FINE AND WILF WORDS FOR ANY PERIODS II

R. Tijdeman and L.Q. Zamboni

Abstract. Given positive integers n, and $p_1, \dots, p_r$, we define a fast word combinatorial algorithm for constructing a word $w = w_1 \cdots w_n$ of length n, with periods $p_1, \dots, p_r$, and on the maximal number of distinct letters. We show that the constructed word, which is unique up to word isomorphism, is a pseudo-palindrome, i.e., is a fixed point of an involutory antimorphism.

1. Introduction

Let $w = w_1 \cdots w_n$ 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 $i = 1, \dots, n-p$. 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 $\gcd(p_1, p_2)$ 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 $\gcd(p_1, p_2)$. 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 (cf. [L2] Ch. 2). 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 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 $P = \{p_1, p_2, \dots, p_r\}$ be a set consisting of r distinct positive integers. We call P a *period set* if $\gcd(p_1, p_2, \dots, p_r) \notin \{p_1, p_2, \dots, p_r\}$. 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 $p_1, p_2, \dots, p_r$ also has period $\gcd(p_1, p_2, \dots, p_r)$.

In 2003 the authors [TZ] introduced a fast algorithm for constructing extremal Fine and Wilf words relative to any period set $P = \{p_1, p_2, \dots, p_r\}$, i.e., words of length $L(P) - 1$ having periods $p_1, p_2, \dots, p_r$, but not period $\gcd(p_1, p_2, \dots, p_r)$, and on the largest possible number of distinct symbols. This same construction was later described by Constantinescu and Ilie [CI] and by Š. Holub [H]. There are several reasons for requiring these words to be on the largest possible number of symbols. First of all, it is shown in [TZ] that such words are unique up to word isomorphism. Secondly, any other word of the same length having periods $p_1, p_2, \dots, p_r$ is the morphic image (under a letter-to-letter morphism) of the

2000 Mathematics Subject Classification: primary 68R15, secondary 11B85

extremal Fine and Wilf word. The authors also showed that as in the case of two periods, if $\gcd(p_1, \ldots, p_r) = 1$, then the non-constant word of maximal length having periods $\{p_1, \ldots, p_r\}$ and on the largest possible number of symbols, is a palindrome.

Let w be a finite word on a non-empty alphabet A. We define $\sharp w$ 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 $\sharp w \geq \sharp v$. We note that if w is a FW word relative to a set P, and $|w| < L(P)$, then $\gcd(p_1, p_2, \ldots, p_r)$ is not a period. This follows directly from the definition of $L(P)$ and the fact that w is assumed to be on the largest possible number of symbols. We call a word w a FW word if w is a FW word relative to some set P.

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. 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 $\gcd(8, 11)$. 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 a composition $\theta = \tau \circ R = R \circ \tau$ where τ is an *involutory permutation* of the alphabet A, i.e., τ satisfies $\tau^2(a) = a$ for all $a \in A$.

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 related to the one originally given in [TZ].

In Section 2 we describe a naive approach in the construction of FW words. In Section 3 we present the algorithm that 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. For further properties of (extremal) FW words see [TZ1,TZ2].

2. A naive description of FW words

In this section we describe a simple but inefficient method for constructing FW words of a given length relative to a given set. 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 $\{1, 2, \dots, n\}$, and for each $p \in P$, we put an undirected edge labeled p between vertices x and y if and only if $|x - y| = p$. 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, when $P = \{8, 11\}$ 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 mentioned in the Introduction. 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}. Hence w' is the extremal FW word for period set $\{8, 11\}$.

3. The algorithm

Let $P = \{p_1, \dots, p_r\}$ be a period set, and n a positive integer. We now describe a multi-dimensional generalization of the Euclidean algorithm which we use to efficiently construct a FW word w of length n relative to the period set P. This algorithm, which we call Algorithm A, reveals the structure of the constructed FW word. This structure is essential for the proof that the word is unique up to word isomorphism and is a pseudo-palindrome.

ALGORITHM A
Input: positive integers $n, p_1, \dots, p_r$.
Reduction
(R0) (Initialization) For $i = 1$ to r, put
  $p_i[0] := p_i$;
  $k := 0$;
  $n[0] := n$.
(R1) Let i be the smallest index 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
  $p[k] := p_i[k]$;
else
  goto (E0).
(R3) For $j \in \{1, \dots, r\}$ different from i, 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 $j = 1$ to r
if $p_j[K] > N$ then
  $h[j] := p_j[K] - N;$
  $v[j] := v_{j1}v_{j2} \dots v_{jh[j]};$
else
  $h[j] := 0;$
  $v[j] := \epsilon.$
(E1) For $k = K - 1$ down to 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 $w := w[0]$.

As in [TZ], we can organize the values $p_i[k]$ (for $1 \le i \le r$), $n[k]$ and $g[k]$ arising in the Reductive part of Algorithm A in a rectangular array or *tower* as illustrated in Example 1 below. We number the row by k and the column of p_j by j. The value $p[k]$ in row k is called a *pivot* and is underlined. Observe that only those factors $v[j]$ appear in w for which $j = g[k]$ for some k, i.e., column j contains a pivot. 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. Algorithm A constructs a FW word w of length n relative to the period set P. The word w is unique up to word isomorphism. It is a pseudo-palindrome with respect to the involutary antimorphism ϕ which fixes each 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 89 relative to the period set $P = \{45, 66, 75, 85\}$. We label the periods: $p_1 = 45, p_2 = 66, p_3 = 79, p_4 = 85$. 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 row j, we identify the first 'period' column i (counting from the left) containing the smallest positive p-value and mark it by underlining $p[j]$. In passing from row j to row $j + 1$, 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 row j to row $j + 1$. Any 0 in a period column at row j stays 0 at row $j + 1$. The algorithm terminates when the current n-value would no longer be positive when we would continue. In Example 1 we stop at row 5, since $p[5] = p_2[5] = 2 \ge 2 = n[5]$ so that otherwise $n[6] = n[5] - p[5] = 0$.

Applying Algorithm A we consecutively obtain: $n[0] = 89, p[0] = 45, g[0] = 1, n[1] = 44, p[1] = 21, g[1] = 2, n[2] = 23, p[2] = 13, g[2] = 3, n[3] = 10, p[3] = 6, g[3] = 4, n[4] = 4, p[4] = 2, g[4] = 2, n[5] = 2, K = 5, N = 2$.
Next we construct words $v[0], v[1], \dots, v[4]$ consisting of $N = 2, h[1] = 1, h[2] = 0, h[3] = 3, h[4] = 2$ distinct letters, respectively:
$v[0] := xy, v[1] := a, v[2] := \epsilon, v[3] := bcd, v[4] := ef$, say.

Then we construct the FW word in six steps where we concatenate $w[k+1]$ with another word to obtain $w[k]$ so that we reach the length $n[k]$, for $k = 5, 4, 3, 2, 1$:

k	$p_1[k]$	$p_2[k]$	$p_3[k]$	$p_4[k]$	$n[k]$	$g[k]$
0	$\underline{45}$	66	79	85	89	1
1	45	$\underline{21}$	34	40	44	2
2	24	21	$\underline{13}$	19	23	3
3	11	8	13	$\underline{6}$	10	4
4	5	$\underline{2}$	7	6	4	2
5	3	2	5	4	2	

Table 1: Result of the Reduction procedure

$w[5] := v[5] = xy$ (of length 2)
$w[4] := w[5]v[g[4]]w[5] = xyxy$ (of length 4)
$w[3] := w[4]v[g[3]]w[4] = xyxyefxyxy$ (of length 10)
$w[2] := w[3]v[g[2]]w[3] = xyxyefxyxybcdxyxyefxyxy$ (of length 23)
$w[1] := w[2]v[g[4]]w[5]v[g[3]]w[4]v[g[2]]w[3] = xy(xyefxyxybcdxyxyefxyxy)^2$ (of length 44)
$w[0] := w[1]v[g[0]]w[1] = w[1]aw[1]$ (of length 89).

Note that $g[4] = g[1]$ and that the suffix $w'[1]$ of $w[1]$ is exactly the subword which was concatenated after the previous time that column $g[4]$ was chosen for underlining. We conclude that $\sharp w$ equals $2 + 1 + 0 + 3 + 2 = 8$. Observe that w is a pseudo-palindrome with respect to the function ϕ defined by $\phi(x) = y, \phi(y) = x, \phi(a) = a, \phi(b) = d, \phi(c) = c, \phi(d) = b, \phi(e) = f, \phi(f) = e$, hence $\phi(xy) = xy, \phi(a) = a, \phi(bcd) = bcd, \phi(ef) = ef$.

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 $u = u_1 \cdots u_m$ be a word with s distinct letters and periods $q_1 < \cdots < q_r$. Put $u' := u_1 \cdots u_{m-q_1}$. 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 $q_1, q_2 - q_1, \ldots, q_r - q_1$.*

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 $j = 2, \ldots, r$. □

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

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

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

Lemma 3. *There is a k_0 (with $-1 \leq k \leq K$) such that $p_j[k] - n[k] \leq 0$ for $k \leq k_0$ and $p_j[k] - n[k] = |v[j]|$ for $j > k_0$.*

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 $p_j[k] \le n[k]$, then $p_j[k-1]-n[k-1] \le p_j[k]-n[k] \le 0$. The result follows by induction. □

Lemma 4. *The word w constructed according to Algorithm A has length n and periods $p_1, \dots, p_r$.*

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

Suppose the constructed word $w[k+1]$ has length $n[k+1]$ and periods $p_1[k+1], p_2[k+1], \dots, p_r[k+1]$. 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 $\{q_1, \dots, q_r\} = \{p_1[k+1], \dots, p_r[k+1]\}$, we obtain that $w[k]$ has periods $p_j[k+1]+p[k] = p_j[k]$ for $j \ne g[k]$.

Suppose $n[k] \le 2n[k+1]$. 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 $p_1[k], \dots, p_r[k]$. Furthermore $w[k]$ has length $n[k+1]+p[k] = n[k]$.

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

Lemma 5. *The word w constructed according to Algorithm A 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 $p_j[k]-n[k] \le 0$ for $k \le k_0$. If $0 \le k_0 < K$, 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 and $g[k]$ is not defined for these values of k. □

Lemma 6. *A FW word w can be generated as follows:*
$w[K+1] := \epsilon$; for $k \ge 0$
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. We proceed by induction on $K-k$. By definition $w[K] = v[0]$ which consists of N distinct letters. Let $0 \le k < K$. If $n[k] > 2n[k+1]$, then

$$w[k] = w[k+1]v[g[k]]w[k+1].$$

We have

$$n[k] = n[k+1]+p[k] \ge p[k] = p_{g[k]}[k].$$

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 row 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 omitted in $w'[k+1]$ has non-negative 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 row after k where $g[l] = g[k]$, hence $p_{g[k]}[l] = p[l]$. Therefore $n[k+1] - p[k] = n[l] - p[l] = n[l+1]$. It follows that the prefix of $w[k+1]$ which is omitted in $w'[k+1]$ equals $w[l+1]$. Thus $w[k+1] = w[l+1]w'[k+1]$. □

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, Algorithm A 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 $t[1] \cdots t[K]$ 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 which fixes each v_i. According to Lemma 6 the word w is constructed inductively by extensions of the form $w \mapsto wuw$ and $w \mapsto wv$, with $w = uv$. Suppose that in both cases $\phi(w) = w, \phi(u) = u$. In the former case we obtain

$$\phi(wuw) = \phi(w)\phi(u)\phi(w) = wuw,$$

in the latter case

$$\phi(wv) = \phi(v)\phi(w) = \phi(v)w = \phi(v)uv = \phi(v)\phi(u)v == \phi(uv)v = \phi(w)v = wv.$$

So it follows by induction that the final FW word is a pseudo-palindrome with respect to ϕ. □

5. Some more examples

In this section we consider more examples which reveal variations in the outcome of Algorithm A and the 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 word isomorphism the FW word w is a concatenation of subwords $v[0] = xy, v[1] = a, v[2] = bc, v[3] = d, v[4] = ef$. The word w begins with

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

$$bcxyaxyefxyaxybcxyaxydxyaxybcxyaxyefxyaxybcxyaxy|axy$$

Note that the periods 198 and 207 have no underlined entries in their corresponding columns. Of course, every word of length 200 has period 207, and period 198 is induced by period 99. Hence these periods are a consequence of the other periods 99, 147, 174, 188 and the length of the word. If we omit one of the remaining periods, then another FW word results. The choice $p_1 < p_2 < \cdots p_r$ is not essential for finding the FW word w, but it ensures that only significant columns get underlined numbers. Otherwise the 3 in column 5 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. This happens when two columns join, that is reach the same positive value at the same row. In the present case the column 3 joins the column 1 at row 7. The choice to proceed with column 1 and not with column 3 is merely the authors' choice and not mathematically prescribed. Similarly the columns 2, 4 and 5 join at row 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 which joins column 1 whereas $v[1]$ starts with a. Similarly, the word $w'[4]$ starts with e because of the underlined number 21 in column 5 which joins column 2 at row 6 whereas $v[2]$ starts with e, 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.

Remark. Obviously, if P is a period set for which Algorithm A generates the word w and R is the set of all periods of w of length $\leq |w|$, then every set Q with $P \subset Q \subset R$ is also a period set of w. The question is therefore: which structure does the minimal period set of w have? But the initial idea of the authors that the period set of minimal cardinality is uniquely determined, is wrong: The extremal FW word of length 16, $w = aaaabaaaaaabaaaa$ has period sets $\{7, 12, 13\}$ and $\{7, 12, 15\}$, but, by the Theorem of Fine and Wilf, it has no period set of two elements.

Acknowledgement. The authors 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) (2008), in press and available on line at http://dx.doi.org/10.1051/ita:2007064..

[BLLZ2] M. Bucci, A. de Luca, A. De Luca, L.Q. Zamboni, *On different generalizations of episturmian words*, Theoret. Comput. Sci. **393** (2008), 23–36.

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

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

[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.

[LL] A. de Luca, A. De Luca, *Pseudopalindrome closure operators in free monoids*, Theoret. Comput. Sci. **362** (2006), 282–300.

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

[TZ1] R. Tijdeman, L. Zamboni, *A period free version of the Fine and Wilf theorem*, to appear.

[TZ2] R. Tijdeman, L. Zamboni, *Characterizations of words with many periods*, to appear.

R. Tijdeman

Leiden University
Mathematical Institute
Postbus 9512
2300 RA Leiden
The Netherlands

L.Q. Zamboni

Université de Lyon
Université Lyon 1
CNRS UMR 5208 Institut Camille Jordan
43, blvd du 11 novembre 1918
F-69622 Villeurbanne Cedex, France.

Reykjavik University
School of Computer Science
Kringlan 1
103 Reykjavik, Iceland

E-mail address:
`tijdeman@math.leidenuniv.nl`
`zamboni@math.univ-lyon1.fr`
`lqz@ru.is`