.

k	$p_1[k]$	$p_2[k]$	$p_3[k]$	$p_4[k]$	$p_5[k]$	$p_6[k]$	$n[k]$	$g[k]$
0	$\underline{99}$	147	174	188	198	207	200	1
1	99	$\underline{48}$	75	89	99	108	101	2
2	51	48	$\underline{27}$	41	51	60	53	3
3	24	21	27	$\underline{14}$	24	33	26	4
4	10	$\underline{7}$	13	14	10	19	12	2
5	$\underline{3}$	7	6	7	3	12	5	1
6	3	4	3	4	0	9	2	

Hence up to isomorphism the FW word w is a concatenation of $v[0] = xy, v[1] = a, v[2] = bc, v[3] = d, v[4] = ef$. The word w begins in

$$xy|axy|bcxyaxy|efxyaxybcxyaxy|dxyaxybcxyaxyefxyaxybcxyaxy|$$

$$bcxyaxyefxyaxybcxyaxydxyaxybcxyaxyefxyaxybcxyaxy|axy$$

Note that the periods 198 and 207 have no underlined entries in their corresponding columns. Hence these periods are a consequence of the other 'essential' periods 99, 147, 174, 188 and the length of the word: Of course, every word of length 200 has period 207, and period 198 is induced by period 99. The choice $p_1 < p_2 < \ldots, p_r$ is not essential, but it secures that only significant columns get underlined numbers. Otherwise the 3 in the column of $p_5 = 198$ may have been underlined and the fact that 198 is an 'insignificant' period would have been hidden.

If we would have chosen $n = 204$ instead, all the numbers in the column of n would have been increased by 4 and the table would have been continued as follows:

k	$p_1[k]$	$p_2[k]$	$p_3[k]$	$p_4[k]$	$p_5[k]$	$p_6[k]$	$n[k]$	$g[k]$
6	$\underline{3}$	4	3	4	0	9	6	1
7	3	$\underline{1}$	0	1	0	6	3	2
8	2	$\underline{1}$	0	0	0	5	2	2
9	$\underline{1}$	1	0	0	0	4	1	1
10	1	0	0	0	0	3	0	1

Hence $\sharp w = 1, v[1] = a$, say, and the FW word w is the constant word of length 204. The extremal FW word relative to the period set P has length 202 and has 2 distinct letters.

Example 3. Let $n = 651$, and $P = \{325, 485, 561, 603, 624\}$. We set $p_1 = 325, p_2 = 485, p_3 = 561, p_4 = 603, p_5 = 625$. We obtain the following table:

k	$p_1[k]$	$p_2[k]$	$p_3[k]$	$p_4[k]$	$p_5[k]$	$n[k]$	$g[k]$
0	$\underline{325}$	485	561	603	624	651	1
1	325	$\underline{160}$	236	278	299	326	2
2	165	160	$\underline{76}$	118	139	166	3
3	89	84	76	$\underline{42}$	63	90	4
4	47	42	34	42	$\underline{21}$	48	5
5	26	21	$\underline{13}$	21	21	27	3
6	13	$\underline{8}$	13	8	8	14	2
7	$\underline{5}$	8	5	0	0	6	1
8	5	3	0	0	0	1	

Hence $\sharp w = 7, v[0] = x, v[1] = abcd, v[2] = ef$. The FW word is given by

$$w = x|abcdx|efxabcdx|abcdxefxabcdx|efxabcdxabcdxefxabcdx|$$
$$(efxabcdxabcdxefxabcdx)^2|abcdxefxabcdx$$
$$(efxabcdxabcdxefxabcdx)^3|((efxabcdxabcdxefxabcdx)^4$$
$$abcdxefxabcdx((efxabcdxabcdxefxabcdx)^3|$$

once repeated without the first x.

In this example all periods are significant, although the columns corresponding to the periods $561, 603, 624$ do not contribute new symbols to w. In fact the period 561 joins the period 325 at level 7. The choice to proceed with column 1 and not with column 3 is merely the authors' choice and not mathematically prescribed. Similarly the periods 603 and 624 join 485 at level 6. The combinations are reflected in the word w. The word $w'[5]$ starts with a because of the underlined number 13 in column 3 of period 561 which joins period 325 in column 1 whereas $v[1]$ starts with a, the word $w'[4]$ starts with e because of the underlined number 21 in column 5, the word $w'[3]$ starts also with e because of the underlined number 42 in column 4, $w'[2]$ starts with a because of the underlined number 76 in column 3, $w'[1]$ starts with e because of the underlined number 160 in column 2, and $[w'[0]$ starts with a because of the underlined number 325 in column 1.

Example 4. The following examples show that different sets of periods can generate the same FW word. Let $n = 17$ and $P = \{3, 16\}$. The resulting FW word is *aabaabaabaabaabaa*. The same FW word is generated by the choice $n = 17$ and $P = \{6, 9, 16\}$. It is an open question whether for a given FW word w the set of periods with minimal cardinality generating w is unique.

Acknowledgement The authors would like to thank Danielle Heckman for many interesting discussions concerning the material in this paper.

References

[AZZ] V. Anne, L.Q. Zamboni, I. Zorca, *Palindromes and pseudo-palindromes in episturmian and pseudo-palindromic infinite words*, Words 2005, Publications du LaCIM (S. Brlek, C. Reutenauer eds.) **36.** (2005), 91-100.

[BLLZ1] M. Bucci, A. de Luca, A. De Luca, L.Q. Zamboni, *On some problems related to palindrome closure*, Theor. Inform. Appl. (RAIRO) (2007), to appear.

[BLLZ2] M. Bucci, A. de Luca, A. De Luca, L.Q. Zamboni, *On different generalizations of episturmian words* (2007), preprint.

[CMR] M.G. Castelli, F. Mignosi, A. Restivo, *Fine and Wilf's theorem for three periods and a generalization of Sturmian words*, Theor. Computer Sc. **218** (1999), 83-94.

[CI] S. Constantinescu, L. Ilie, *Generalised Fine and Wilf's theorem for arbitrary number of periods*, Theor. Computer Sci. **339** (2005), 49-60.

[LL] A. de Luca, A. De Luca, *Pseudopalindrome closure operators in free monoids*, Theor. Comput. Sc. **362** (2006), 282-300.

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

[H] Š. Holub, *On multiperiodic words*, Theor. Inform. Appl. **40** (2006), 583-591.

[J] J. Justin, *On a paper by Castelli, Mignosi, Restivo*, RAIRO: Theoret. Informatics Appl. **34** (2000), 373-377.

[L1] M. Lothaire (ed.), *Combinatorics on Words*, Addison Wesley, Reading MA, 1983.

[L2] M. Lothaire (ed.), *Algebraic Combinatorics on Words*, Cambridge University Press, 2002.

[TZ] R. Tijdeman, L. Zamboni, *Fine and Wilf words for any periods*, Indag. Math. **N.S. 14** (2003), 135-147.

R. Tijdeman

Leiden University
Mathematical Institute
Postbus 9512
2300 RA Leiden
The Netherlands

L.Q. Zamboni

Department of Mathematics
P O Box 311430
University of North Texas
Denton, TX 76203-1430
U.S.A.

E-mail address:
tijdeman@math.leidenuniv.nl
luca@unt.edu